AGENT-BASED MODELS AND ECONOMIC POLICY

edited by Jean-Luc Gaffard and Mauro Napoletano

AGENT-BASED MODELS AND ECONOMIC POLICY

ofce

Revue de l'OFCE / *Debates and policies*

OFCE

L'Observatoire français des conjonctures économiques est un organisme indépendant de prévision, de recherche et d'évaluation des politiques publiques. Créé par une convention passée entre l'État et la Fondation nationale des sciences politiques approuvée par le décret n° 81.175 du 11 février 1981, l'OFCE regroupe plus de 40 chercheurs français et étrangers, auxquels s'associent plusieurs *Research fellows* de renommée internationale (dont trois prix Nobel). « Mettre au service du débat public en économie les fruits de la rigueur scientifique et de l'indépendance universitaire », telle est la mission que l'OFCE remplit en conduisant des travaux théoriques et empiriques, en participant aux réseaux scientifiques internationaux, en assurant une présence régulière dans les médias et en coopérant étroitement avec les pouvoirs publics français et européens. Philippe Weil préside l'OFCE depuis 2011, à la suite de Jean-Paul Fitoussi, qui a succédé en 1989 au fondateur de l'OFCE, Jean-Marcel Jeanneney. Le président de l'OFCE est assisté d'un conseil scientifique qui délibère sur l'orientation de ses travaux et l'utilisation des moyens.

Contact
OFCE, 69 quai d'Orsay 75340 Paris cedex 07
Tel. : +33(0)1 44 18 54 87
mail : revue@ofce.sciences-po.fr
web : www.ofce.sciences-po.fr

Dépôt légal : octobre 2012 ISBN : 978-2-312-00316-0
N° ISSN 1265-9576 – ISSN en ligne 1777-5647 –

Index

AGENT-BASED MODELS AND ECONOMIC POLICY

edited by Jean-Luc Gaffard and Mauro Napoletano

IMPROVING THE TOOLBOX
NEW ADVANCES IN AGENT-BASED AND COMPUTATIONAL MODELS

Jean-Luc Gaffard and Mauro Napoletano

Are current economic models well equipped to provide useful policy prescriptions? Many economists would have certainly answered, "yes" before the recent Global Recession. This economic crisis has not only demonstrated the importance of banking and financial markets for the dynamics of real economies. It has also revealed the inadequacy of the dominant theoretical framework. Standard models have indeed failed to forecast the advent of the crisis. In addition, they have been unable to indicate a therapy able to restore economic growth.

Since the onset of the crisis, the discontent towards the dominant approach to economic modeling has flourished.[1] Criticism has been mainly directed towards the over-simplicity of standard models in general, and of Dynamic Stochastic General Equilibrium Models (DSGEs) in particular. Most features that have played a key role in generating the crisis, such as heterogeneity of agents, markets, and regulatory frameworks, financial innovation, securitization, are by and large overlooked in standard macro-models. A second kind of dissatisfaction is related to the hyper-rationality of individuals. Markets (and financial markets in particular) are plenty of people acting on the basis of overconfidence, heterogeneous beliefs, imperfect knowledge of the states of the world, and of the consequence of humans' actions, etc. These features are not present in standard macro models, which build on the assumption of a representative individual

1. Interestingly, this time critiques have not only come from "heterodox" schools of thought. Critiques have also been raised by scholars who made a significant use of the ingredients of standard models in the past (see e.g. Caballero, 2010, Krugman, 2009, Stiglitz, 2011) as well as by leading policy-making authorities (see e.g. Trichet, 2010).

knowing all the characteristics of the economy and able to replicate whatever human intelligence can do (Leijonhufvud, 1993). A third concern is the assumption of equilibrium. Standard models typically focus on states of the economy in which all markets clear. In contrast, the crisis has shown the possibility of situations in which some markets (and the market for labor in particular) do not clear. Standard models ignore the problems that would result from reactions of agents to such market disequilibria. They are therefore badly equipped to study how the economy behaves during crises.

A natural way to follow in face of the problems exposed in the previous section would be departing from the representative agent paradigm, thereby introducing heterogeneity of agents' characteristics and behavior, and allowing for markets that do not clear. All the aforementioned characteristics add new degrees of complexity to macroeconomic analysis. As eloquently expressed by Tesfatsion (2006):

> *"The modeler must now come to grips with challenging issues such as asymmetric information, strategic interaction, expectation formation on the basis of limited information, mutual learning, social norms, transaction costs, externalities, market power, predation, collusion, and the possibility of coordination failures."*

Exploiting the growing capabilities of computers, Agent-Based Models (ABMs) analyze economic processes as dynamical systems of heterogeneous interacting agents (Epstein and Axtell, 1996; Tesfatsion, 2006). In ABMs repeated interactions among agents over time induce continuously changing microeconomic patterns, the aggregation of which generates a dynamics for the macroeconomic variable of interest (Pyka and Fagiolo, 2007).

This special issue gathers contributions of leading scholars in Agent-Based and computational economics and shows the applicability of this methodology to several issues both in micro- and in macroeconomics. This introduction aims to provide some guidelines to the different contributions in the issue by organizing them around 4 main themes.

1. The critique to standard economic models and the structure of Agent-Based Models

As we mentioned above standard economic models (and DSGEs in particular) are badly equipped to analyze some key issues that emerged in the last economic crisis. The contributions of **Kirman** (2012), **Ashraf, Gershman and Howitt** (2012) and of **Fagiolo and Roventini** (2012) provide a clear and thorough account of the critiques to mainstream macroeconomic models. Kirman (2012) discusses the historical evolution of modern macro, and shows how DSGEs are the final outcome the particular path followed by mainstream economics in the last century. The task of the economist in this tradition (and of the macroeconomist in particular) is to make assumptions about individual preferences and technologies in line with some axioms that are characteristic of the "rational" agent and to build models on that basis. Next, one finds the equilibrium of the system and examines the characteristics of such equilibrium states. In contrast, ABMs allow one to dispense with the restrictive assumptions of standard models and put at the center of the analysis the heterogeneity of economic agents and the evolution of the network of interactions among them. The article of Fagiolo and Roventini (2012) discusses in details the building blocks of Agent-Based Models.

The analysis of the mechanisms that govern (and coordinate) economic interactions in a decentralized economy is essential not only for understanding the generation of crises. It also helpful to provide useful policy insights beyond those provides by standard models, and to analyze the effects of technical change in the economy. The first point is vividly stressed in Ashraf, Gershman and Howitt (2012) and Fagiolo and Roventini (2012), whereas the second one is extensively discussed in the paper of **Amendola, Gaffard and Saraceno** (2012).

2. Analytical vs. simulation methods in ABMs

Agent-Based Models are typically more complex than standard models, and this implies that one must often employ computer simulations for their analysis. The contribution of Fagiolo and Roventini (2012) discusses the different problems that emerge in the use of simulation techniques for the analysis of ABMs, and the different solutions that have so far been proposed in the literature.

One critique against ABMs is that the use of computer simulations and the complicated structure of ABMs often prevent one from detecting the mechanisms generating results in the model. In our opinion, the fact that an overly complicated structure may blur causal mechanisms is a quite general remark applicable to any model, rather than a specific and unavoidable fallacy of ABMs (see also Napoletano, Gaffard and Babutsidze, 2012, for more discussion). However, even in very complicated ABMs—causal mechanisms can be detected through counterfactual simulation analyses. Indeed, ABMs often allow one to control the presence of some dynamics in the model through an appropriate setting of the parameters, and to test how results are different when such dynamics are switched off/on. This technique is widely exploited by most papers in this issue. Furthermore, one should consider computer simulations and analytical results as possible complements rather than substitutes. This last methodological point emerges with great clarity in the contributions of **Delli Gatti *et al.*** (2012) and of **Babutsidze** (2012). More precisely, the paper of Delli Gatti et al shows how the interactions among heterogeneous agents can be successfully modeled by employing master equations techniques. This analytical investigation is then coupled with simulation techniques that show the reaction of the systems to different microeconomic shocks. On related grounds, the paper of Babutsidze studies analytically the properties of the equilibrium of an economy populated by heterogeneous agents using asymmetric (S,s) rules to set prices, and then uses simulation techniques to analyze the far-from-the-equilibrium dynamics.

One of the strongest critiques to DSGE models has concerned their failure to forecast the advent of the crisis (Wieland and Wolters, 2012, summarize the debate on this issue). At the same time, forecasting appears difficult in ABMs due to the inherent non-linearities implied by the complex interactions among heterogeneous agents (see also Dawid and Neugart (2011), and Grazzini *et al.* (2012), for more discussions). **Barde** (2012) makes an important step in overcoming the hurdles involved in forecasting exercises with ABMs, by showing the possibility of successfully applying Maximum Entropy techniques to predict model outcomes in the Agent-Based models of Kirman (1993) and of Abrams and Strogatz (2003).

3. Agents' heterogeneity, micro- and macroeconomic dynamics

The recent crisis has shown that *distributions matters* (see also Stiglitz, 2011). One instance of this is the market for credit, where the distribution of information between borrowers and lenders plays a key role. Asymmetric information, credit contracts, and the possible bankruptcy of agents are all elements that significantly affect aggregate dynamics in the papers of **Cincotti, Raberto and Teglio** (2012), **Ashraf, Gershman and Howitt** (2012), **Delli Gatti *et al.*** (2012) in this special issue. Furthermore, the crisis and the associated surge in unemployment also generated sharp inequalities within the population: some individuals have seen their income falling either because they got unemployed or because of falling wages in a situation of depressed labor demand. Reduced incomes by a significant fraction of the population would normally lead to a fall in consumption demand. The interplay between inequality and demand and its consequences for both the micro and the macro-economic dynamics are central the central themes of the papers of **Napoletano *et al.*** (2012) and of **Patriarca and Vona** (2012). The former paper shows that, independently from the investment behavior of firms, steady growth of the economy requires a balance in the distribution between profits and wages. The paper of Patriarca and Vona studies instead the relations between income inequality and the effectiveness of policies aimed at introducing green technologies and shows that when income inequality is taken into account carbon taxes may have the paradoxical effect of reducing (rather than increasing) the number of consumers of green products.

Heterogeneity among agents in ABMs does not only relate to agents' asymmetries in either characteristics or in behavior. It may also involve differences in production processes over time. Accordingly, ABMs are suitable for analyzing the interplays between aggregate long-run growth and the processes of structural transformations and qualitative change in the economy. The papers of **Ciarli** (2012), **Amendola, Gaffard and Saraceno** (2012) and **Saviotti and Pyka** (2012) are devoted to this crucial issue.

4. Policy analyses under different institutional scenarios

ABMs are not only models that explicitly account for agents' heterogeneity. They have another advantage that is more related to policy design, and is represented by their finer description of the economy compared to standard models. In modern economies very articulated and heterogeneous institutional arrangements often govern the functioning of key markets (e.g. labor and credit markets). The same macroeconomic policy can have different effects according to the different institutional setting in which it is implemented (Stiglitz, 2011). The papers of **Cincotti, Raberto and Teglio** (2012) and of **Ashraf, Gershman and Howitt** (2012) use this feature of ABMs to study of bank regulation policies. The paper of **Napoletano *et al.*** (2012) studies the effects of wage-flexibility policies under different institutional regimes characterizing the investment behavior of firms. **Guerci and Sapio** (2012) compare the effects of wind power supply on prices in the Italian electricity market between a scenario in which plant-level and demand data are calibrated on real-data and a scenario where wind electricity output is progressively scaled up to the Italian wind potential, *i.e.* he maximum amount of wind energy that could in principle be produced given the geographical characteristics of the Italian territory.

This brief overview of the papers in this special issue clearly illustrates the great flexibility and the great potential of ABMs for the analysis of key issues emerged in the recent crisis as well as of other important problems both in micro- and macroeconomics. Certainly, the use of ABMs in economics involves new problems and challenges for researchers, especially fore those more accustomed with standard modeling approaches. At the same time ABMs and computational models allow one to avoid the straitjackets, of standard models in the analysis of important real situations and allow policy analyses under more realistic scenarios. These last features should be considered as key improvements to the toolbox of both micro- and macroeconomists.

References

Abrams D.M., and Strogatz, S.H., 2003. "Modeling the dynamics of language death". *Nature,* 424:900.

Amendola, A., Gaffard, J.L., and Saraceno F., 2012. "Production process heterogeneity, time to build, and macroeconomic performance", *Revue de l'OFCE, Debates and policies* 124, this issue.

Ashraf, Q., Gershman, B., and Howitt, P. 2012. "Macroeconomics in?a Self-Organizing Economy", *Revue de l'OFCE,* this issue.

Babutsidze, Z., 2012. "Asymmetric (S,S) Pricing: Implications For Monetary Policy", *Revue de l'OFCE, Debates and policies* 124, this issue.

Barde, S., 2012. "Of Ants And Voters. Maximum Entropy Prediction Of Agent-Based Models With Recruitment", *Revue de l'OFCE, Debates and policies* 124, this issue.

Caballero, R. 2010. "Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge Syndrome", *Journal of Economic Perspectives* 24(4), 85–102.

Ciarli, T., 2012. "Structural Interactions and Long-Run Growth. An Application of Experimental Design to Agent-Based Models", *Revue de l'OFCE, Debates and policies* 124, this issue.

Cincotti, S., Raberto, M., and Teglio A., 2012. "Macroprudential Policies in an Agent-Based Artificial Economy", *Revue de l'OFCE, Debates and policies* 124, this issue.

Dawid, H., and Neugart, M., 2011. "Agent-based Models for Economic Policy Design," *Eastern Economic Journal*, vol. 37(1), pages 44-50.

Delli Gatti, D., Di Guilmi, C., Gallegati M., and Landini S., 2012. "Reconstructing Aggregate Dynamics In Heterogeneous Agents Models. A Markovian Approach", *Revue de l'OFCE, Debates and policies* 124, this issue.

Epstein, J., Axtell, R. 1996. *Growing Artificial Societies: Social Science from the Bottom Up*, The MIT Press.

Fagiolo, G., and Roventini, A., 2012. "Macroeconomic policy in DSGE and agent-based models", *Revue de l'OFCE, Debates and policies* 124, this issue.

Grazzini, J., Richiardi, M., and Sella, L., 2012. "Indirect estimation of agent-based models. An application to a simple diffusion model,", *LABORatorio R. Revelli Working Papers Series* 118, LABORatorio R. Revelli, Centre for Employment Studies.

Guerci, E., and Sapio A., 2012. "High wind penetration in an agent- based model of the electricity market. The case of Italy", *Revue de l'OFCE, Debates and policies* 124,, this issue.

Kirman A, 1993. "Ants, rationality and recruitment", *Quarterly Journal of Economics,* 108:137–156.

Kirman, A., 2012. "Can Artificial Economies Help Us Understand Real Economies", *Revue de l'OFCE, Debates and policies* 124, this issue.

Krugman, P. 2009. "How Did Economists Get It So Wrong?", *The New York Times*.

Leijonhufvud, A. 1993. "Towards a not-too-rational macroeconomics", *Southern Economic Journal*, 1-13.

Napoletano, M., Gaffard, Jean-Luc, and Babutsidze, Z., 2012. "Agent Based Models: A New Tool for Economic and Policy Analysis" *OFCE briefing paper* No3/March.

Napoletano, M., Dosi, G., Fagiolo G., and Roventini, A., 2012. "Wage Formation, Investment Behavior and Growth Regimes: an Agent-Based Analysis", *Revue de l'OFCE, Debates and policies* 124, this issue.

Patriarca, F., and Vona F., 2012. "Environmental taxes, inequality and technical change", *Revue de l'OFCE, Debates and policies* 124, this issue.

Pyka, A. and Fagiolo, G. 2007. "Agent-based modelling: a methodology for neo-Schumpeterian economics", *Elgar companion to Neo-Schumpeterian Economics*, 467.

Saviotti, P.P., and Pyka A., 2012. "On the Co-Evolution of Innovation and Demand: Some Policy Implications", *Revue de l'OFCE*, this issue.

Stiglitz, J. 2011. "Rethinking macroeconomics: what failed and how to repair it", *Journal of the European Economic Association*, vol. 9(4), pages 591-645, 08.

Tesfatsion, L. 2006. "Agent-based computational economics: A constructive approach to economic theory", *in* L. Tesfatsion & K. Judd, ed., 'Handbook of Computational Economics', Elsevier, pp. 831–880.

Trichet, J. 2010. "Reflections on the nature of monetary policy non-standard measures and finance theory", *in* "opening address at the 6th ECB Central Banking Conference", *Frankfurt am Main*.

Wieland, W., and Wolters, M., 2012. "Macroeconomic model comparisons and forecast competitions", *VoxEU*, 13 February.

CAN ARTIFICIAL ECONOMIES HELP US UNDERSTAND REAL ECONOMIES?

Alan Kirman
GREQAM, Aix Marseille Université, EHESS

This paper argues that the path followed by modern macroeconomic theory excludes the analysis of major endogenous movements in macroeconomic variables. Rather than persist with models based on the idea that the economy behaves like a rational individual we should build models of the economy as a complex system of interacting agents. Profiting from the advances made in computer science we can now build agent based or computational models which allow us to model the out of equilibrium behaviour of such systems. They allow us to remove many of the restrictive assumptions of standard models and to incorporate the heterogeneity of economic agents and the evolution of the network that governs the interactions between the individuals and firms in the economy. Such models can help fill the theoretical void with which policymakers declare that they have been faced in the current crisis.

Keywords: crisis, complex system, heterogeneity, networks.

> "*First, we have to think about how to characterise the homo economicus at the heart of any model. The atomistic, optimising agents underlying existing models do not capture behaviour during a crisis period. We need to deal better with heterogeneity across agents and the interaction among those heterogeneous agents. We need to entertain alternative motivations for economic choices. Behavioural economics draws on psychology to explain decisions made in crisis circumstances. Agent-based modelling dispenses with the optimisation assumption and allows for more complex interactions between agents. Such approaches are worthy of our attention.* "
>
> Jean-Claude Trichet (2010).

Recently, considerable dissatisfaction has been expressed not only by Trichet but also by Bernanke, Turner and many others in policymaking positions, with economic theory in general and macroeconomic theory in particular. This leads to two questions, why does macroeconomic theory seem to have become, as Trichet and others suggest, so irrelevant particularly in times of crisis and to what extent are agent based or computational economic models a more satisfactory alternative?

Before answering these questions it is worth observing that computational or algorithmic models have a long and distinguished tradition in economics. One has only to think of Von Neumann's work and of Scarf's contributions, particularly to the computation of economic equilibria and the contributions of Dantzig and Kuhn in using mathematical programming, to see this. What agent based modelling is doing is to renew the tradition of using an algorithmic approach to model economic phenomena. The recent, rapid development of computer science explains the resurgence of this approach. Now computational models have two significant advantages. Firstly there is a wealth of data on the behaviour of individuals and from this we can categorise different behavioural types and use these as a basis for building agent based models. Secondly, given the progress that has been made in computational power and capacity we are now capable of simulating the dynamics of very large systems of heterogeneous interacting agents. This is, for example, the ambition of the EU project, Futur-ICT.

1. The evolution of modern macroeconomics

But to return the first question as to how macroeconomic theory seems to have become so irrelevant, it is worth looking at the recent evolution of the discipline of economics. Economic theory has developed, as a criterion for "rigorous" analysis that our models should be based on the intellectual bedrock of the "rational individual" or homo oeconomicus. The rationality of economic agents is not defined by the intuitive idea that individuals do not act against their own interest, but rather, that they have preferences, which satisfy the axiomatic structure typified by that used in the Arrow Debreu model. For firms it is even simpler, they choose that combination of inputs and outputs which maximises their profit. If we take time into account we assume that our agents, consumers and firms, have infinite horizons and that they discount the future appropriately. Lastly if the future is uncertain, they know the nature of that uncertainty, they know the probability distribution of future events, they have "rational expectations". This is the basis for the underlying theoretical model of modern economics the "general equilibrium model" initiated by Walras, improved by Pareto and honed to perfection by their successors and which culminated in the Arrow Debreu model and was closed by assuming rational expectations. The task of the economist in this tradition is, therefore, to make assumptions about individual preferences and technologies and to build models, particularly macroeconomic models on that basis. One then finds the equilibrium of the system and examines the characteristics of such equilibrium states. Whenever objections as to the realism of the model are made, the response is to modify some aspect of it to take the criticism into account but not to put the whole exercise in doubt.

The problems with this approach are well known and need not be elaborated here. Suffice it to say that we know that restricting ourselves to only making assumptions on individual preferences and technologies will allow us to say nothing about how one might get to an equilibrium nor whether such a state is unique. Thus, all that we can say is that, under our assumptions, an equilibrium will exist. But, since the idea that the only way to construct "scientific" models is to limit ourselves to assumptions on individuals, all that we can do is assume the economy to be in equilibrium all the time. This is precisely where modern macroeconomics has arrived. The answer to building models which allow us to say something about how economies function out of equilibrium and how they might move to equilibrium seems to be simple, add assumptions about the way in which people interact and the institutions that govern them, and this is precisely what Lucas suggested.

> "*Applications of economic theory to market or group behaviour require assumptions about the mode of interaction among agents as well as about individual behaviour.*" Lucas (1988).

However, curiously, rather than do this, and maybe change the basic approach to economics, the solution that has been adopted is to assume that the economy as a whole acts like a rational individual, an assumption for which there is no theoretical justification, (see Kirman (1992). Perversely, the idea of enlarging the model to incorporate assumptions about how it is organized has been considered to be "unscientific" whilst the assumption that the economy acts like an individual was not. Not only was this intrinsic contradiction ignored, the underlying assumption that we should only make assumptions about individuals was actually elevated to the status of a principle. Lucas, some years before making the observation just mentioned, explicitly rejected the idea of adding parameters to the basic model to allow even for an adjustment process. In fact at that time he said,

> *"Now, I am attracted to the view that it is useful, in a general way, to be hostile toward theorists bearing free parameters, so that I am sympathetic to the idea of simply capitalizing this opinion and calling it a Principle."* Lucas (1980, p. 709).

But, by doing so he made it impossible for economists, who followed his lead, to study out of equilibrium phenomena. Since, with the assumptions that he considered scientific, all that we could show was that an equilibrium exists, the economy should be studied in that state. Even if one considers dynamics, the idea would be that the economy simply evolves through a sequence of equilibrium, thus business cycles are equilibrium phenomena. The fact that, at another point of time, Lucas suggested that we needed additional assumptions on the organisation of the economy in addition to the assumptions on individuals, did not deviate macroeconomic theorists from the path which he had encouraged them to pursue.

Following this theoretical path has had important consequences for the way in which macroeconomics has been developed. Despite Lucas' observations, it is generally assumed in macroeconomic models that the way in which the economy or markets are organised, as such, has little impact on economic outcomes. Individuals participate in anonymous markets in which they are price takers and little is said about who sets the prices and how. When exceptions are made to this basic structure it is significant that economists refer to "imperfect competition" and market "imperfections". Thus there is a benchmark model in which individuals interact only through the price system and other situations in which individuals react to each other are thought of as deviations from the norm. Direct interaction, and the externalities that go with it are either declared to be the subject of game theory or are incorporated with difficulty into a modified GE model.

Our attraction for the idea of economics as a "science" which stems from Walras' obsession with showing that we could develop a general internally consistent mathematical model of the economy, has driven us into a corner. The attitude of many theoretical economists to real world economic problems is directly in the spirit of Bourbaki from whom Debreu took his inspiration. As Bourbaki[1] said,

> "*Why do applications [of mathematics] ever succeed? Why is a certain amount of logical reasoning occasionally helpful in practical life? Why have some of the most intricate theories in mathematics become an indispensable tool to the modern physicist, to the engineer, and to the manufacturer of atom-bombs? Fortunately for us, the mathematician does not feel called upon to answer such questions.*" (Bourbaki, *Journal of Symbolic Logic* 1949, 2).

Thus, in that spirit, the furthering of economy theory was seen as an avenue to more advanced models and not as a pursuit of explanations of economic phenomena. We became totally preoccupied by the properties of the economy in an equilibrium state. But, given the important results established in the '70s[2] it became clear that we had, within our "scientific" models, to abandon the concern with how the equilibrium prices are established and how the economy evolves towards equilibrium. There was almost no consideration of the idea that the economy might never be in equilibrium in the standard sense. So theorists have concentrated on the properties, in particular, the efficiency, of equilibrium states. They have insisted on the rigour of the analysis, but much less on the realism of the assumptions. In the end, the mathematical road that we followed petered out some 40 years ago in pure theory and has only remained in macroeconomic theory.

2. An alternative approach

Keynes once remarked, and this remark has been widely cited recently, that economists should become more like dentists, using such knowledge of the economy that they have acquired to improve the health of the patient particularly in times of crisis. Colander (2011) refers to this as, an "engineering", rather than a "scientific" approach. Rather than developing general theories which have the ambition of giving us a clear, if simplified, vision of the economy as

1. Nicolas Bourbaki was, of course, a pseudonym for a group of mathematicians mainly based in Paris.
2. The results in question are those of Sonnenschein (1972), Mantel (1974) and Debreu (1974) himself.

a whole and independent of context, he argues that we should concentrate on developing models capable of providing us with recommendations for specific situations. Unconcerned with the basic physics that underlies the tools that he uses, the engineer tries to solve problems, often relying on generous approximations to give him self a margin of safety.

I would like to argue that we should not necessarily abandon the idea of a general model, after all Keynes called his major contribution, "The General Theory…". This does not mean that we should aim at a model of everything, but rather, that we can establish some basic principles on which our models should be built. However, if we are to have such basic foundations they should be radically different to those that we are in the habit of using. Our basic model, as suggested in Miller and Page (2007), Epstein (2007) and Kirman (2010) should be one of a complex system of interacting agents who learn and adapt (for an early exposition, see e.g Aoki (1996)). Their behaviour and their interaction with each other generates aggregate phenomena, from which they again learn or to which they adapt. There are positive feedbacks, and we have no reason to believe that such a system will converge to a stationary equilibrium. We have therefore to study economies that are out of equilibrium and how they evolve. In such systems the aggregate behaviour is not like that of an individual and the way in which individuals react to each other will be an important feature of the evolution of the economy. Two things should be emphasised. Firstly, we cannot assume the aggregation problem away as is currently done in macroeconomics. Secondly, we need to understand out of equilibrium dynamics.

I claim therefore that we have to turn things inside out and bring direct interaction to the centre of the stage. Furthermore, I claim that we should radically simplify our models of individuals and that, in so doing, we may still observe interesting and complicated aggregate behaviour which is, however, the result of the aggregation itself and not of the complicated behaviour of some "representative individual". We should treat the economy as a complex system and, as in other disciplines, we should not expect the system to behave like an individual.

The way to do this is by building models of simple individuals who follow simple rules and interact with each other just as molecules in biological systems or particles in physical systems. This is at the heart of agent based models or computational economics. Indeed, the now widely accepted definition of this approach is that given by Leigh Tesfatsion[3], (2002) where she says the following,

> *"Agent based computational economics is the computational study of economic processes modelled as dynamic systems of interacting agents."*

The fact that we are talking about a computational approach does not mean that we are abandoning any attempt to obtain analytical results, but does mean that a reasonable way to proceed is to try to obtain formal results in simple models and, where this proves too difficult in more general models, to use simulations and to see whether the results persist in those contexts. The formal analysis is more likely to come from statistical physics (see e.g. Blume (1993)), discrete mathematics and computer science, than the sort of mathematics that we use, in general, in theoretical economics. This does not make it any less rigorous. Lucas' principle, as he stated it at the time, was based on the false premise that our assumptions on individuals are, in some sense, scientific. In fact, those assumptions have no special status. They come from the introspection of economists and not from the long and careful examination of how individuals actually behave. We have become familiar with them and this has made them acceptable. But there is no reason that we should not adopt different formal models that were originally used to explain the behaviour of systems of interacting particles or molecules.

The usual argument against this is that humans have intentions and are forward looking and cannot therefore be modelled as one would model molecules or inanimate particles. This misses the essential point, if we can describe the rules that an individual follows and the way, in which he interacts with other individuals, we can use the formal models developed elsewhere to understand what the outcomes will be. We do not need to know what the deep psychological motives for an individual's actions are. Consider the argument that individuals are forward looking, and think of the problem of forecasting. In all our models individuals map past history into forecasts of the future. Once they have a forecast of the

3. Those who are interested in the agent based modelling approach cannot do better that to go to Tesfatsion's website devoted to this issue, http://www.econ.iastate.edu/tesfatsi/amulmark.htm

future they take an action, so in the end, they just have an algorithm, which maps past history into actions. There is nothing intrinsic which prevents us from building simple agents or robots that do this. We can choose what we consider to be the appropriate level of sophistication for the mapping from the past to actions; we can also model the reactions of other agents to an individual's choice of actions. What is more we can let the agent learn about the rules that he uses and we can find out if our simple creatures can learn to be the sophisticated optimisers of economic theory.

In doing this we are not straying far from what has been recommended by our illustrious predecessors and the leaders of our profession. The first idea that I am suggesting is that we have to treat the economy as a complex system. But Herb Simon (1962) already described a complex system when explaining how he thought economic theory should develop and he said,

> *"Roughly by a complex system I mean one made up of a large number of parts that interact in a non-simple way. In such systems, the whole is more than the sum of the parts, not in an ultimate metaphysical sense, but in the important pragmatic sense that, given the properties of the parts and the laws of their interaction, it is not a trivial matter to infer the properties of the whole. In the face of complexity, an in-principle reductionist may be at the same time a pragmatic holist."* Herbert Simon (1962, p. 267).

The second argument that I would make is that we should dispense with the a priori assumptions about rationality and optimisation, which are so central to economic theory. But, if you think that this might be heretical consider what Robert Lucas (1988), had to say on the subject:

> "*In general we view or model and individual as a collection of decision rules (rules that dictate the action to be taken in given situations) and a set of preferences used to evaluate the outcomes arising from particular situation-action combinations. These decision rules are continuously under review and revision: new decisions are tried and tested against experience, and rules that produce desirable outcomes supplant those that do not. I use the term "adaptive" to refer to this trial-and-error process through which our modes of behaviour are determined.*"

However, Lucas then goes on to argue that we can safely ignore the dynamics of this process since,

> "*Technically, I think of economics as studying decision rules that are steady states of some adaptive process, decision rules that are found to*

> *work over a range of situations and hence are no longer revised appreciably as more experience accumulates*.".

In general, however, one cannot assume convergence to some equilibrium but one has to look at the dynamic evolution of the economy resulting from the interaction between agents. One is also interested in knowing how the state of the system evolves over time and not only whether it settles down to what might be thought of as some sort of equilibrium. Here I am taking a different position from Lucas and arguing that one cannot assume that all the adaptation has taken place in the past but that we are faced, in economics, with many situations in which individuals are constantly adapting to change and thereby generating change. Thus, not only the relevant time scale but also the process itself is very different from that relevant for biological evolution, which is too often used by simple analogy. Indeed when explaining the difference between standard and computational or agent based economic models, Farmer and Foley (2010) explain,

> "*An agent-based model is a computerized simulation of a number of decision-makers (agents) and institutions, which interact through prescribed rules. The agents can be as diverse as needed — from consumers to policy-makers and Wall Street professionals — and the institutional structure can include everything from banks to the government. Such models do not rely on the assumption that the economy will move towards a predetermined equilibrium state, as other models do. Instead, at any given time, each agent acts according to its current situation, the state of the world around it and the rules governing its behaviour."* Farmer and Foley (2010) p.685.

In fact, the economy is a noisy system that may not show any tendency to settle to a steady state. Thus the argument that individuals learn to achieve equilibria does not take into account the fact that the environment about which they are learning is largely composed of other agents who are also learning. This undermines the basic arguments made by Evans and Honkapohja, (2001) who argue that learning in macroeconomics leads to equilibrium states.

If we take the view that most of the dynamics of the economy is due to the interaction between the heterogeneous agents in the economy, this means taking a very different view of business cycles and crises. Rather than putting these down to some exogenous, (technological) shock, the shocks would be idiosyncratic ones which affect individuals or firms and are transmitted to others. The system is not occasionally knocked off course and then returns to its steady state path but internally generates movements and from time to time, phase changes. As Wilhem Buiter, a former member of the Bank of England's Monetary Policy Committee and now chief economist of Citigroup, says,

> *"Those of us who worry about endogenous uncertainty arising from the interactions of boundedly rational market participants cannot but scratch our heads at the insistence of the mainline models that all uncertainty is exogenous and additive."* Buiter (2009).

3. Business cycles and crises

What is necessary then is to build models in which much of the volatility of macroeconomic variables is accounted for by microeconomic shocks. A step in this direction has been made by Gabaix (2011) who suggests that simply making assumptions about the heterogeneity of firms more consistent with the observed data, can make a major difference. He points out, in a recent paper, that if one accepts, consistently with the empirical evidence, that the distribution of firm sizes in the U.S. follows a power law, then the idiosyncratic movements of the largest 100 firms appear to explain about one-third of variations in output growth. The underlying assumption about firm sizes, that he refers to as the "granular" hypothesis, suggests that a number of macroeconomic questions can be clarified by looking at the behaviour of large firms. He actually details the history of a number of major incidents at the firm level that were translated into movements of aggregate output. The underlying idea is not new and in the past several economists have proposed mechanisms that generate macroeconomic shocks from purely microeconomic causes. A pioneering paper is by Jovanovic (1987), whose models generate non-vanishing aggregate fluctuations owing to a multiplier proportional to $\sqrt{N}$, the square root of the number of firms in the economy. However, Jovanovic's results have been criticised as empirically implausible. Yet a number of economists followed the route he mapped out. For example Scheinkman, and Woodford (1994) applied the physical theory of self-organizing criticality to explain aggregate fluctuations. Their approach however, generates fluctuations which are more fat-tailed than in reality, with infinite means. Again Nirei (2006) proposed a model where aggregate fluctuations arise from the (s,S) rules developed by Scarf and others at the firm level. Each of these examples is close, in spirit, to the complex systems approach that I am recommending here. Still we have a large residual to explain in terms of the overall size of aggregate shocks. This leads naturally to the second point, that individual shocks may be transmitted to other individuals producing contagion effects, and thus relatively small shocks may be amplified, through the network of linkages between firms. This suggests that we have to pay much more attention to the network structure of the economy than is typically done.

4. Networks

As I have said, it is remarkable that Lucas observed that we had to develop hypotheses about the organisation of the interaction between agents in direct contradiction with the principle that he invoked when he argued that we should not add any assumptions to those we make about the characteristics of individuals. Nevertheless, it is directly in line with the basic argument of this paper, which is that we have to model the economy as a system in which rather simple individuals, organisations or enterprises interact with each other. The complicated outcomes that we observe at the aggregate level are not due to the complicated nature of the individuals but rather to the structure of the system within which they interact. However, once one argues that it is the interaction between agents that is primordial in determining the aggregate outcomes then one has also to be much more specific about the way in which those interactions are structured.

In particular the network of relationships between individuals, banks and even nations are of primary importance. They do not figure in macroeconomic models but have been extensively studied by economists (see e.g. Jackson (2008), Goyal (2007), and Vega Redondo (2007).)

This means modelling the network of links between the individuals and institutions, specifying the nature of the nodes and links and, in addition, establishing criteria for determining their robustness. Here, however, we find an interesting problem with the economist's approach to networks. Economists wish to develop a very general theory and, in particular, one which is based on individual maximising behaviour. As Goyal (2007) says,

> *"The distinctiveness of the economics approach to networks lies in the different methodology used. These differences can be traced to a substantive methodological premise in economics: social and economic phenomena must be explained in terms of the choices made by rational agents."*

In fact if, as I have claimed, and others do, (see Haldane and May (2011), that we have to view networks as an integral part of the vision of the economy as a complex adaptive system, what we need is a very different basis for our models. Again Lucas' recommendation that we interest ourselves in the structure of interactions, reveals the difficulties in dealing with phenomena such as networks and trying to stick to the conventional macroeconomic approach. Firstly, it underlines the economist's quest for a very general abstract model which will encompass many of the empirical phenomena that we observe and, secondly, the need that theoretical macroeconomists feel to base the model on the same micro-foundations that have been shown to be inadequate as a basis for a general, empirically verifiable model.

A different approach is needed. For example if we now couple the role of networks with the explanations of the origins of aggregate shocks, proposed by Gabaix (2011) we can move a long way to explaining a large part of the volatility of macroeconomic variables and the onset of crises. As Gabaix says,

> *" It would be interesting to exploit the hypothesis that the financial crisis was largely caused by the (ex post) mistakes of a few large firms, e.g., Lehman and AIG. Their large leverage and interconnectedness amplified into a full-fledged crisis, instead of what could have been a run-of-the-mill (sic) that would have affected in a diffuse way the financial sector." Gabaix* (2011, p.764).

Thus what we need to do is to integrate considerations of the interconnectedness of the network into explaining macroeconomic evolution. This has been the subject of a considerable theoretical and empirical literature. Durlauf (1993) generated macroeconomic uncertainty with idiosyncratic shocks and local interactions between firms. His results are driven by the nonlinear interactions between firms. This sort of result based on diffusion across a network, coupled with the « granularity » of firm size advanced by Gabaix (2011) is a powerful basis for examining the dynamics of the macro-economy. The argument would then be that the skewed size distribution of firms together with their strong inter-linkage underlies much of macroeconomic fluctuations. The interdependence may be due to the input output structure of the economy but might also be linked to the ownership structure of the large firms in an economy. If we are interested in global fluctuations then we need, for the latter explanation, a detailed analysis of the ownership structure of the largest international firms, or Trans-national corporations (TNCs) as they are sometimes called. Vitali *et al.* (2011). use a large data base on the share holdings of firms to establish the nature of the network of international ownership. Their findings, which have provoked considerable discussion, are remarkable. As the authors explain

> *"Nearly 40 percent of the control over the economic value of TNCs in the world is held, via a complicated web of ownership relations, by a group of 147 TNCs in the core, which has almost full control over itself."* Vitali *et al.* (2011).

Unsurprisingly, three-quarters of these companies are banks.

5. The emergence of the networks structure

An important observation is that there is no evidence that this structure was intentional, there was no giant conspiracy, rather the network evolved endogenously in this way. The network is an emergent phenomenon, characteristic of complex systems. However, concentrated power in the hands of a few has clear implications for global financial stability as recent events have shown. What is worth observing is that, starting from the empirical evidence, the authors were able to build a picture of the structure of the network and then, to emphasise the implications of that structure for the stability of the network. Building on this approach could potentially help policymakers and economists to find ways to stabilize financial markets.

The important thing to notice here is that two statistical characteristics of the network of the international network of firms allow one to draw conclusions as to the evolution and fragility of the international economy. This does not depend on finding microfounded explanations of those statistics although a number of efforts have been made to do This is at the opposite end of the spectrum to the idea that fluctuations at the macroeconomic level are essentially due to large unspecified (technological?) shocks to the economy as a whole.

Why do I describe this as belonging to the realm of "agent based|" or computational modelling? What we are looking at here is a system of different interacting agents, and one where the differences between the agents can be measured. Starting from the distribution of the sizes of the firms in the particular system in question, one already obtains a measure of the aggregate volatility induced by idiosyncratic shocks. If one then couples this with the network structure of the firms, one can then model the aggregate impact of individual shocks. But both the size distribution of firms and the nature of the network are initially inferred from the data, so this is an inductive rather than a deductive approach. The robustness of the conclusions to the specification of the distribution and of the network structure can be examined by simulating a model in which the distribution of firm sizes is varied and where the network structure can be modified.

To take another example, the properties of the international financial network have been examined by Haldane, the director of the Bank of England responsible for financial stability (Haldane (2009)). In this case, the nodes correspond to countries and the size of the nodes to the total amount of foreign assets held by the country corresponding to the node in question. A link between countries means that at least one of the two holds the assets of the other and these links are weighted by the sum of the mutually held assets. Typically one would define a minimum threshold for such assets to constitute the basis for a link. One can calculate a number of statistics to characterise the structure of the graph, the empirical degree distribution for example, and the proportion of the total weight of all the links made up by the total of the weights associated with the links emanating from the largest nodes. Whilst the connectivity of the global financial network has increased remarkably in recent years (see Nier et al. 2007), the degree distribution3 has changed and has become more skewed with a few nodes having very high degree and a group of nodes becoming very central. To quote Haldane (2009) of the Bank of England, when talking about these developments in the banking network before the global financial crisis, he says:

> *"This evolution in the topology of the network meant that sharp discontinuities in the financial system were an accident waiting to happen. The present crisis is the materialisation of that accident."* Haldane (2009, p. 4).

Again note that the reasoning is inductive. Start with an empirical phenomenon, establish its features and then examine the consequences for the evolution of the system of those features. Here what one is looking at is a network that emerged from a particular evolution of trading relationships which were mutually advantageous. What we see is how it can become fragile, without those who participate in it realizing what is going on. The importance of this for economists is clear. Interaction and the networks through which it operates have to be analysed since they play a large role in determining aggregate economic phenomena. Furthermore, understanding the evolution of the structure of the networks that make up the economy is not just an intellectual exercise; it is important for very practical reasons and policy makers are coming to appreciate this. For as Haldane says,

> *"Deregulation swept away banking segregation and, with it, decomposability of the financial network. The upshot was a predictable lack of network robustness. That is one reason why Glass-Steagall is now back on the international policy agenda. It may be the wrong or too narrow an answer. But it asks the right question: can network structure be altered to improve network robustness? Answering that question is a mighty task for the current generation of policymakers. Using network resilience as a metric for success would help ensure it was a productive one."* Haldane (2009).

When models that would address these questions are proposed, they are often described as following an engineering methodology rather than a scientific one. Rather than demanding the total internal consistency which characterises current economic models such a methodology would use a much broader and looser set of assumptions that would blend economic and non-economic considerations. In this view, all aspects of the problem necessary to arrive at an answer to the economic problem posed would be included in the applied economist's research. Thus, for example, if judgments about tradeoffs of individual's welfare were necessary, the economic engineer would develop as his objective a method for making those judgments as possible. Again this sort of approach is not new. Earlier economists, who took a more engineering approach, were quite willing to develop models involving interpersonal welfare comparisons as Colander (2007) points out. He gives the example, of Irving Fisher (1927) and Ragnar Frisch (1932) who developed a statistical method for making interpersonal comparisons of wants; they justified their models by pragmatism. Fisher posed the rhetorical question about whether the necessary assumptions can be used, and answered: "To all these questions I would answer 'yes'—approximately at least. But the only, or only important, reason I can give for this answer is that, in actual practical human life, we do proceed on just such assumptions." He continues: "Philosophical doubt is right and proper, but the problems of life cannot, and do not, wait." (Fisher, 1927). As Colander says, maximizing a non-operational social welfare function is not a policy goal of engineering research.

Whilst many would accept to accept the idea that an engineering approach may be useful in providing solutions to, or understanding of, very specific economic problems they would ask that it should also allow us to better understand more general economic phenomena. A first example of an area where this approach may be particularly useful is that of the market.

6. Markets and their organisation

Of all economic institutions the market is probably the most ancient and the most historically documented. If any feature of the economy is emphasised in analysing economic phenomena it is surely the market. Indeed, as Braudel observed,

> *"Ce très vieux type d'échange se pratiquait dejà à Pompei, à Ostie ou à Timgad la Romaine, et des siècles, des millénaires plus tôt : la Grèce ancienne a eu ses marchés; des marchés existent dans la Chine classique comme dans l'Egypte pharaonique, dans la Babylonie où l'échange était si précoce... En Ethiopie, les marchés par leurs origines se perdent dans le temps."* Braudel (1979).

Yet as Douglas North remarked,

> *"It is a peculiar fact that the literature on economics…contains so little discussion of the central institution that underlies neoclassical economics—the market."* North (1977, p.710).

One has only to think of the standard vision, in economic theory, of a market to see why there is such a gulf between what Braudel is describing and modern economic theory. What is described is a system in which the actors act according to a system of rules, which constrains them, and this, in turn, generates the aggregate economic outcomes. These actors are anonymous and their relations with others are not considered. Financial markets are often analysed on the basis of such a vision. Yet, those who participate in, who regulate or study actual market mechanisms have a very different view. For example Aboulafia argues that markets are essentially social institutions in his well-known study of financial markets, indeed he says,

> *"Markets are socially constructed institutions in which the behavior of traders is suspended in a web of customs, norms, and structures of control...Traders negotiate the perpetual tension between short-term self-interest and long-term self-restraint that marks their respective communities."* M. Aboulafia (1997).

Kuhn goes further and argues that individual relationships and trust are necessary for the functioning of markets. For him, it is clear that,

> *"Markets are not self-operating, objective mechanical objects. They are, rather, a complex set of constraints, rules, rights, regulations, and laws, guiding human participants in making their multiple, various trades, purchases, and exchanges. The motivating force that generates benign market outcomes is the willingness of all to obey the guidelines and deal openly—transparently—with each other. Invisible to the naked eye are the common social bonds of trust among all, strangers and acquaintances alike. The bonds of trust are what create and sustain truly efficient, effective markets."* Kuhn (1995).

In another context Alan Greenspan, Chairman at the time of the Federal Reserve, has remarked that,

> *"It is hard to overstate the importance of reputation in a market economy. To be sure, a market economy requires a structure of formal rules—a law of contracts, bankruptcy statutes, a code of shareholder rights—to name but a few. But rules cannot substitute for character. In virtually all transactions, whether with customers or with colleagues, we rely on the word of those with whom we do business. If we could not do so, goods and services could not be exchanged efficiently. Even when followed to the letter, rules guide only a small number of the day-to-day decisions required of corporate management. The rest are governed by whatever personal code of values corporate managers bring to the table."* Greenspan (2003).

This poses a problem for those who would like to model the way markets really function. The anonymous market poses few problems, for one just has to specify the rules which individuals follow when they are faced with the prices given by some unspecified market mechanism. However if we take the previous remarks seriously, we are faced with the idea that individuals build up relations of confidence with each other and this seems more like a subject for psychologists or, at least, for behavioural economists.

Furthermore, if we specify who interacts with whom in markets, we simply have to specify the graph in which the various buyers and sellers are linked and we return to the sort of network analysis I have already described. Moreover, if we are to explain how these links form and are sustained, the task is even more difficult. An argument that is often used is that in the large anonymous markets of today, relationships are not longer important, therefore we do not need to worry about the network linking traders together nor how it is formed. The quotes that I have given suggest that this, even today, is far from being the case. Traders in financial markets such as the Forex market use a very limited number of other traders, despite the fact that they are faced with thousands of quotes at any point in time. Again, in a recent paper, Puri *et al.* (2011), analyzed the importance of retail consumers' banking relationships for loan defaults using a dataset of over one million loans by savings banks in Germany. They find that loans of retail customers, who have a relationship with their savings bank prior to applying for a loan, default significantly less than customers with no prior relationship. Thus relationships play a very significant role.

Two remarks are worth making here. Firstly, not long ago such an analysis would have been almost impossible but, as I have remarked, the abundance of data together with the progress in computational capacity now allows us to undertake such exercises. Secondly, the same advances now permit us to build models in which individuals learn with whom they wish to interact and within which one can study the consequences of such interactions. In this way, artificial markets can contribute to the quest for the explanation of some of the features of the complex market structures that we observe. Conventional macroeconomic models are not concerned with the details of how markets function, but a better knowledge of market microstructure may be very useful in explaining some of the evolution of financial markets, for example. However, what we are talking about here is the emergence of certain phenomena and the dynamic evolution of the structure of the relations within markets, where, by their very nature, many transactions are on a repeated basis. Such problems do not sit well in the context of models that aim at the analysis of steady states of systems in which agents only interact through some anonymous unspecified "market".

However, artificial markets on their own are what their name suggests, artificial. The three traditional approaches for economists are, theoretical, experimental and empirical. The idea here is that the fourth approach that I am recommending, that of using agent based models to construct artificial markets which are then simulated, can complement the other approaches each of which has its weaknesses.

What are the drawbacks of theoretical models? The first and most important weakness is that they have to be greatly simplified in order to make them analytically tractable. Miller and Page (2007) in describing the purification of economic models, cite a Chinese philosopher, who says,

"Water which is too pure has no fish." Ts'ai Ken Tan.

That is, in reducing our models to a minimum to be able to solve them analytically, for example to characterise their equilibria we may actually purify away the phenomena we are interested in. The second weakness is that the assumptions are often made for analytic tractability rather than for economic realism. Artificial markets can help here by providing results in more general analytically intractable situations and then seeing if these results coincide with those obtained in the simpler case which could be handled analytically.

A second approach that has developed rapidly in recent years is that of experimental economics. Leaving aside the fact that macroeconomics may not be the most suitable subject for laboratory experiments,[4] one could ask, in general, what are the limitations of experiments? Once again, this time, in order to make the situation understandable for the subjects, one has to simplify. Furthermore, the situation with which the subjects are faced is extremely unnatural. Often they believe that they have a problem to solve for which there is a « right » answer, thus rather than reacting naturally, they try to outguess the experimenter. The first lesson, it seems to me, that we learn is that, even when faced with a well specified problem in a controlled environment, subjects frequently do not behave as theory would predict. Thus, my own view is that this teaches us that individuals are noisier and less consistent than we assume them to

4. Although it is only fair to observe that the formation of bubbles in asset markets and the formation of expectations have been the subject of a considerable number of experiments, (see e.g Duffy (2008)).

be and that we should incorporate this into our models. Some economists are more ambitious and would like to know if, other than using simplified theoretical models, one could not develop another benchmark for the rationality against which to evaluate subjects' behaviour in experiments. Again artificial models could be used to provide such a benchmark.

Finally, why can one not content oneself with working with empirical data directly? Doing this can enable us to establish some « stylised facts » or statistical regularities but gives us no idea as to the structure that generated them. It is precisely to get an understanding of how the system functions, that we typically build a model, and then we are faced again with the choice between a theoretical model and its agent-based counterpart. Although we can usefully employ both approaches, the typical theoretical models is developed before looking at the facts, while, what I would argue is that we should use salient features of the empirical data as our benchmark, and then construct models, which reproduce some of these. This is the essence of the agent based approach which is essentially data driven and more inductive than deductive.

As a very simple example, in Weisbuch *et al.* (2000), and Kirman and Vriend (2001) we wished to explain the strong loyalty of many buyers to sellers in the Marseille fish market. We first developed an extremely primitive theoretical model in which people simply learn from their previous experience and then, in consequence, change their probability of visiting different sellers as a result of their experience. We then went on to simulate more elaborate versions of the model and were still able to reproduce the salient feature. Models of this sort that attribute very little computational ability or general reasoning capacity to individuals are capable of reproducing specific features of real markets. Since then a literature on this sort of model for these markets has developed. (see e.g. Sapio *et al.*, 2011).

As a further illustration, consider another example of a market, but this time for a financial asset, (Anand *et al.* (2010)), where we once again started with an empirical phenomenon, the collapse of the price of asset backed securities early in the current crisis. We first constructed a simple theoretical model to capture the essence of the phenomenon and then ran simulations of a more general dynamic model in which the agents act in the same way as in the theoretical model to see if the model evolves to the states predicted by the theory.

Our goal was to model the general mechanism whereby investors, as a rule, trade securities without giving due diligence to fundamental information that is, they do not check on the toxicity of the asset. The rationale motivating investors, is simply that it is profitable to adopt this rule, because other investors have already adopted it.

The market consists of agents, who, in the case of the sub-prime crisis, we can think of as the banks who were both the issuers and the investors in these Asset Backed Securities,(ABS). Each agent decides whether or not to follow a rule, which is to purchase an ABS, relying on signals from the rating agencies, without independently evaluating the fundamental value of underlying assets. If enough other participants do so, the agent becomes convinced, not irrationally, that the ABS is highly liquid and hence easy to trade.

Let us assume that the ABS is toxic with a certain probability. By toxic I mean, for example, that the underlying asset was incorrectly graded and that the original borrower of loan has already defaulted or has a higher probability of defaulting. Agents are linked together with trading partners in a financial network. This captures the fact that the secondary market for trading ABS and other credit derivatives is not centralized but instead takes place over-the-counter with traders in one firm directly calling up other traders to sell their securities.

When an agent *i* receives an offer to buy a new ABS, she considers whether or not to follow the rule. The line of reasoning she pursues is to first determine the probability that, if she adopts the rule and subsequently attempts to re-sell the security, the potential buyer, agent *j* will refuse to buy the security. This will be because agent *j* does not follow the rule and, as such, may verify that the underlying asset is toxic and, hence, not liquid. Each agent now calculates the expected gain to him of following the rule given the rules chosen by the neighbours in his network and adopts the rule if the expected pay-off is higher than that obtained by not adopting it and checking.

It is not difficult to find the equilibria of this simple market, in terms of the rule being followed, and there are two, one of which is always an equilibrium, and the other which only appears above a certain critical value for the probability of default on the underlying asset. In the first equilibrium no banks check on the underlying assets whilst in the second all banks do so. Now in order to test the stochastic stability of the two equilibria we ran simulations in which agents noisily learn (they use reinforcement learning, (see Bush and Mosteller, 1955 or for a more sophisticated version Camerer and Ho, 1999), which rule is more profitable. What transpires from the simulations, is that the system always converges to the no-checking equilibrium if the probability of default is low enough, but a small increase in that probability, can lead the market to collapse into the equilibrium in which everyone checks. Thus a small change in one of the parameters of the model can lead to catastrophic consequences at the aggregate level.

Indeed, what we did was to examine the co-evolution of the default rates on mortgages and the prices of securities backed by those mortgages. The default rates steadily increased but this was not reflected in the price of assets until they suddenly collapsed and the interbank market froze. Thus, a continuous change at one level led to a discontinuous change at the aggregate level. Whilst we could establish the existence of the equilibria of the model analytically, we had to resort to simulations to see to which equilibria the learning process converged.

This underlines an important message. As soon as we are interested in real economic phenomena we cannot avoid examining how the economy behaves out of equilibrium and the characteristics of the states through which it passes, or to which it settles. This sort of "bounded rationality" approach has received a lot of attention but is often dismissed for its lack of rigour. In reality, the analysis of the evolution of the "state" of the market in the model can be perfectly rigorous given the specific choice of rules for the agents. Yet, it is still the case that building artificial markets or economies, in which agents have simple rules of behaviour, is not widely accepted in economics. The reason seems to me simple; choosing rules of thumb for agents is regarded as *ad hoc*. However, as I have already explained, we have come to accept that the restrictions that we impose on the preferences of individuals, unlike other behavioural rules, are not *ad hoc*. Therefore, if we replace those assumptions, which, by their very nature, cannot be empirically tested, by other rules, we are subject to the criticism that we lose the rigour of "proper micro foundations".

One of the arguments of the proponents of agent based modelling is that, unlike the more standard economic models, their approach is largely innocent of theoretical pre-suppositions. In many agent based models, the individuals learn, as Lucas (1988) would have them do, which rules to use, and this is not dictated a priori. However, it should be noted that the very specification of the rules amongst which the agent chooses has an impact on the outcomes. Ideally, one would like to start, as in the "artificial life" approach, with agents who are totally ignorant (see Arthur et al. (1997). However, this would imply that they would somehow generate a set of rules with which they would experiment. This pushes the analysis back many stages to a very fundamental level. What is done in most agent based models is to provide the agents with a set of rules and simply note that this, to some extent, conditions the outcomes of the process.

Still we are faced with the criticism that artificial markets, are not "scientific". Let me simply repeat that choosing the best element of a well defined preference order is not necessarily a reasonable assumption when both the order and the set of alternatives are highly complicated, and that something is to be gained from simplifying our account of individuals' behaviour in complicated situations[5]. Whether the specification of the set of rules available to individuals is more ad hoc than the standard assumptions on preferences and technologies is a subject for legitimate debate.

The message here is rather simple. Markets are an important feature of all economies. Each market is characterised by an organisation and structure that will have an impact on the outcomes observed. In general it is difficult to capture all but the simplest feature of such markets in theoretical markets. Even in the case of the simplest markets, those for perishable goods, standard models do not seem to be well adapted to shedding light on the nature of the economic outcomes that one might expect. Curiously enough the particular example which I have mentioned, that of the Marseille fish market does exhibit rather a lot of regularity at the aggregate level. Nevertheless this is not due to individuals behaving, in isolation in a regular way as in the standard competitive model.

5. I have developed this sort of argument at length in Kirman (2006) where I suggest that we have gone down the wrong route in modelling demand.

The complicated organisation of this sort of model breaks any simple link between individual and aggregate behaviour. A number of the specific features of this market, such as the special trading relationships that have developed are difficult to account for in the standard framework.

Artificial or agent based markets are particularly useful in studying situations where the interaction and organisation make simple theoretical analysis too difficult. To repeat, it is not legitimate, I would argue, to dismiss these models as "*ad hoc*". Firstly, one can develop a theoretical model in a very restricted case and then simulate the model to see if the conclusions hold up in a more general case. Secondly, one can use a simplified version of the artificial market in which the solution should be obvious to see if it functions correctly before moving on to the more general framework in which the situation is more difficult to predict. This allows us to do more than simply confirm standard theoretical results but also to detect those features which emerge from the additional structure in the artificial markets. Finally, armed with this information, one can then, return to the empirical data to check them. In this approach therefore there is a constant feedback between the data and the model construction. The data is not just used to test or validate the theoretical model, but plays an active part in its conception and construction.

7. Conclusion

The sort of argument made in this paper in favour of agent based models and computational models in general is often interpreted as an argument against a "scientific approach" to economics. I would argue that this is based on a false notion of what science consists of. Whereas economists have insisted, in the recent past, on a very particular approach to the development of formal economic models it is now time to explore the possibility of different but no less rigorous avenues. As mathematics moves into a more computational mode, economics cannot afford to stand aside and insist on the sort of Bourbakian axiomatics that have dominated the field in recent years. Many leading mathematicians, such as Avi Widgerson at the Institute for Advanced Study in Princeton, have argued forcefully that a computational revolution is taking place in that discipline.

Whereas computer science was regarded as marginal in the field of mathematics in the past, theoretical computer science is now considered to be an integral part of the field. The dismissal of the computational approach by many economic theorists overlooks this and, as I said at the outset, it also forgets that there has been a long and distinguished part of the evolution of economic theory that focused on an algorithmic approach to economic problems. What agent based modelling is doing, is to renew the tradition of using an algorithmic approach to model economic phenomena. Its advantages are clear since it focuses on the essentially dynamic nature of the economy and allows for the explicit introduction of heterogeneity into the models, rather than vainly trying to reduce aggregate activity to that of an individual. None of this is an argument for a less analytical approach to economics but it is an argument for entertaining the possibility of other types of analysis, without any a priori restriction on the field in which they originated. We were wedded to physics and then to mathematics in the past and it seems likely that computer science is more likely to play an increasing role in constructing economic models in the future. Developing and using such models is surely to be preferred to a situation in which theoretical models are abandoned in times of crisis and policymaking reverts to "judgement and experience" alone, to cite Trichet (2010). Agent based modelling is not just an intellectual exercise. As Farmer and Foley (2009) say, « Policy-makers can thus simulate an artificial economy under different policy scenarios and quantitatively explore their consequences. ». Although we will not be able to predict precisely the onset of the next crisis at least we may be better prepared to deal with it.

References

Aboulafia M., 1997, *Making Markets: Opportunism and Restraint on Wall Street*. Cambridge: Harvard University Press.

Aoki M., 1996, *A New Approach to Macroeconomic Modelling*. New York: Cambridge University Press.

Arthur B. *et al.*, 1997, "Asset Pricing Under Endogenous Expectations in an Artificial Stock Market". in *The Economy as an Evolving Complex System II*, Edited by W. B. Arthur, S.N. Durlauf, and D. Lane, Addison Wesley.

Blume L., 1993, "The Statistical Mechanics of Strategic Interaction". *Games and Economic Behaviour*, 5: 387-424.

Buiter W., 2009, "The Unfortunate Uselessness of Most `State of the Art' Academic Monetary Economics" *Financial Times*, March 3.

Bush R.-R., & F. Mosteller, 1955, *Stochastic Models for Learning.* New York: Wiley.

Camerer & Ho, 1999, *Experience-Weighted Attraction learning in normal-form games*. Econometrica, 67: 827-873.

Durlauf, S., 1993, "Non Ergodic Economic Growth". *Review of Economic Studies*, 60: 349-366.

Duffy J. 2008, "Macroeconomics: A Survey of Laboratory Research". Chapter prepared for the *Handbook of Experimental Economic,s* Vol. 2, J. Kagel and A.E. Roth, Eds.

Epstein J. M., 2007, *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton: Princeton University Press.

Evans G.W., & Honkapohja, S., 2001, *Learning and Expectations in Macroeco- Nomics*. Princeton and Oxford: Princeton University Press.

Farmer J. D., & D. Foley, 2009, "The Economy Needs Agent-Based Modelling". *Nature*, 460: 685-686.

Goyal S., 2007, *Connections: An introduction to the economics of networks*. Princeton, NJ: Princeton University Press.

Greenspan A., 2003, "Corporate governance". Remarks at the 2003 Conference on Bank Structure and Competition, Chicago, Illinois May 8.

Haldane A., 2009, "Rethinking the financial network". Speech delivered at the Financial Student Association, Amsterdam

Jackson M., 2008, *Social and economic networks*, Princeton, NJ: Princeton, University Press.

Jovanovic, B., 1987, "Micro Shocks and Aggregate Risk". *Quarterly Journal of Economics*, 102: 395-409.

Kirman A., 1992, "What or whom does the representative individual represent?". *Journal of Economic Perspectives*, 6 (2): 117-136

———, and N. Vriend, 2000, "Evolving Market Structure: A Model of Price Dispersion and Loyalty". *Journal of Economic Dynamics and Control,* 459-502.

———, 2006, "Demand Theory and General Equilibrium: From Explanation to Introspection, a Journey down the Wrong Road". Annual supplement, *History of Political Economy*.

Kuhn J., 2005, "On Today's Business Ethics". EPIC, New York: Columbia University.

Lucas R., 1980, "Methods and Problems in Business Cycle Theory". *Journal of Money Credit and Banking* 12(4): 696-715.

Lucas R., 1988, "Adaptive Behaviour and Economic Theory". *Journal of Business*, 59: 401-426.

Mantel R., 1974, “On the characterisation of aggregate excess demand”. *Journal of Economic Theory,* 7: 348-53.

Miller J. & S. Page, 2007, *Complex Adaptive Systems*. Princeton: Princeton University Press.

Nirei M., 2006, “Threshold Behavior and Aggregate Critical Fluctuations”. *Journal of Economic Theory,* 127: 309-322.

North D., 1977, “Markets and other allocation systems in history”. *Journal of European Economic History,* 6: 703–16.

Robbins L., 1935, *An Essay on the Nature and Significance of Economic Science*. London Macmillan.

Sapio S, A. Kirman & G. Dosi, eds, 2011, “Special Section on Fish Markets”. *Journal of Economic Behavior and Organisation*, 30: 1-68.

Scheinkman J. A. & M. Woodford, 1994, “Self-Organized Criticality and Economic Fluctuations”. *The American Economic Review* Vol. 84, 2: 417-421

Simon H, 1962, “The Architecture of Complexity”. *Proceedings of the American Philosophical Society* 106: 467-482

Solow R., 2007. “Reflections on the Survey”. In David Colander, *The Making of an Economist, Redux*. Princeton: Princeton University Press.

Sonnenschein H., 1972, “Market excess demand functions”. *Econometrica*, 40: 549-63.

Vega-Redondo F., 2007, *Complex Social Networks*, Econometric Society”. Monograph Series, Cambridge: Cambridge University Press.

Vitali S., J-B. Glattfelder & S. Battiston, 2011, *The Network of Global Corporate Control*. PLoS ONE 6(10): e25995. doi:10.1371.

Weisbuch G., A. Kirman & D. Herreiner, 2000, "Market Organisation and Trading Relationships. *Economic Journal*, 110: 411-436.

MACROECONOMICS IN A SELF-ORGANIZING ECONOMY*

Quamrul Ashraf
Williams College
Boris Gershman
American University
Peter Howitt
Brown University

This paper emphasizes the importance of considering the mechanisms that coordinate economic transactions in a decentralized economy, namely the role played by a self-organizing network of entrepreneurial trading firms, for theories aimed at guiding macroeconomic policy. We review a research program that aims to understand how, and how well, trading activities are coordinated in various circumstances by employing agent-based computational (ACE) models of stylized economies where these activities take place in a self-organizing network of markets created and operated by profit-seeking business firms. We discuss how such a research program can yield important policy-relevant insights, beyond those that can be offered by conventional dynamic stochastic general equilibrium (DSGE) models, into several macroeconomic phenomena including the emergence of monetary equilibria in a decentralized economy, the microfoundations of the multiplier process, the costs of a higher trend rate of inflation, and the role of the banking system in economic crises.

Keywords: Self-organizing trade networks, Market coordination, Agent-based computational economics, Emergence of monetary equilibria, The multiplier process, Costs of inflation, Bank regulation, Economic crises.

* We thank participants of the REPLHA International Conference, Milan, October 13, 2011, for helpful comments. The research reported on here was partly funded by the NSF.

This paper is largely about macroeconomic theory, but it is motivated by some of the most important policy challenges that we are confronted with in the extraordinary world we are living in today. Dealing effectively with these challenges requires a conceptual framework that focuses on those parts of the economic system that matter most for the question at hand. We are not alone in thinking that such a conceptual framework is not provided by the theories that guide macroeconomic policy in most countries today. What we aim to do in this paper is to describe a line of research that we have been pursuing, one that goes back to joint work between Howitt and his former teacher and co-author, Robert Clower, and is aimed at providing a more appropriate conceptual framework for thinking about some important macroeconomic policy issues.

Our starting point is one of the oldest and most important ideas in economics, going back at least to Adam Smith, namely the idea that a decentralized economic system is self-organizing. It is capable of "spontaneous order," in the sense that a globally coherent pattern of transactions can result from purely local interactions, without the intervention of a central coordinator. Indeed, like an anthill, a free market economy can organize transactions into patterns that are beyond the comprehension of any of its individual participants. We would like to understand how this self-organization takes place. Specifically, what is the process that coordinates the exchange activities of millions of independent transactors in a decentralized economy?

The reason why these questions are critical for understanding macroeconomic policy is that an economy's coordination mechanism works better at some times than others. Even Smith and Hayek recognized that the automatic workings of the decentralized economy could sometimes be improved by collective intervention. Consider, for example, the increase in unemployment that takes place during a deep recession. Unemployed workers who used to be employed are just as willing and able to work as before, the fall in aggregate output that accompanies recession has enhanced the scarcity value of the output they could potentially produce if employed, and yet the market for their services has somehow disappeared. The coordination mechanism that had previously allowed them and those with a taste for their output to realize their potential gains from mutually advantageous exchange is no longer allowing them to do this, even though those gains are if anything larger than before.[1] Macroeconomic policy to deal with unemployment thus amounts to fixing a mechanism that has malfunctioned, and a highly complex mechanism at that. And attempts to fix this broken mechanism without first understanding how it is supposed to work normally are likely to be as successful as medieval medicine was in treating bacterial infections.

The main premise of our research has been that the role of coordinating transactions in a decentralized economy is performed, for better or worse, by a self-organizing network of business firms that seek profit from creating and operating the markets through which others transact. To use a phrase that Clower once coined, business firms are the visible fingers of the invisible hand. Economics has a long tradition of regarding exchange as a do-it-yourself affair, in which people with goods and services to sell trade directly with the ultimate demanders of those goods and services. But a little reflection on the experience of daily life is enough to persuade most people that exchange in a market economy is not a do-it-yourself affair. People are not like the actors in a typical monetary search model, who when hungry go wandering aimlessly in hopes of randomly encountering someone with surplus food. They go to a grocery store or to a restaurant. When in search of clothing they visit a tailor or a clothing store. They lend surplus funds through the intermediation of a bank, arrange for long-distance travel by using facilities provided by a travel agent, and so on and so forth. Most of us also sell our labor services to an economic entity, either a private business or a government agency (the latter of which would not exist in a purely decentralized economy) whose primary purpose is

---

1. Of course there are some macroeconomists who would claim that unemployment can only arise when productivity falls or tastes change toward more leisure, in which case there is no malfunction; it's not that the gains from trade are unexploited but rather that the gains have disappeared. But this is a view that we do not accept. Recessions are not periods of technological regress or of contagious laziness. Although there may be shifts in demand from one sector to another, or changes in the overall level of aggregate demand, there is nothing inherent about such shifts that would imply that the gains from trade have shrunk in all sectors of the economy, yet in the typical recession unemployment rises in all sectors of the economy. Instead, it seems clear to us that some kind of market failure takes place in recessions.

to purchase various such services, organize them into production, and sell the resulting output. So to understand how, and how well, exchange activities are coordinated in a decentralized economy, the first place to look is to this self-organizing network of firms that constitutes the institutional structure through which we all conduct our daily business.

Now some would argue that a good economic theory involves abstraction, and that if we want to model the transactions process of a modern economy we should maybe abstract from business firms, by assuming that people who work for businesses trade their output directly with those having a taste for the workers' output. This is the stance taken by search theoretic models of money, which typically assume that trade takes place directly between ultimate suppliers and ultimate demanders, who meet in random non-repeated encounters without the aid of any intermediary. Presumably the rationale for this way of looking at the transactions process is the same as the rationale for abstracting from money in the theory of value. On the surface, what we see is people trading goods and services for money, but the deeper underlying reality that we see once we pierce the veil of money is that people are ultimately trading goods and services for other goods and services, with money acting only as a device for executing these ultimate exchanges.

The analogy between money and firms is a useful one. But as John Stuart Mill once observed, there is nothing more insignificant than money, except when it goes wrong. By the same token, the fact that people find it convenient to trade through shops rather than directly with one another is perhaps of little significance for understanding the long-run structure of relative prices. But when something goes wrong with the network of firms that people normally rely upon, then abstracting from the existence of such firms is as unhelpful as it would be to ignore money when trying to understand inflation.

Moreover, a good case can be made that, by recognizing that production takes place in firms but not recognizing their role in coordinating transactions by creating and operating markets, we are ignoring their most important activity, as measured by the value of resources they use. Wallis and North (1986), for instance, have shown that more than 50 percent (by value) of the primary resources used up in the course of economic activity in the United

States are devoted not to production activities but to providing transactions services, services that would be of no use to a Robinson Crusoe with no trading partners. The value added in Finance, Insurance, and Real Estate, for example, is typically much more than that of the entire manufacturing sector, as is the value added in Retail and Wholesale Trade. Moreover, much of the input to the manufacturing sector is best construed as being used up in the production of transactions that help people realize gains from trade rather than being used up in transforming inanimate objects. We have in mind the inputs of lawyers, sales people, those engaged in personnel, marketing, and advertising, and so forth, all of whom are undertaking activities whose main purpose is to facilitate transactions.

We made reference at the start of this paper to the theories now guiding macroeconomic policy. We were referring, of course, to the broad class of rational-expectations equilibrium models generally known as DSGE, for dynamic stochastic general equilibrium. DSGE started out 4 decades ago as a reaction against the Keynesian economics that had been the dominant paradigm of the profession in the 1950s and 60s, a reaction that was first expressed by new classical economists, like Lucas and Sargent, and later by real business cycle theorists. The early proponents of DSGE argued that an equilibrium model built on a slightly modified Walrasian conceptual framework, in which markets clear everywhere and always, could account for important short-run as well as long-run macroeconomic phenomena. But soon Keynesian economists began developing their own versions of DSGE, which consisted of rational expectations equilibrium models in which not all prices were perfectly flexible, and by now DSGE has become the dominant paradigm agreed upon by all sides of the great macroeconomic debates.

Of course there are many serious criticisms one might make of DSGE, and many of them have been made in the literature. The criticism we consider most important for present purposes is that existing DSGE models, even those with imperfectly flexible prices, are built on a conceptual foundation that pays little or no attention to the way in which economic transactions are organized. To borrow a phrase from Jevons, they ignore the institutions and processes that make up the mechanism of exchange.

When we examine DSGE models, looking for what else might go wrong with the market mechanisms that coordinate economic transactions, we find that in most of them there is no such mechanism. In models with perfect competition, the setting of prices is left to a mysterious outside agent called the auctioneer, whose behavior is left largely unexplained. But that is just the tip of the iceberg. There is no description of how trades are arranged. Even if we accept that the auctioneer can provide everyone with a price vector such that the sum of desired demands equals total supply for each tradable object, there is no account of how buyers and sellers are matched up with one another and how the trades that people have planned will be executed. When demand and supply are not equal, the theory offers no guidance as to who gets to trade how much and with whom, no indication of how people learn about trading opportunities, about who creates and maintains the shops and other facilities through which they trade, about how bids and offers are transmitted, and so on and so forth.

The canonical model of Woodford (2003), which forms the basis of the estimated New Keynesian DSGE models, now used in central banks around the world, makes less use of the mysterious auctioneer, inasmuch as many prices are set by a given set of monopolistically competitive firms who are explicitly motivated to maximize their shareholders' wealth. So far, so good. But, there is no account of where these price-setting monopolists come from, how they maintain their monopolies against the threat of entry, how people decide to trade with one set of firms rather than another, how firms manage to coordinate with their suppliers and customers, what happens when one of them goes out of business in a recession, and so forth. Instead, all transactors are in continuous touch with each other through the intermediation of these firms, whose existence is merely assumed, and who take care of enough details of the transactions process that the other people in the model are connected only through the market prices that they take as given from the firms. As a result, there is nothing that can go wrong in the transactions process other than some mistake in price-setting.

In essence, these New Keynesian DSGE models are providing the same diagnosis that economists have given for generations; unemployment rises because wages and prices are slow to adjust to shifts in demand and supply. This is the answer provided by classical economists from Hume through Marshall. It is still the answer offered by modern Keynesian economics. Indeed it is now even the answer that has been finally accepted by most proponents of the real-business-cycle school of macroeconomics, who admit the need for wage-and-price stickiness to account for various features of the business cycle.[2]

The problem with this time-honored tradition of blaming wage-and-price stickiness is not that the assumption of stickiness is factually incorrect. On the contrary, the stickiness of wages and prices is one of the most well-documented facts of macroeconomics. Instead, as Leijonhufvud (1968) forcefully pointed out, the problem is that, first, the experience of the Great Depression in the United States shows clearly that the downturn that started in 1929 did not come to an end until wages and prices started to rise, that is, until the reflation that was clearly a deliberate policy move on the part of the Roosevelt administration started to take place. If lack of wage and price flexibility had caused the downturn, then it would have taken deflation rather than reflation to cure the unemployment problem. Second, as Keynes argued in Chapter 19 of the *General Theory*, and as Fisher had already argued in *Debt Deflation Theory of Depressions*, there are many reasons for believing that wage and price flexibility would actually make fluctuations in unemployment larger rather than smaller.

So when unemployment rises in a recession, something has gone wrong with the process by which economic transactions become organized, something that goes beyond the mere stickiness of wages and prices, something that we think can only be discovered by investigating simple stylized models of economies in which trading activities take place, in and out of equilibrium, in a self-organizing network of markets that are created and operated by profit-seeking business firms, and by asking how, and how well, those activities are coordinated in various circumstances. What we would like to do in this paper is to give the reader an idea of what kind of model that research has led us to construct, and why we think this class of models provides a more solid framework for analyzing certain policy questions than does any DSGE model currently in use.

2. See Chari and Kehoe (2006) for example.

It turns out that this research agenda is one for which Agent-Based Computational Economics (ACE) is particularly well suited for two main reasons. First, by endowing each agent with a set of relatively simple adaptive behavioral rules that allow the agent to operate intelligently in an unknown environment, an ACE model gives the system a chance to achieve some semblance of order without giving anyone the kind of systemic knowledge that would allow him to act as a central coordinator. A rational expectations equilibrium might or might not emerge from the interaction of these rules. If it does, then we have discovered something about at least one possible mechanism that produces that kind of spontaneous order whereas, if the system fails to approximate a rational expectations equilibrium, we will have discovered something about the conditions under which a spontaneous order is likely to require some kind of collective intervention. The second reason for using ACE is that models of spillovers between multiple markets that are not in supply-demand equilibrium are notoriously difficult to analyze. The attempts by Barro and Grossman, Benassy, Malinvaud, and others in the 1970s to understand what some called "general disequilibrium analysis," building on the original contributions of Clower and Patinkin, made little progress largely because the problem quickly became analytically intractable. ACE can deal with this kind of intractability by substituting simulation and Monte-Carlo results for unattainable analytical results.

1. Self-organization of trading firms

At the heart of all our work is a parable concerning the spontaneous emergence of a more-or-less self-regulating network of markets operated by profit-seeking business firms. The details of this parable were first laid out in the form of an ACE model by Howitt and Clower (2000). The rest of this section briefly describes the Howitt-Clower model of a self-organizing economy.

Time comes in discrete "weeks" indexed by t. There are n non-storable goods and m households. In the computer simulations, $n = 10$ and $m = 2160$. Each household is of a type (i, j), where i and j are distinct goods; such an agent receives a weekly endowment of one unit of good i and wishes to consume only good j. There is a symmetric distribution of types in the population, with the same number b (= 25 in the simulations) of each type of household.

It is assumed that households can trade only through organized trading facilities called "shops." A new shop is created whenever a household chooses to act on a random opportunity to enter business (*i.e.*, to become an "entrepreneur"). Each shop trades only two goods, which we assume must be the endowment and consumption goods of the "owner" (the entrepreneur who opened the shop). Once opened, a shop of type (i, j) will make weekly postings of two offer prices—an amount of i offered for each unit of j delivered by anyone who chooses to do so, and an amount of j offered for each unit of i delivered. Those households seeking to trade any good for which they do not own a shop must form a "trading relationship" with such a shop. They form such relationships by searching, both directly through the space in which the shops are located and indirectly by querying randomly met people about the shops with which they have relationships.

Each week the computer simulation takes the households through a fixed routine involving 5 stages, each of which represents what we take to be an important aspect of the trading process. First, there is an entry process in which a random selection of households is faced with an opportunity to become an entrepreneur. There is a fixed setup cost of opening the shop, and there will be a fixed weekly overhead cost of maintaining the shop in operation if opened. So before deciding whether or not to open, the potential entrepreneur will conduct "market research" by querying two other households—one that might want to trade i for j and one that might want to trade j for i, asking each if they would choose to form a trading relationship with the proposed shop if it were to open, at the posted prices that the entrepreneur has decided on (more on these prices below). Both households will answer the query using the same criterion (to be described below) they will use when searching for better trading opportunities in the next stage of the program. If both answer affirmatively, then the potential entrepreneur will indeed open. Otherwise, the opportunity will be allowed to lapse.

In choosing his shop's prices, the potential entrepreneur uses a simple form of full-cost markup pricing. First, an estimate is formed of how much of each good will be delivered by the shop's suppliers/customers. Each of these initial estimates is taken from a uniform distribution between *1* and *X*, where *X* represents the state of "animal spirits." The entrepreneur then calculates, given the fixed costs of operation and the margin required to compensate for the sunk cost of setting up the shop, the combination of offer prices that would just allow the shop to break even in terms of each of the goods it trades if its estimates were correct. The bigger the delivery estimates, the bigger those breakeven offer prices will be because of the economy of scale implicit in the fixed costs. These breakeven prices are the prices the shop will post.

The second stage of weekly activities is one in which households search for trading relationships. Each household can have only two relationships at a time—one with a shop (outlet) trading the household's endowment good and the other with a shop (source) trading its consumption good. In some cases, the same shop can serve as both source and outlet (double coincidence). The household wants to maximize weekly consumption. In the double coincidence case, weekly consumption good will be the shop's offer price for the household's endowment good. Otherwise, if the household has a source and outlet both trading the same complementary good, then it can engage each week in indirect exchange using the common complementary good as an exchange intermediary, and weekly consumption will be the product of the outlet's offer price for the endowment good and the source's offer price for the complementary good. In all other cases, weekly consumption will be zero. During the search stage, a household always forms a new relationship with any shop that would allow weekly consumption to be raised at currently posted prices. Whenever a new source or outlet is chosen, the relationship with the old source or outlet must be severed.

The third stage of the Howitt-Clower model is a trading stage, in which each household, in random order, visits the shops with which it has a relationship, first delivering its endowment to the outlet and then, if possible, using the entire sales proceeds to buy the consumption good from the source. During this stage it is assumed that all planned trades can be executed by the shops, regardless of the shop's inventory position.

The fourth stage of this process is the exit stage. During the course of trading, a shop trading good *j* will have had an amount *x* of good *j* delivered to it and will have had to pay out a certain amount of good *j* to its suppliers of the other good it trades. In addition, it will have to pay out some fixed amount to cover its overhead cost. The remainder is available for the owner's consumption. If this is negative, then the owner will have had to engage in home production of the good just to stay in business. This home production, or negative consumption, represents a loss that the shopowner does not want to incur indefinitely. Any shop that has incurred a loss in either of the goods it trades will exit at this stage with some fixed probability.

The fifth and final stage of weekly activity is where expectations and prices are updated by each shop. Delivery expectations are adjusted using the simplest form of adaptive expectations. That is, the expectation for each delivery amount is adjusted by some fixed fraction of the gap between what was actually delivered this week and what had been expected this week. Prices are then updated using the same full-cost markup procedure used when the shop opened, but with the newly revised delivery estimates.

The model sketched above can be described as a stochastic process. Provided that animal spirits are not too large and that fixed costs are not too large, the process will possess several absorbing states. These absorbing states correspond to stable shop networks that provide a coherent pattern of trading activity throughout the system.

One such absorbing state is a stationary barter equilibrium, in which there exists in operation exactly one shop for each unordered (*i, j*) pair. Each household of type (*i, j*) either owns the shop of that type or has a trading relationship with it, and delivers its unit endowment to it each week in exchange for consumption. Delivery estimates of each shop equal *b*, the number of households of each type, and prices are constant at the values that allow the shops to break even at those estimates.

Other absorbing states are stationary monetary equilibria, in which one good *k* has emerged as a universal medium of exchange, being traded in each shop, and in which there is exactly one shop trading each other good. Every non-shopowner who is not endowed with *k* and does not consume *k* engages in indirect exchange using *k* as an exchange intermediary. The others are able to consume using just one shop. Delivery estimates have adjusted to equal actual amounts, and prices are constant.

In order for both barter and monetary absorbing states to exist, it is necessary that animal spirits not be too large. This ensures that the state is never disturbed by new entrants who pass the market research test during entry because their delivery estimates are so optimistic that they can undercut the prices charged by incumbents.

Howitt and Clower show that the monetary equilibrium in which fixed costs are lowest (a shop's fixed costs are assumed to vary according to the identity of the goods traded in the shop) allows the maximal feasible total consumption. As in the barter stationary equilibrium, since each household has an outlet, all endowment is used either for consumption or for defraying a shop's fixed costs. The total number of fixed costs is minimized in a monetary equilibrium since it uses only n-1 shops as opposed to the $n(n-1)/2$ shops needed to support the barter equilibrium, and using the least cost exchange intermediary obviously allows the most to be consumed of the money commodity.

Howitt and Clower show that this model is capable of self-organization under a wide set of parameter values. Specifically, they report the results of 6,000 simulation runs. Each run starts in autarky, with no shops and hence no trading relationships in place, and continues for 20,000 weeks or until a monetary equilibrium has been reached. They find that in almost all runs an absorbing state is reached unless fixed costs were set too high or too low.[3] Furthermore, they find that the only absorbing states that ever emerge are monetary equilibria. This latter result arises from the "network externality" implicit in the above sketch. That is, once a few shops trading the same complementary good have emerged, the survival chances of other shops are greatly enhanced if they also trade that complementary good since this allows them to attract more suppliers/customers and hence makes it easier for them to defray the fixed cost while offering competitive prices. The fact that the same model that was capable of "growing" market organization also happened to exhibit a feature of all economies of record that orthodox theories have trouble accommodating without artificial

3. The problem with low fixed costs appears to be that they weaken the network externality that helps to promote the achievement of the efficient monetary equilibria.

contrivances, namely monetary exchange, adds to the plausibility of this model's account of self-organization.

2. The multiplier process

One implication of the above account of self-organization is that there is more that can go wrong with the trading process than just having disequilibrium prices. In particular, the failure of a business firm will always disrupt the economy at least locally for some time, and there is nothing much that speedy price adjustment can do to compensate for this shock because the disappearance of a shop constitutes a loss of organizational capital, which can only be rectified by the successful entry of a new replacement shop regardless of the prices that are charged by the surviving shops.

Indeed, the random nature of the entry process can lead to a shakeout period that makes failure more likely for other firms, causing a cumulative contraction or multiplier process in the economy. This is because a supplier of shop A that disappears will also be the customer of some other shop B. If several of those suppliers suddenly lose their source of income, then shop B will be faced with a drop in demand that threatens its existence. If it was on the verge of making losses, then this might be enough to put it under as well.

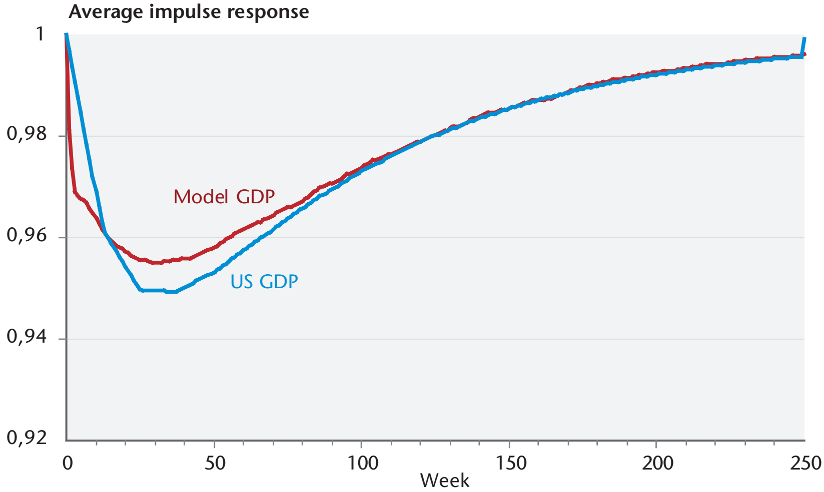
Howitt (2006) showed more systematically how this can happen in a slightly modified version of the Howitt-Clower model. In this modified model, it was assumed that a convention had already been reached that good *1* was the universal medium of exchange. Thus an entrepreneur of type (i, j) can only open a shop of type $(i, 1)$. The model was started in a monetary equilibrium and subjected to a reallocative shock of the sort that occurs when people reduce their demands for some products, without immediately signaling to anyone what they are planning to demand instead of these products. This is the classic coordination problem that Keynes wrestled with. Consumers may decide to spend less than their income, but this does not amount to a specific demand for future consumption. Instead, their future demands remain latent, and entrepreneurs must somehow discover them through trial and error. Likewise, unemployed workers' notional demands remain undiscovered until some entrepreneurs find it in their interest to employ the workers and thereby provide them with the means of making their demands effective.

To portray such a shock in the above system, Howitt supposed that at a certain date some fraction of the population switches from consuming one good to another. To preserve the aggregate structure, he supposed that the total number of each type remains constant, so that for every *i*-consumer that becomes a *j*-consumer there is a *j*-consumer that switches to *i*. At the time of this shock, each switcher is suddenly without a source, and his former source loses a customer. The switcher may continue to sell his endowment to his outlet but he does not spend his sales proceeds. GDP falls because of the reduced goods consumption of the switchers that no longer show up to their former sources, and because of the reduced good-1 consumption of the entrepreneurs whose operating surplus in good *1* suddenly falls.

Figure 1 below shows the average impulse response, over 10,000 runs, of aggregate real GDP relative to the full-capacity level achieved in the initial equilibrium, in the case where 12 percent of the population made the switch. The blue curve in this figure represents the actual impulse-response of HP-filtered GDP under the univariate AR(2) model estimated by Chari, Kehoe, and McGrattan (2000), assuming 50 weeks per year. As it turns out, this simple model, where deviation amplification comes just from the cumulative process of shop failures, does a good job of tracking the actual hump-shaped impulse response pattern of US GDP!

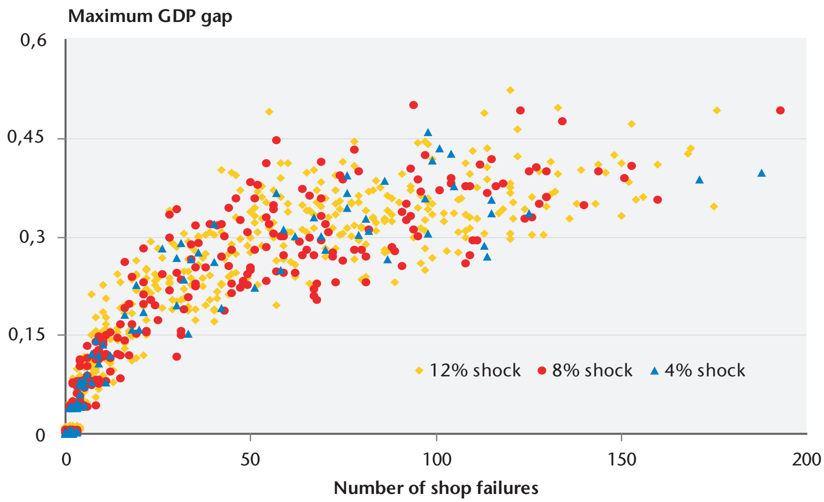
Howitt verified that this hump-shaped impulse response pattern arises from the cumulative process of shop failures by showing that if shop exits were eliminated (by modifying the code to set the exit probability of a loss-making firm equal to zero), then the economy always snapped back into equilibrium within a short number of weeks. Moreover, as Figure 2 demonstrates, there was a strong positive correlation between the size of the displacement in GDP and the number of shop failures in the 5 years following the shock. He also demonstrated that allowing for greater wage-and-price flexibility by having a higher speed of adaptation in delivery estimates did nothing systematically to reduce the amplitude of the impulse response.

Figure 1. Impulse response of real GDP to a 12 percent shock



Source: Howitt, 2006.

Figure 2. Maximum GDP gap and shop failures following the shocks



Source: Howitt, 2006.

3. Costs of inflation

Ashraf, Gershman, and Howitt (2012) have built a somewhat less stylized version of the Howitt-Clower model to analyze one of the important policy issues that has proven particularly intractable for orthodox theory. Specifically, they address the question of the extent to which an economy's macro performance is enhanced by having a lower trend rate of inflation.

According to conventional theory, the answer to this question is "not much." The case for low inflation in modern macro theory comes from various DSGE studies that have confirmed the optimality of Friedman's rule, which is to reduce inflation to the point where the nominal rate of interest equals zero. But the saving that would arise in principle from following this rule consists of the elimination of a tax on non-interest-bearing money holdings, a saving that almost all published research estimates to be a trivial fraction of GDP because the base of this tax is just a tiny fraction of total wealth in any advanced economy.

New Keynesian DSGE models, in which money as a means of exchange and store of value plays no essential role, offer another possible reason for targeting low inflation, namely the inefficiency that comes from having a wider dispersion of relative prices for no reason other than the fact that different sellers are at different stages of the price-change cycle; those with more recent price changes will tend to have higher relative prices because they have made the most recent adjustment to inflation. In these models, the optimal trend rate of inflation is clearly zero, except possibly for second-best public finance reasons (Phelps, 1972) or risk-sharing considerations (Levine, 1991) that might argue for a positive rate.

Howitt and Milionis (2007) show that this argument is especially dependent on the Calvo pricing model that everyone agrees is particularly unconvincing. Specifically, in a conventionally calibrated model, once the trend inflation rate reaches 10 percent, over 35 percent of aggregate output is produced by the 0.3 percent of firms that are selling at a price below marginal cost! These firms would certainly want to either raise their price or curtail production if it were not for the fact that they have not recently been visited by the Calvo fairy, but the model requires them anyway to produce however much is demanded at their obsolete prices. Replacing the

Calvo model by a Taylor model with as much as a 7 quarter lag between price changes gets rid of this counterintuitive feature of the model and has no firms selling below marginal cost, but it also reduces the cost of a 10 percent rate of inflation to about 1.5 percent of aggregate consumption. Moreover, if one keeps the assumption of Calvo pricing but reinserts lagged inflation in the Phillips Curve, as central bank DSGE models typically do, by invoking the usual indexation story that price setters not visited by the Calvo fairy adjust their prices as a function of lagged inflation, then the cost of inflation in DSGE models is almost entirely eliminated, because indexation greatly reduces the extent to which inflation raises price dispersion (Billi, 2011).

Despite the failure of conventional theory to account for significant costs, central bankers around the world continue to attach the highest priority to maintaining very low inflation. Before concluding that central bankers are being wrongheaded, we believe that one needs to explore non-conventional theoretical reasons why inflation might be costly. In particular, we have used a much-expanded version of the Howitt-Clower model to explore the suggestion by Heymann and Leijonhufvud (1995) to the effect that inflation impedes the coordination mechanism that Howitt and Clower focus on and that conventional theory takes as functioning perfectly at all times at no cost.

In particular, the results of the preceding section concerning the cumulative process of shop failures suggest an unconventional mechanism through which inflation might really matter for macro performance. Specifically, the higher is the rate of inflation the more difficult it is for the firms that operate markets to remain in business, because of the well-known tendency of inflation to induce noise into the price system. Thus, an environment with higher inflation is likely to have a higher incidence of such cumulative contractions and, hence, a worse overall macro performance.

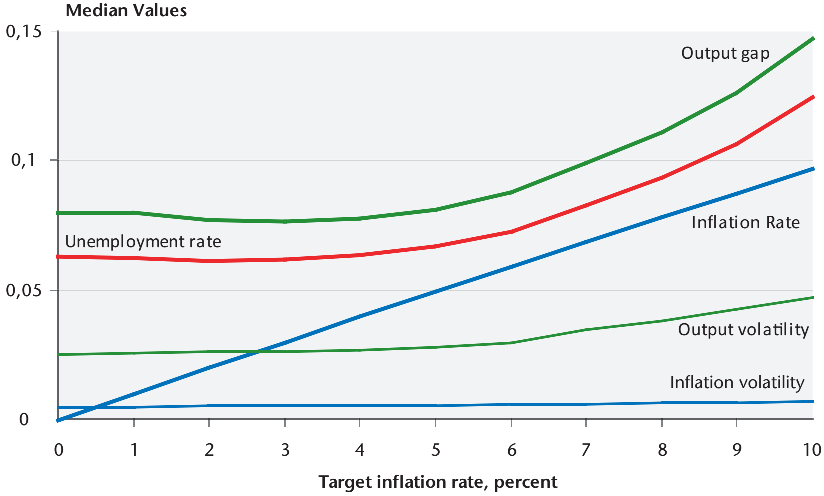
The changes made by Ashraf, Gershman, and Howitt (2012) include making the goods durable, allowing each household to have two consumption goods, introducing fiat money instead of commodity money, having staggered price setting (but making it state dependent instead of having a Calvo-type Poisson process delivering price change opportunities), having government bonds and a central bank that conducts open market operations using a Taylor rule, having a retail sales tax at a rate adjusted annually to pay the interest on the government debt, and having continual shocks and high enough animal spirits that there is no absorbing state for the economy. Moreover, parameter values were (roughly) calibrated to US economic data. In this model, peoples' endowments of any good *j* are interpreted as a type of labor services capable of producing good *j*, so that people without an outlet can be interpreted as unemployed.

The trend rate of inflation in the model is the inflation target implicit in the central bank's Taylor rule. In the baseline calibration, this inflation target was set equal to 3 percent, and the economy achieved almost exactly 3 percent inflation on average across all runs and all years. To address the issue of the macro consequences of trend inflation, the paper simulated the artificial economy 10,000 times for each integer value of the inflation target from 0 to 10 percent. The main results are depicted in Figure 3 below. As shown in this figure, the median performance of the economy deteriorates steadily by all reported measures when trend inflation rises above 3 percent. It also shows that this deterioration is highly significant in economic terms when trend inflation reaches the 10 percent level. We know of no conventional analysis that provides such powerful support for the idea that a central bank can improve an economy's performance simply by choosing a low inflation target.

Ashraf, Gershman, and Howitt use the model to explore the reasons for this effect of inflation, and find strong evidence that the above mentioned link between inflation, price dispersion, and shop failure was at work. In particular, as Figure 4 shows, increases in trend inflation produce monotonic increases in price dispersion and monotonic decreases in the median number of shops in existence.

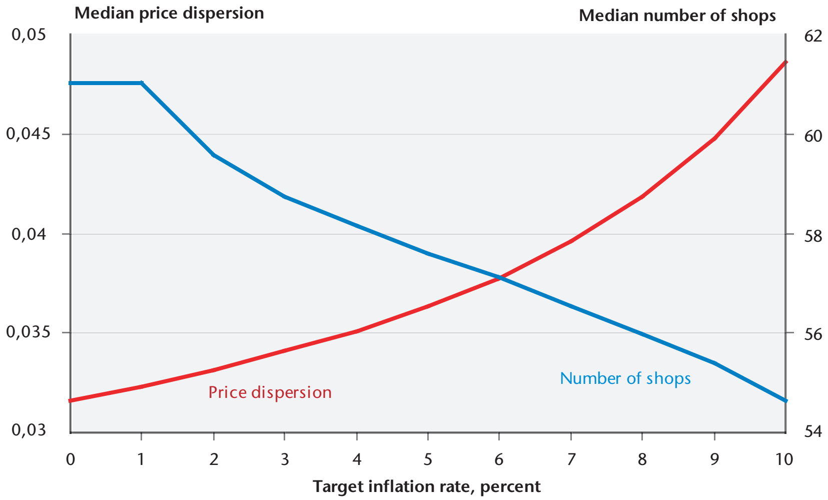
They also show that the effects work even if the zero lower bound on nominal interest rates is suppressed and even if efforts are made to take into account the Lucas critique by allowing critical parameters in peoples' decision rules, like the markup parameter in price setting, to vary systematically with the inflation rate. In short, it seems that, according to this particular ACE model, inflation does create a big macroeconomic cost by impairing the self-organizing capacity of the economic system.

Figure 3. Economic performance and target inflation



Source: Ashraf, Gershman and Howitt, 2012.

Figure 4. Price dispersion, number of shops, and target inflation



Source: Ashraf, Gershman and Howitt, 2012.

4. Banks and economic crises

In Ashraf, Gershman, and Howitt (2011), we explored the role of banks in a self-organizing economy. Banks are of course a kind of shop in themselves, being intermediaries between buyers and sellers of credit. But they also play a critical role in the creation and destruction of the other shops in an economy's trading network by providing or withholding finance. In this paper, we explored the effect that banks have on macro performance, much as we did in the other paper where we explored the effects of trend inflation.

The model used in this paper was a further extension of the one used to study inflation. We added a fixed number of banks, who make commercial loans to shops, with full recourse, secured by collateral in the form of fixed capital and inventories. The banks charge a fixed spread over their deposit rate, and always allow customers to borrow up to a regulatory maximum loan-to-value ratio. The banks also invest in government bonds. Households hold their non-money wealth in the form of bank deposits, which bear the same interest rate as government bonds. Banks are also subject to capital requirements, and are sanctioned when their capital is inadequate. The government in this extended model acts not just as central bank and fiscal authority but also as bank regulator and deposit insurer. When a bank is found to be insolvent, the government injects enough capital to make all deposits good and to restore its capital adequacy, and finds a new owner for the bank, much as the FDIC now does routinely for small banks that fail in the United States.

What this paper shows is that banks matter a great deal for economic performance. The baseline model is calibrated to US data and simulated 10,000 times, first with and then without banks. The difference is remarkable. With banks, the median simulation run had an annual average unemployment rate of 5.9 percent. Without banks it was 11 percent. Similarly, the volatility of the output gap was 2.8 percent with banks and 6.2 percent without.

We also explored in this paper one of the aspects of economic performance that almost always escapes orthodox DSGE analysis, namely the prospect that an economy can perform reasonably well most of the time but can, on occasion, go completely out of control. On average, across all 10,000 runs, the rate of inflation and the output gap were roughly constant after a 20-year initialization period. In the case of most individual runs, the output gap never deviated by more than 20 percent from its average value. But in a small fraction of runs (about 5 percent) the economy at some point diverged radically from this pattern and exhibited wild fluctuations of the output gap, with GDP falling in some cases even to zero.

In an attempt to contribute to the literature on “macroprudential” regulation, we used this model to explore the ways in which bank regulation affects the economy. In particular, we ran experiments in which we varied either the maximal loan-to-value ratio (LTV) or the capital adequacy ratio (CAR). What we found surprised us somewhat in two senses. First, in terms of median results, neither of these regulatory parameters seemed to matter. For example we compared the median results of our baseline model, with LTV = 0.5 and CAR = 0.08, to those from a “risky” scenario in which LTV = 0.9 and CAR = 0.02. The median results were quite similar across scenarios in terms of all macro indicators except bank failures. Thus, it seems that in “normal times” the economy is not much affected by bank regulations.

But when we sorted the 10,000 runs by average GDP over the 40 years of each run, we found another surprise. In the worst decile of runs (the 10 percent of runs with the lowest average GDP), we found that the regulatory parameters mattered a lot, and that in fact these “bad times” were a lot *less bad* under the “risky” scenario! Thus, what would normally seem like risky bank regulation from a microprudential point of view turned out to alleviate the problems that arise in those small number of cases where the economy was spinning out of control. For example, the average unemployment rate in the worst decile of runs was 7.2 percent in the “risky” scenario *versus* 8.9 percent in the baseline, and output volatility was 3.4 percent *versus* 5.2 percent.

Exploring the source of this last surprise, we discovered that it worked the way it did because banks provide the self-organizing economy with not just a financial accelerator, as emphasized by the literature started by Williamson (1987) and Bernanke and Gertler (1989), but also with a financial stabilizer. That is, when the economy is starting to become disorganized it has a critical need for entrepreneurship to replace shops that have failed. Bank finance is essential for facilitating this entrepreneurship. But when banks are heavily regulated they are less able to provide this finance just when it is most needed, that is, just when firms are running short of unpledged collateral and banks are finding themselves short of capital. The more of their capital they are able to devote to financing entrepreneurship and the less collateral entrepreneurs are required to put up, the more banks can play this essential role of averting a cumulative collapse to the economy's trading network.

Of course, all of these results must be heavily qualified by noting in particular that we have postulated banks that do not engage in proprietary trading, and are not influenced by the moral hazard issues associated with too-big-to-fail. Nevertheless, the results do illustrate the potential for properly regulated banks to help prevent an economy from leaving what Leijonhufvud (1973) calls the "corridor" of stability. The results also illustrate the new perspective that can be had from looking at economic fluctuations from the point of view of a self-organizing economy.

References

Ashraf Quamrul, Boris Gershman & Peter Howitt, 2011, *Banks, Market Organization and Macroeconomic Performance: An Agent-Based Computational Analysis*. Unpublished, Brown University.

———, 2012, *How Inflation Affects Macroeconomic Performance: An Agent-Based Computational Investigation*. Unpublished, Brown University.

Bernanke Ben & Mark Gertler, 1989, "Agency Costs, Net Worth, and Business Fluctuations." *American Economic Review,* 79: 14-31.

Billi Roberto M., 2011, "Optimal Inflation for the U.S. Economy." *American Economic Journal: Macroeconomics,* 3: 29-52.

Chari V. V. & Patrick J. Kehoe, 2006, "Modern Macroeconomics in Practice: How Theory is Shaping Policy." *Journal of Economic Perspectives,* 20: 3-28.

Chari V. V., Patrick J. Kehoe, & Ellen R. McGrattan, 2000, "Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem?" *Econometrica* 68: 1151-79.

Heymann Daniel, & Axel Leijonhufvud, 1995, *High Inflation.* Oxford: Clarendon Press.

Howitt Peter, 2006, "The Microfoundations of the Keynesian Multiplier Process." *Journal of Economic Interaction and Coordination* 1: 33-44.

Howitt Peter, & Robert Clower, 2000, "The Emergence of Economic Organization." *Journal of Economic Behavior and Organization* 41: 55-84.

Howitt Peter, & Petros Milionis, 2007, *On the Robustness of the Calvo Price Setting Model*. Unpublished, Brown University.

Leijonhufvud Axel, 1968, *On Keynesian Economics and the Economics of Keynes: A Study in Monetary Theory*. New York: Oxford University Press.

Leijonhufvud Axel, 1973, "Effective Demand Failures." *Swedish Journal of Economics* 75: 27-48. (now *Scandinavian Journal of Economics*).

Levine David K., 1991, "Asset Trading Mechanisms and Expansionary Policy." *Journal of Economic Theory* 54: 148-64.

Phelps Edmund S., 1972, *Inflation Policy and Unemployment Theory*. New York: Norton.

Wallis John J., & Douglass C. North, 1986, "Measuring the Transaction Sector in the American Economy, 1870-1970." In *Long-Term Factors in American Economic Growth*, edited by Stanley Engerman and Robert Gallman. Chicago: University of Chicago Press.

Williamson Stephen D., 1987, "Financial Intermediation, Business Failures, and Real Business Cycles." *Journal of Political Economy* 95: 1196-1216.

Woodford Michael, 2003, *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.

MACROECONOMIC POLICY IN DSGE AND AGENT-BASED MODELS[1]

Giorgio Fagiolo
Sant'Anna School of Advanced Studies
Andrea Roventini
University of Verona, Sant'Anna School of Advanced Studies, OFCE

The Great Recession seems to be a natural experiment for macroeconomics showing the inadequacy of the predominant theoretical framework—the New Neoclassical Synthesis—grounded on the DSGE model. In this paper, we present a critical discussion of the theoretical, empirical and political-economy pitfalls of the DSGE-based approach to policy analysis. We suggest that a more fruitful research avenue to pursue is to explore alternative theoretical paradigms, which can escape the strong theoretical requirements of neoclassical models (e.g., equilibrium, rationality, representative agent, etc.). We briefly introduce one of the most successful alternative research projects—known in the literature as agent-based computational economics (ACE)—and we present the way it has been applied to policy analysis issues. We then provide a survey of agent-based models addressing macroeconomic policy issues. Finally, we conclude by discussing the methodological status of ACE, as well as the (many) problems it raises.

Keywords: Economic Policy, Monetary and Fiscal Policies, New Neoclassical Synthesis, New Keynesian Models, DSGE Models, Agent-Based Computational Economics, Agent-Based Models, Great Recession, Crisis.

> "*Tous les événements sont enchaînés dans le meilleur des mondes possibles*."
> Voltaire, Candide

At the dawn of 2008 a large number of contributions claimed that monetary—and, more generally, economic—policy was finally becoming more of a science (Mishkin, 2007; Galí and

1. Thanks to Herbert Dawid, Antoine Mandel, Mauro Napoletano, Julia Taddei, an anonymous referee and to the participants to the GSDP Workshop, Paris, September 2011 for their comments. All usual disclaimers apply.

Gertler, 2007; Goodfriend, 2007; Taylor, 2007). Almost at the end of the Great Moderation, these authors argued that both the academic world and central banks had finally reached an overall consensus not only on the contingency rules to implement in alternative situations, but also on the fact that "the practice of monetary policy reflects the application of a core set of "scientific principles" (Mishkin, 2007, p.1). These scientific principles, in turn, derived from the so-called New Neoclassical (Goodfriend, 2007; Woodford, 2009) grounded upon Dynamic Stochastic General Equilibrium (DSGE) models[2]. What is more, the available toolbox of economic policy rules was deemed to work exceptionally well not only for normative purposes, but also for descriptive ones. For example, Taylor (2007) argued that "while monetary policy rules cannot, of course, explain all of economics, they can explain a great deal" (p.1) and also that "although the theory was originally designed for normative reasons, it has turned out to have positive implications which validate it scientifically" (abstract). Given these Panglossian premises, scientific discussions on economic policy seemed therefore to be ultimately confined to either fine-tuning the "consensus" model, or assessing the extent to which elements of art (appropriable by the policy maker) still existed in the conduct of monetary policy (Mishkin, 2007)[3].

Unfortunately, as it happened with two famous statements made, respectively, by Francis Fukuyama 1992 about an alleged "end of history", and by many physicists in the recent debate on a purported "end of physics" (see, e.g., Lindley, 1994), these positions have been proven to be substantially wrong by subsequent events. The "perfect storm" which followed the bankruptcy of Lehman Brothers on September 15, 2008 brought financial markets on the edge of collapse causing in turn the worst recession developed economies have ever seen since the Great Depression. In 2012, the risks for the world economic system have not finished yet as the crisis is now menacing the sovereign debt of European countries and the very survival of the Euro.

2. For an introduction, see Clarida *et al.* (1999), Woodford (2003) and Galí and Gertler (2007). *Cf.* also Colander (2006c) for an historical perspective.
3. At the opposite, according to Howitt (2011) "macroeconomic theory has fallen behind the practice of central banking" (p. 2). On the same camp, Mankiw (2006) thinks that macroeconomists should not behave as scientist but as engineers trying to solve practical problems. See also Summers (1991) for an extremely pessimistic view on the possibility of taking *any* economic model seriously econometrically. On these points see also Mehrling (2006).

What is worse, mainstream DSGE-based macroeconomics appear to be badly equipped to deal with the big turmoil we are facing. As Krugman (2011) points out, not only orthodox macroeconomists did not forecast the current crisis, but they did not even admit the possibility of such event and, even worse, they did not provide any useful advice to policy makers to put back the economy on a steady growth path (see also Stiglitz, 2011). On the same line, Delong (2011) reports that when the former U.S. secretary Lawrence Summers was recently asked what economics can offer to understand the crisis, he quoted the works of Bagehot, Minsky and Kindleberger, three dead men whose most recent book is 33 years old. This is so because the DSGE approach "has become so mesmerized with its own internal logic that it has begun to confuse the precision it has achieved about its own world with the precision that it has about the real one" (Caballero, 2010, p. 85).

In that respect, the Great Recessions has revealed to be a natural experiment for economic analysis, showing the inadequacy of the predominant theoretical frameworks. Indeed, an increasing number of leading economists claim that the current "economic crisis is a crisis for economic theory" (Kirman, 2010; Colander *et al.*, 2009; Krugman, 2009, 2011; Caballero, 2010; Stiglitz, 2011; Kay, 2011; Dosi, 2011; Delong, 2011). The basic assumptions of mainstream DSGE models, e.g. rational expectations, representative agents, perfect markets etc., prevent the understanding of basic phenomena underlying the current economic crisis[4].

In this paper, we argue that instead of performing Ptolemaic exercises (Stiglitz, 2011; Dosi, 2011; Caballero, 2010) trying to add additional "frictions" to fix the problems of DSGE models, economists should consider the *economy as a complex evolving system*, *i.e.* as an ecology populated by heterogenous agents whose far-from-equilibrium interactions continuously change the structure of the system (more on that in Kirman, 2010; Dosi, 2011; Rosser, 2011). This is the starting point of agent-based computational economics

4. More precisely, in Section 3 we argue that the DSGE policy apparatus is plagued by a long list of serious problems concerning theoretical issues (*i.e.*, having to do with formal inconsistencies of the model—given its assumptions), empirical difficulties (*i.e.*, related to empirical validation of DSGE models) and political-economy issues (*i.e.*, concerning the absence of any justification for the often unrealistic and over-simplifying assumptions used to derive policy implications). See also Colander (2006b).

(ACE, Tesfatsion, 2006a; LeBaron and Tesfatsion, 2008). Bounded rationality, endogenous out-of-equilibrium dynamics, direct interactions, are the tenets of ACE which allow to catch many of the features of the current crisis (e.g. asset bubbles, resilience of interbank network, self-organized criticality, financial accelerator dynamics, etc; see Section 4 for more details).

On the normative side, due to the extreme flexibility of the set of assumptions regarding agent behaviors and interactions, ACE models (often called agent-based models, ABMs) represent an exceptional laboratory to perform policy exercises and policy design. Indeed, as Section 4 shown, an increasing number of macroeconomic policy applications have been already devised and explored concerning fiscal and monetary policies, bank regulation and central bank independence.

Certainly, also in the ACE approach there are still open issues that should be addressed. The most important ones concern empirical validation, over-parametrization, estimation and calibration. Nevertheless, the success of ACE models in delivering policy implications while simultaneously explaining the observed micro and macro stylized facts are encouraging for the development of a new way of doing macroeconomic theory.

The rest of the paper is organized as follows. Section 1 surveys the approach to policy of the New Neoclassical Synthesis. In Section 2 we discuss the main theoretical and empirical difficulties of DSGE models. In Section 3 we instead introduce the ACE paradigm and in Section 4 we briefly review some policy macroeconomic applications in this field. Section 5 concludes by telegraphically accounting for some methodological issues related to policy in ACE models and the ensuing research avenues that these problems open up.

1. Policy in the DSGE framework

Let us begin by presenting how policy analysis is usually carried out in DSGE models.

The clash between the two competing business cycle theories—the Real Business Cycle (RBC) perspective (see e.g. King and Rebelo, 1999) and the New Keynesian paradigm (*cf.* Mankiw and Romer, 1991)—ended in the last decade with the development of a New Neoclassical Synthesis (NNS)[5]. In a nutshell, the canonical model employed by the NNS paradigm is basically a RBC dynamic stochastic general equilibrium (DSGE) model with monopolistic competition, nominal imperfections and a monetary policy rule (see Clarida e*t al.*, 1999; Woodford, 2003: Galí and Gertler, 2007, for a more detailed exposition of the NNS approach).

In line with the RBC tradition, the starting point of the new vintage models is a stochastic version of the standard neoclassical growth model with variable labor supply: the economy is populated by an infinitely-lived representative household, who maximizes its utility under an intertemporal budget constraint, and by a large number of firms, whose homogenous production technology is hit by exogenous shocks. All agents form their expectations rationally (Muth, 1961). The New Keynesian flavor of the model stems from three ingredients: money, monopolistic competition and sticky prices. Money has usually only the function of unit of account and its short-run non-neutrality is guaranteed by the nominal rigidities introduced by sticky prices. As a consequence, the central bank can influence the real economic activity in the short run by manipulating the interest rate. The RBC scaffold of the model allows one to compute the "natural" level of output and of the real interest rate, that is the equilibrium values of the two variables under perfectly flexible prices. The "natural" output and interest rate constitute a benchmark for monetary policy: the central bank cannot persistently push the output and the interest rate away from their "natural" values without creating inflation or deflation. Note that the assumption of imperfect competition (and of other real rigidities) implies that the "natural" level of output is not socially efficient.

Analytically, the NNS model can be represented by three equations:[6] the expectation-augmented IS equation, the New Keynesian Phillips (NKP) curve, and a monetary policy rule. The expectation-

5. This term was first introduced by Goodfriend and King (1997). Woodford (2003) labeled the approach as "Neo Wicksellian". As stated by Galí and Gertler (2007) the term "New Keynesian" is the most used, even if earlier New Keynesian models were very different from the ones of the New Neoclassical Synthesis.

6. For a formal derivation of the NNS model see Goodfriend and King (1997); Clarida *et al.* (1999); Woodford (2003); Galí (2008).

augmented IS equation constitutes the aggregate-demand building block of the NNS model. Assuming perfect capital markets and taking a log-linear approximation around the steady state, one can derive the IS equation from the goods market-clearing condition and the Euler equation of the representative household:

$$\tilde{y}_t = E_t \tilde{y}_{t+1} - \sigma(i_t - E_t \pi_{t+1} - r_t^n), \tag{1}$$

where $\tilde{y}$ is the output gap (*i.e.*, the percentage gap between real output and its "natural" level), σ is the intertemporal elasticity of substitution of consumption, i is the nominal interest rate, π is inflation, r^n is the "natural" interest rate and E_t stands for the (rational) expectation operator taken at time t. Note that in line with the traditional IS-LM model, the IS equation postulates a negative relation between the output gap and the interest rate gap.

The aggregate-supply building block of the NNS model boils down to a New Keynesian Phillips curve. Starting from the Dixit and Stiglitz (1977) model of monopolistic competition and the Calvo (1983) model of staggered prices (with constant probability of price adjustment), one gets that in any given period firms allowed to adjust prices fix them as a weighted average of the current and expected future nominal marginal cost. The NKP curve can be obtained by combining the log-linear approximation of the optimal price-setting choice, the price index and the labor-market equilibrium:

$$\pi_t = \kappa \tilde{y}_t + \beta E_t \pi_{t+1} + u_t, \tag{2}$$

where β is the subjective discount factor of the representative household and κ depends both on the elasticity of marginal cost with respect to output and on the sensitivity of price adjustment to marginal cost fluctuations (*i.e.*, frequency of price adjustment and real rigidities induced by price complementarities). The term u is usually considered a "cost-push shock": it captures the fact that the natural level of output may not coincide with the socially efficient one for the presence of real imperfections such as monopolistic competition, labor market rigidities, etc. The presence of u implies that inflation does not depend only on the presence of a positive output gap, but also on other factors affecting firms' real marginal costs (the output gap appears in Equation (2) because in the underlying model there is a positive relation between $\tilde{y}$ and the log deviation of real marginal cost from its natural level).

The model just sketched leads to a system of two difference equations (*cf.* Equations 1 and 2) and three unknowns: the output gap, inflation, and the nominal interest rate. In order to solve the system, one has to append a rule to determine the nominal interest rate. This is the role reserved to monetary policy. The choice of the optimal monetary policy rule is usually carried out adopting a welfare criterion: taking a second-order Taylor series approximation of the utility of the representative household, one can derive a welfare loss function for the central bank that is quadratic in inflation and in deviations of output from its socially efficient level (see Woodford, 2010). The NNS model is often closed with "simple" rules such as the Taylor (1993) rule[7] (see Taylor and Williams, 2010, for a survey; more on that in Section 2.3 below):

$$i_t^{\tau} = r_t^n + \varphi_\pi \pi_t + \varphi_y \tilde{y}_t, \tag{3}$$

where i^{τ} is the interest rate target of the central bank, $\phi_y > 0$ and $\phi_\pi > 1$.

Medium scale DSGE models (see e.g. Christiano *et al.*, 2005; Smets and Wouters, 2003, 2007) are usually expanded to account for investment dynamics. Moreover, given the strong interactions between financial markets and the real economy showed by the Great Financial crisis, a new vintage of DSGE models (see e.g. Curdia and Woodford, 2010, 2011; Christiano *et al.*, 2011; Gertler and Kiyotaki, 2010; Gertler and Karadi, 2011) has tried to embody the credit market into the canonical model.

Before performing policy exercises with DSGE models, one should assess their empirical performance and calibrate their parameters. At this stage, different type of shocks (e.g. government spending and private consumption disturbances) are usually added to the model to improve the estimation. Since the assumption of forward-looking agents implies that standard DSGE models are not able to match the econometric evidence on the co-movements of nominal and real variables (e.g., the response of output and infla-

7. Originally, the Taylor rule was just designed for normative purposes, *i.e.* to provide recommendations on how monetary policy should be carried out (Taylor, 2007). Later the Taylor rule assumed a positive role given its good explanatory and predictive power.

tion as to a monetary policy shock is too fast to match the gradual adjustment showed by the corresponding empirical impulse-response functions), they are usually extended introducing a great deal of "frictions"—often not justified on the theoretical ground—such as predetermined price and spending decisions, indexation of prices and wages to past inflation, sticky wages, habit formation in preferences for consumption, adjustment costs in investment, variable capital utilization, etc..

From an econometric perspective, the Equations (1-3) of the DSGE model are naturally represented as a vector auto-regression (VAR) model. The estimation of the resulting econometric model is usually carried out either with a limited information approach or by full-information likelihood-based methods.

Limited information approach. The strategy of the limited information approach to estimate and evaluate DSGE models is usually the following (e.g., Rotemberg and Woodford, 1999; Christiano *et al.*, 2005):[8]

1. Specify the monetary policy rule and the laws of motion for the shocks.
2. Split the parameters in two sets and calibrate the parameters in the first set providing some theoretical or empirical justifications for the chosen values.
3. Fix the timing of the endogenous variables in order to allow the interest rate to respond to contemporaneous output and inflation, while the latter variables are only affected by lagged interest rate. Under this assumption one can estimate via OLS the coefficients of the monetary policy rule and the impulse-response functions of the three variables to a monetary policy shock.
4. Recover the second set of parameters by minimizing the distance between the model-generated and empirical impulse-response functions.
5. Finally, given the structural parameter values and the VAR, identify the other structural shocks by imposing, if necessary, additional restrictions.

The empirical performance of the model is then measured by comparing the impulse-response functions generated by the model with the empirical ones.

8. See also Christiano *et al.* (2010) for a limited information Bayesian approach.

Full information approach. The full information approach was initially discarded to estimate DSGE models because maximum likelihood methods deliver implausible estimates. However, with the introduction of Bayesian techniques, the full information approach regained popularity and it is now commonly employed (see e.g. Smets and Wouters, 2003, 2007). Bayesian estimation is carried out according to the following steps[9]:

1. Place if necessary some restrictions on the shocks in order to allow later identification. For instance Smets and Wouters (2003) assume that technology and preference shocks follow an independent first-order autoregressive process with i.i.d. Gaussian error terms, whereas "cost-push" and monetary policy shocks are i.i.d. Normal white noise processes.
2. Employ the Kalman filter to compute the likelihood function of the observed time series.
3. Form the prior distribution of the parameters by choosing their initial values through calibration, preliminary exploratory exercises, and/or to get some desired statistical properties.
4. Combine the likelihood function with the prior distribution of the parameters to obtain the posterior density, which is then used to compute parameter estimates.

One can then assess the empirical performance of the estimated DSGE model comparing its marginal likelihood[10] with the one of standard VAR models (*i.e.* the Bayes factor) and the model-generated cross-covariances vis-á-vis the empirical ones.

Once one has recovered the parameters of the model by estimation or calibration and has identified the structural shocks, policy-analysis exercises can finally be carried out. More specifically, after having derived the welfare loss function, one can assess the performance of the subset of "simple" policy rules that guarantee the existence of a determinate equilibrium or the more appropriate

9. See Schorfheide (2011) for a discussion of recent advances in the econometrics of DSGE models, current challenges, and possible ways of future research.

10. Following Smets and Wouters (2003), the marginal likelihood of a model A is:

$$\int_{\theta} p(\theta \mid A) p(Y_T \mid \theta A) d\theta ,$$

where $p(\theta \mid A)$ is the prior density for model A and $p(Y_T \mid \theta A)$ is the likelihood function of the observable time series, Y_T conditional on model A and the vector of parameter θ.

parametrization within the class of optimal monetary policy rules. This can be done via simulation, by buffeting the DSGE model with different structural shocks and computing the resulting variance of inflation and the output gap and the associated welfare losses of the different monetary policy rules and parameterizations employed (see e.g. Rotemberg and Woodford, 1999; Galí and Gertler, 2007). In practice, assuming that the DSGE model is the "true" data generating process of the available time series, one is evaluating how the economy portrayed by the model would react to the same structural shocks observed in the past if the monetary policy followed by the central bank were different.

2. Policy with DSGE models: A safe exercise?

DSGE models are plagued by at least three classes of problems which could potentially undermine the usefulness of performing policy-analysis exercises in such a framework. More specifically, DSGE models are subject to theoretical, empirical, and political-economy problems that we shall discuss in the next sections.

2.1. Theoretical issues

From a theoretical perspective, DSGE models are general equilibrium models (GE) rooted in the Arrow-Debreu tradition with some minor non-Walrasian features (e.g., sticky prices). Hence, they share with traditional GE models their same problems and weaknesses. Even if there is a vast and widely-known literature within the neoclassical paradigm dealing with the theoretical issues affecting GE models (see e.g. Kirman, 1989), we briefly recall what are the major problems at hand.

To begin with, sufficient conditions allowing for the existence of a general equilibrium do not ensure neither its uniqueness nor its stability. In addition, the well-known results obtained by Sonnenschein (1972), Debreu (1974) and Mantel (1974) show that one can never restrict agents' characteristics (e.g., endowments, preferences, etc.) in such a way to attain uniqueness and stability. What is more, Kirman and Koch (1986) show that even if agents are almost identical (*i.e.*, same preferences and almost identical endowments), uniqueness and stability cannot be recovered.

In this framework, the strategy followed by neoclassical economists to get stable and unique equilibria is to introduce a representative agent (RA). If the choices of heterogeneous agents collapse to the ones of a representative individual, one can circumvent all the problems stemming from aggregation and provide GE macroeconomic models with rigorous Walrasian micro-foundations grounded on rationality and constrained optimization. However, the RA assumption is far from being innocent: there are (at least) four reasons for which it cannot be defended (Kirman, 1992).[11] First, individual rationality does not imply aggregate rationality: one cannot provide any formal justification to support the assumption that at the macro level agents behave as a maximizing individual. Second, even if one forgets the previous point and uses the RA fiction to provide micro-foundations to macroeconomic models, one cannot safely perform policy analyses with such models, because the reactions of the representative agent to shocks or parameter changes may not coincide with the aggregate reactions of the represented agents. Third, even if the first two problems are solved, there may be cases where given two situations *a* and *b*, the representative agent prefers *a*, whereas all the represented individuals prefer *b*.

Finally, the RA assumption introduces additional difficulties at the empirical level, because whenever one tests a proposition delivered by a RA model, one is also jointly testing the RA hypothesis. Hence, the rejection of the latter hypothesis may show up in the rejection of the model proposition that is being tested. This last point is well corroborated by the works of Forni and Lippi (1997, 1999), who show that basic properties of linear dynamic microeconomic models are not preserved by aggregation if agents are heterogeneous (see also Pesaran and Chudik, 2011). To cite some examples, micro-economic co-integration does not lead to macroeconomic co-integration, Granger-causality may not appear at the micro level, but it may emerge at the macro level, aggregation of static micro-equations may produce dynamic macro-equations. As a consequence, one can safely test the macroeconomic implica-

11. A discussion of the limits of the representative assumption in light of the current crisis is contained in Kirman (2010).

tions of micro-economic theories only if careful and explicit modeling of agents' heterogeneity is carried out.

The fact that solving DSGE models leads to a system of difference equations may potentially add another problem to those discussed above. More specifically, one has to check whether the solution of the system of equilibrium conditions of a DSGE model exists and is determinate. If the exogenous shocks and the fluctuations generated by the monetary policy rule are "small", and the "Taylor principle" holds ($\phi_\pi > 1$, see Equation 3), one can prove existence and *local* determinacy of the rational expectation equilibrium of the DSGE model presented in Section 1 (Woodford, 2003).[12] This result allows one to perform comparative "dynamics" exercises (*i.e.* to compute impulse-response functions) in presence of "small" shocks or parameter changes and to safely employ log-linear approximations around the steady state. Unfortunately, the existence of a *local* determinate equilibrium does not rule out the possibility of multiple equilibria at the *global* level (see e.g. Schmitt-Grohé and Uribe, 2000; Benhabib *et al.*, 2001; Ascari and Ropele, 2009). This is a serious issue because there is always the possibility, for instance if the laws of motion of the shocks are not properly tuned, that the DSGE model takes an explosive path, thus preventing the computation of impulse-response functions and the adoption of the model for policy analysis exercises.

2.2. Empirical issues

The second stream of problems is related to the empirical validation of DSGE models. As remarked by Canova (2008), estimation and testing of DSGE models are performed assuming that they represent the true data generating process (DGP) of the observed data. This implies that the ensuing inference and policy experiments are valid only if the DSGE model mimics the unknown DGP of the data.

As mentioned in Section 1, DSGE models can be represented as a VAR of the form:

$$A_0(\varphi)x_t = H_1(\varphi)x_{t-1} + H_2(\varphi)E_t, \tag{4}$$

12. Of course, also other monetary policy rules different from the Taylor rule (*cf.* eq. 3) can lead to a local determinate rational-expectation equilibrium.

where x are both endogenous and exogenous variables, ϕ is the vector of the parameters of the model and E contains the errors. If the matrix A_0 is invertible, one can obtain a reduced-form VAR representation of the DSGE model.

Following Fukac and Pagan (2006), the econometric performance of DSGE models can be assessed along the identification, estimation and evaluation dimensions. Before going in depth with this type of analysis, two preliminary potential sources of problems must be discussed. First, the number of endogenous variables contemplated by DSGE models is usually larger than the number of structural shocks. This problem may lead to system singularity and it is typically solved by adding measurement errors. Second, H_1 and H_2 are reduced rank matrixes. This problem is circumvented by integrating variables out of the VAR (eq. 4) as long as H_1 and H_2 become invertible. This process leads to a VARMA representation of the DSGE model. This is not an innocent transformation for two reasons: i) if the moving average component is not invertible, the DSGE model cannot have a VAR representation; ii) even if the VAR representation of the DSGE model exists, it may require an infinite number of lags (more on that in Fernandez-Villaverde *et al.*, 2005; Ravenna, 2007; Alessi *et al.*, 2007).

Identification. Given the large number of non-linearities present in the structural parameters (θ), DSGE models are hard to identify (Canova, 2008). This leads to a large number of identification problems, which can affect the parameter space either at the local or at the global level. A taxonomy of the most relevant identification problems can be found in Canova and Sala (2005).[13] To sum them up: i) different DSGE models with different economic and policy implications could be observationally equivalent (*i.e.*, they produce indistinguishable aggregate decision rules); ii) some DSGE models may be plagued by under or partial identification of their parameters (*i.e.*, some parameters are not present in the aggregate decision rules or are present with a peculiar functional form); iii) some DSGE may be exposed to weak identification problems (*i.e.*, the mapping between the coefficients of the aggregate decision rules and the structural parameters may be characterized by little

13. See also Beyer and Farmer (2004).

curvature or by asymmetries), which could not even be solved by increasing the sample size.

Identification problems lead to biased and fragile estimates of some structural parameters and do not allow to rightly evaluate the significance of the estimated parameters applying standard asymptotic theories. This opens a ridge between the real and the DSGE DGPs, depriving parameter estimates of any economic meaning and making policy analysis exercises useless (Canova, 2008). For instance, Schorfheide (2008) finds that the parameters of the New Keynesian Phillips curve estimated in 42 DSGE models published in academic journals range from zero to four. In most of the cases, identification problems can only be mitigated by appropriately re-parameterizing the model.[14]

Estimation. The identification problems discussed above partly affect the estimation of DGSE models. DSGE models are very hard to estimate by standard maximum likelihood (ML) methods, because ML estimator delivers biased and inconsistent results if the system is not a satisfying representation of the data. This turns out to be the case for DSGE models (see the evaluation section) and it helps to explain why ML estimates usually attain absurd values with no economic meaning and/or they are incompatible with a unique stable solution of the underlying DSGE model.

A strategy commonly employed when the DSGE model is estimated following the limited-information approach (*cf.* Section 1) consists in calibrating the parameters hard to identify and then estimating the others. Given the identification problems listed above, Canova (2008) argues that this strategy works only if the calibrated parameters are set to their "true" values. If this is not the case, estimation does not deliver correct results that can be used to address economic and policy questions (see also Canova and Sala, 2005).

As we mentioned in Section 1, Bayesian methods are now commonly employed to estimate DSGE models. They apparently solve the problems of estimation (and identification) by adding a (log) prior function to the (log) likelihood function in order to increase the curvature of the latter and obtain a smoother func-

14. Fukac and Pagan (2006) also argue that identification problems are usually partly mitigated by arbitrarily assuming serially correlated shocks.

tion. However, this choice is not harmless: if the likelihood function is flat—and thus conveys little information about the structural parameters—the shape of the posterior distribution resembles the one of the prior, reducing estimation to a more sophisticated calibration procedure carried out on an interval instead on a point (see Canova, 2008; Fukac and Pagan, 2006). Unfortunately, the likelihood functions produced by most DSGE models are quite flat (see e.g. the exercises performed by Fukac and Pagan, 2006). In this case, informal calibration is a more honest and internally consistent strategy to set up a model for policy analysis experiments (Canova, 2008).

All the estimation problems described above stem also from the fact that DSGE models are not conceived to simplify the estimation of their parameters (Canova, 2008). As a consequence DSGE models put too much stress upon the data, using for instance more unobservable that observable variables (Fukac and Pagan, 2006). This requires strong assumptions about the variances in order to get identification and to employ Kalman filter to obtain the likelihood function. The likelihood functions produced by the Kalman filter are correct only if observations are Gaussian, but macroeconomic time series are typically not normally-distributed (Fagiolo *et al.*, 2008).

Evaluation. Evaluating DSGE models means assessing their capability to reproduce as many empirical stylized facts as possible. For instance, following Fukac and Pagan (2006), one can check: i) whether variables with deterministic trend cotrend; ii) whether I(1) variables co-integrate and the resulting co-integrating vectors are those predicted by the model; iii) the consistency (with respect to data) of the dynamic responses (e.g., autocorrelation, bivariate correlations); iv) the consistency of the covariance matrix of the reduced form errors with the one found in the data; v) the discrepancies between the time series generated by the model and real-world ones. In light of the Great Recession, the last point is particularly important: can DSGE models account for the occurrence of rare large shocks?

Fukac and Pagan (2006) perform such exercises on a popular DSGE model. First, they find that co-trending behaviors cannot be assessed because data are demeaned (a practice commonly followed by DSGE modelers). However, the computation of the technology growth rates compatible with the observed output growth rates shows that the possibility of technical regress is very high. Second, there are no co-integrating vectors, because output is the only I(1) variable. Third, the model is not able to successfully reproduce the mean, standard deviations, autocorrelations, bivariate correlations observed in real data. In addition, the DSGE model predicts the constancy of some "great" ratios (in line with the presence of a steady state of the economy), but this is not confirmed by real data. Fourth, many off-diagonal correlations implied by the covariance matrix of the errors are significantly different from zero, contradicting the DSGE model assumption of uncorrelated shocks. Fifth, the tracking performance of the model depends heavily on the assumed high serial correlation of the shocks.

Finally, DSGE models are not able to account for the occurrence of rare economic crises (see Section 2.3 below for a theoretical explanation). This is not surprising since macroeconomic time series distributions are well approximated by fat tail densities (Fagiolo *et al.*, 2008) and DSGE models typically assume Gaussian distributed shocks.[15] Moreover, Ascari *et al.* (2012) find that even assuming fat-tailed Laplace shocks, the distributions of the time series generated by DSGE models have much thinner tails than those observed in real data.

The results just described seem to support Favero (2007) in claiming that modern DSGE models are exposed to the same criticisms advanced against the old-fashioned macroeconometric models belonging to the Cowles Commission tradition: they pay too much attention to the identification of the structural model (with all the problems described above) without testing the potential misspecification of the underlying statistical model (see also Johansen, 2006; Juselius and Franchi, 2007).[16] In DSGE models, "restrictions are made fuzzy by imposing a distribution on them and then the relevant question becomes what is the amount of uncertainty that we have to add to model based restrictions in order to make them

15. An exception is Curdia *et al.* (2011) where shocks are drawn from a Student-t distribution.
16. On the contrary, the LSE-Copenhagen school follows a macroeconometric modeling philosophy orthogonal to the one followed by DSGE modelers. Scholars of the LSE-Copenhagen approach have concentrated their efforts on improving the statistical model in order to structure data with an identified co-integrated VAR that could then be used to produce stylized facts for theoretical models (Johansen and Juselius, 2006; Juselius and Franchi, 2007).

compatible not with the data but with a model-derived unrestricted VAR representation of the data" (Favero, 2007, p. 29). There are many signals of the potential misspecification of the statistical model delivered by DSGE models: the presence of many persistent shocks, the fact that theory-free VAR models of monetary policy need to include additional variables such as commodity price index to match the data, the absurd estimates produced by standard maximum likelihood estimation, etc. (Fukac and Pagan, 2006; Favero, 2007). If the statistical model is misspecified, policy analysis exercises lose significance, because they are carried out in a "virtual" world whose DGP is different from the one underlying observed time-series data.

2.3. Political-economy issues

Given the theoretical problems and the puny empirical performance of DSGE models, one cannot accept the principles of the positive economics approach summarized by the "as if" argument of Milton Friedman (1953). The assumptions of DSGE models can no longer be defended invoking arguments such as parsimonious modeling or matching the data. This opens a Pandora's box, forcing us to consider how policy-analysis exercises performed with DSGE models are influenced and constrained by the legion of underlying assumptions.

DSGE models presume a very peculiar and un-realistic framework, where agents endowed with rational expectations (RE) take rational decisions by solving dynamic programming problems. This implies that: i) agents perfectly know the model of the economy; ii) agents are able to understand and solve every problem they face without making any mistakes; iii) agents know that all other agents behave according to the first two points. In practice, agents are endowed with a sort of "olympic" rationality and have free access to the whole information set. Note, however, that while rational expectations is a property of the economic system as a whole, individual rationality is not a sufficient condition for letting the system converge to the RE fixed-point equilibrium (Howitt, 2011). It is also unreasonable to assume that agents possess all the information required to attain the equilibrium of the whole economy (Caballero, 2010), especially in periods of strong structural transformation, like the Great Recession, that require policies never tried before (Stiglitz, 2011). In presence of structural breaks, the learning process of agents introduces further non-stationarity into the system preventing the economy to reach an equilibrium state (Hendry and Minzon, 2010). No surprise that empirical tests usually reject the full-information, rational expectation hypothesis (see e.g. Guzman, 2009; Coibion and Gorodnichenko, 2011). Assuming agents behaving according to what suggested by the psychological and sociological evidence allow to build models which better account for macroeconomic phenomena (Akerlof, 2002) including the current crisis (Akerlof and Shiller, 2009).

The representative-agent (RA) assumption prevents DSGE models to address distributional issues, which are one of the major cause of the Great Recession and they are fundamental for studying the effects of policies. Indeed, increasing income inequalities induced households to indebt more and more over time paving the way to the subprime mortgage crisis (Fitoussi and Saraceno, 2010; Stiglitz, 2011). In this framework, redistribution matters and different policies have a different impact on the economy according to the groups of people they are designed for (e.g. unemployed benefits have large multipliers than tax cuts for high-income individuals, see Stiglitz, 2011). Moreover, the RA assumption coupled with the implicit presence of a Walrasian auctioneer, which sets prices before exchanges take place, rule out almost by definition the possibility of interactions carried out by heterogeneous individuals.

Besides being responsible for the problems analyzed in Sections 2.1 and 2.2, the RE and RA assumptions strongly reduce the realism of DSGE models. This is not a minor issue when one has to perform policy analyses (on this point *cf.* also Colander, 2006a, p. 5).

As a consequence of the "as if" methodology, the macroeconomics of DSGE models does not appear to be truly grounded on microeconomics (Stiglitz, 2011). For instance, DSGE models do not take into account the micro and macro implications of imperfect information. Moreover, the behavior of agents is often described with arbitrary specification of the functional forms. The common employed Dixit and Stiglitz (1977) utility function provides a bad description of agents' behavior toward risk. Similarly, the Cobb-Douglas production function is not suited for studying income distribution issues.

More generally, within the Neoclassical-DSGE paradigm there is a sort of internal contradiction. On the one hand, strong assumptions such as rational expectations, perfect information, complete financial markets are introduced *ex-ante* to provide a rigorous and formal mathematical treatment of the problems and to allow for policy recommendations. On the other hand, many imperfections (e.g., sticky prices, rule-of-thumb consumers) are introduced *ex-post* without any theoretical justification only to allow DSGE models to match the data.[17] This process is far from being innocuous, for instance Chari *et al.* (2009) point out that the high level of arbitrariness of DSGE models in the specifications of structural shocks may leave them exposed to the Lucas critiques, preventing them to be usefully employed for policy analysis. Adopting less stringent—but in tune with the microeconomic statistical evidence —assumptions may contribute to jointly solve many empirical puzzles without introducing an army of *ad hoc* imperfections.

There are a couple of other internal inconsistencies which could potentially undermine the reliability of the policy prescriptions developed following the DSGE approach. The first one is related to the marginal role (in the best case) that DSGE models reserve to money and banks. This bizarre situation in which models designed for monetary policy analyses do not include money stems from the simplifying assumption that the representative agent behaves respecting the transversality (or No Ponzi Game) condition, *i.e.* she is perfectly creditworthy and never default (Goodhart, 2009). As a consequence, agents face the same interest rate (no risk premia) and all transactions can be undertaken in capital markets without the need of banks. Moreover, since agents can swap IOUs without facing any credit risk, money has only the function of unit of account and it can be ruled out from DSGE models.[18] The abstraction from default risks does not allow DSGE models to contemplate

17. Citing a very provocative sentence of a famous evolutionary economist, this way of theorizing is like claiming that biology stems from thermodynamics equilibrium with some imperfections.
18. When money is present in the utility function of consumers, the transactions requiring money are assumed to be sufficiently unimportant, so for "reasonable" calibrations, money-augmented DSGE models deliver almost the same results of the standard ones (Woodford, 2003, chapter 2). Of course, the unimportance of transactions requiring money, the calibration reasonability and the quantitative discrepancies between standard and money-augmented DSGE models is debatable and subject to the judgement of policymakers.

the conflict between price and financial stability that Central Banks always face (Howitt, 2011):[19] they just care about the nth-order distortions caused by price misallignments which can eventually result in inflation without considering the huge costs of financial crisis (Stiglitz, 2011). No surprise that DSGE models work fine in normal time but they are unequipped not only to forecast but also to explain the current crisis (Goodhart, 2009; Krugman, 2011).

The second potential inconsistency concerns how business cycles arise in the DSGE framework. DSGE models can be employed to assess the impact of different monetary policies because they are genuine business cycle models. However, the theory of business cycles embedded in DSGE models is exogenous: the economy rests in the steady state unless it is hit by a stream of exogenous stochastic shocks. As a consequence, DSGE models do not explain business cycles, preferring instead to generate them with a sort of deus-ex-machina mechanism. This could explain why even in normal times DSGE models are not able to match many business cycle stylized facts or have to assume serially correlated shocks to produce fluctuations resembling the ones observed in reality (*cf.* Zarnowitz, 1985, 1997; Cogley and Nason, 1993; Fukac and Pagan, 2006).[20] Even worse, the subprime mortgage crisis and the ensuing Great Recession clearly shows how bubbles and, more generally, endogenously generated shocks are far more important for understanding economic fluctuations (Stiglitz, 2011). How policymakers can assess the impact of policies in models not explaining business cycles is an open issue. For instance, the Great Recession revealed that the FED's doctrine about cleaning up afterward asset bubbles bursts was patently wrong.

Moving to the normative side, one supposed advantage of the DSGE approach is the possibility to derive optimal policy rules. However, policymakers adopting optimal policy rules face certain costs—the strict assumptions at the root of DSGE models—but

19. As Howitt (2011) puts it, financial accelerator dynamics have recently been introduced in some DSGE models (see Gertler and Kiyotaki, 2010; Brunnermeier *et al.*, 2011, for a survey) just to analyze how shocks can be amplified by financial considerations without any reference to the price-stability trade-off.

20. The highly persistency of the estimated shock processes and the fact that their path is very akin to the path of one of the observable raises concerns about whether shocks capture aggregate uncertainty or mispecification (Schorfheide, 2011).

uncertain benefits. As argued by Galí (2008), optimal monetary policy rules cannot be used in practice, because they require the knowledge of the "true" model of the economy, the exact value of every parameter, and the real time value of every shocks. Moreover, when the "true" model of the economy and the appropriate loss function are not know, rule-of-thumb policy rules may perform better than optimal policy rules (Brock *et al.*, 2007; Orphanides and Williams, 2008).

2.4. Any ways out?

Given the theoretical and empirical problems of DSGE models discussed above, the positive economics approach advocated by Milton Friedman would suggest to remove or change the plethora of underlying assumptions in order to improve the performance of the model.

This recommendation is reinforced by two related observations. First, the assumptions underlying DSGE models become a sort of strait jacket that preclude the model to be flexible enough to allow for generalizations and extensions. Second, the un-realism of these assumptions prevent policymakers to fully trust the policy prescriptions developed with DSGE models.

It is far from clear why within the mainstream DSGE paradigm there is a widespread conservative attitude with no significative attempts to substitute the "Holy Trinity" assumptions of rationality, greed and equilibrium (Colander, 2005) with more realistic ones. For instance, Akerlof (2007) argues that a broader definition of agents' preferences which take into account the presence of realistic norms can violate many neutrality results of neoclassical economics without recurring to imperfections. Moreover, introducing heterogeneous agents or substituting the rationality assumption with insights coming from behavioral economics could substantially change the working of DSGE models, "making monetary policy more of a science" (Mishkin, 2007).

In any case, if neoclassical economists truly enlist themselves among those advocating an instrumentalist approach to scientific research, they should agree that when models display estimation and validation (descriptive) problems such as those exhibited by DSGE ones, the only way out would be to modify the models' assumptions. *A fortiori*, this should be the recommendation that an instrumentalist researcher would provide if, in addition, the model, as happens in the DSGE case, would also display problems on the normative side.

This is exactly the research avenue that a growing number of scholars have been pursuing in the last two decades. Dissatisfied with standard macroeconomic, micro-founded, general- equilibrium-based neoclassical models like those discussed above, they have begun to devise an entirely new paradigm labeled as "Agent-Based Computational Economics" (ACE).[21] The basic exercise ACE tries to perform is building models based on more realistic assumptions as far as agent behaviors and interactions are concerned, where *more realistic* here means rooted in empirical and experimental micro-economic evidence. For example, following the body of evidence provided by cognitive psychologists (see for example, among a vast literature, Kahneman and Tversky, 2000), the assumptions of perfect rationality and foresight are replaced with those of bounded rationality and adaptive behavior. More generally, ACE scholars share the view that agents in the model should have "the same information as do the economists modeling the economy" (Colander, 2006a, p. 11). Similarly, insights from network theory (e.g., Albert and Barabasi, 2002) and social interactions (e.g., Brock and Durlauf, 2001) suggest to move away from the unrealistic and oversimplifying assumptions concerning agents interactions typically employed in neoclassical models and allow for direct, non-trivial interaction patterns. Finally, the widespread evidence on persistent heterogeneity and turbulence characterizing markets and economies indicate to abandon crazy simplifications such as the representative agent assumption, as well as the presumption that economic systems are (and must be observed) in equilibrium, and to focus instead on out-of-equilibrium dynamics endogenously fueled by the interactions among heterogenous agents.

21. The philosophical underpinnings of ACE largely overlap with those of similar, complementary, approaches known in the literature as "Post Walrasian Macroeconomics" (Colander, 2006b) and "Evolutionary Economics" (Nelson and Winter, 1982; Dosi and Nelson, 1994).The overlap is often so strong that one might safely speak of an emerging "heterodox synthesis". Historically, the first attempt to develop agent-based economics can be traced back to Marshal (Leijonhufvud, 2006).

In other words, ACE can be defined as the computational study of economies thought as complex evolving systems (Tesfatsion, 2006a). Notice that neoclassical economics, on the contrary, typically deals with economies conceived as simple, linear, homogeneous and stationary worlds. It should not come as a surprise that the class of models used by ACE to explore the properties of markets, industries and economies (called agent-based models, ABMs) are far more complicated—and harder to analyze—objects than their neoclassical counterparts. In the following Section we will therefore begin by outlying the basic building blocks of ABMs. Next, we will address the question how ABMs can be employed to deliver normative implications. Then, we will briefly review some examples of policy exercises in ABMs. Some final remarks about pro and cons of using ABMs for policy analysis will be left for the concluding section.

3. Agent-Based Models and economic policy

3.1. Building blocks of ABMs

The last two decades have seen a rapid growth of agent-based modeling in economics. An exhaustive survey of this vast literature is of course beyond the scope of this work.[22] However, before proceeding, it is useful to introduce the main ten ingredients that tend to characterize economic AB models.

1. *A bottom-up perspective*. A satisfactory account of a decentralized economy is to be addressed using a bottom-up perspective. In other words, aggregate properties must be obtained as the macro outcome of a possibly unconstrained micro dynamics going on at the level basic entities (agents). This contrasts with the top-down nature of traditional neoclassical models, where the bottom level typically comprises a representative individual and is constrained by strong consistency requirements associated with equilibrium and hyper-rationality.

22. This and the following subsections heavily draw Pyka and Fagiolo (2007) and Fagiolo *et al.* (2007b). For further details see, among others, Dosi and Egidi (1991), Dosi *et al.* (2005), Lane (1993), Tesfatsion and Judd (2006), Colander (2006a) and Tesfatsion (2006b).

2. *Heterogeneity*. Agents are (or might be) heterogeneous in almost all their characteristics.
3. *The evolving complex system (ECS) approach*. Agents live in complex systems that evolve through time. Therefore, aggregate properties are thought to emerge out of repeated interactions among simple entities, rather than from the consistency requirements of rationality and equilibrium imposed by the modeler.
4. *Non-linearity*. The interactions that occur in AB models are inherently non-linear. Additionally, non-linear feedback loops exist between micro and macro levels.
5. *Direct (endogenous) interactions*. Agents interact directly. The decisions undertaken today by an agent directly depend, through adaptive expectations, on the past choices made by other agents in the population.
6. *Bounded rationality*. The environment in which real-world economic agents live is too complex for hyper-rationality to be a viable simplifying assumption. It is suggested that one can, at most, impute to agents some local and partial (both in time and space) principles of rationality (e.g., myopic optimization rules). More generally, agents are assumed to behave as boundedly rational entities with adaptive expectations.
7. *The nature of learning*. Agents in AB models engage in the open-ended search of dynamically changing environments. This is due to both the ongoing introduction of novelty and the generation of new patterns of behavior; but also on the complexity of the interactions between heterogeneous agents (see point 5 above).
8. *"True" dynamics*. Partly as a consequence of adaptive expectations (*i.e.*, agents observe the past and form expectations about the future on the basis of the past), AB models are characterized by true, non-reversible, dynamics: the state of the system evolves in a path-dependent manner.[23]
9. *Endogenous and persistent novelty*. Socio-economic systems are inherently non-stationary. There is the ongoing introduction

23. This has to be contrasted with the neoclassical approach, where agents hold rational expectations and, as Mehrling (2006, p. 76) puts it, "the future, or rather our ideas about the future, determines the present".

of novelty in economic systems and the generation of new patterns of behavior, which are themselves a force for learning and adaptation. Hence, agents face "true (Knightian) uncertainty" (Knight, 1921) and are only able to partially form expectations on, for instance, technological outcomes.

10. *Selection-based market mechanisms*. Agents typically undergo a selection mechanism. For example, the goods and services produced by competing firms are selected by consumers. The selection criteria that are used may themselves be complex and span a number of dimensions.

3.2. The basic structure of ABMs

Models based on (all or a subset of) the ten main ingredients discussed above typically possess the following structure. There is a population—or a set of populations—of agents (e.g., consumers, firms, etc.), possibly hierarchically organized, whose size may change or not in time. The evolution of the system is observed in discrete time steps, $t = 1,2,\ldots$. Time steps may be days, quarters, years, etc.. At each t, every agent i is characterized by a finite number of micro-economic variables $\underline{x}_{i,t}$ (which may change across time) and by a vector of micro-economic parameters $\underline{\theta}_i$ (that are fixed in the time horizon under study). In turn, the economy may be characterized by some macroeconomic (fixed) parameters Θ.

Given some initial conditions $\underline{x}_{i,0}$ and a choice for micro and macro parameters, at each time step $t > 0$, one or more agents are chosen to update their micro-economic variables. This may happen randomly or can be triggered by the state of the system itself. Agents selected to perform the updating stage collect their available information about the current and past state (*i.e.*, micro-economic variables) of a subset of other agents, typically those they directly interact with. They plug their knowledge about their local environment, as well as the (limited) information they can gather about the state of the whole economy, into heuristics, routines, and other algorithmic, not necessarily optimizing, behavioral rules. These rules, as well as interaction patterns, are designed so as to mimic empirical and experimental knowledge that the researcher may have collected from his/her preliminary studies.

After the updating round has taken place, a new set of micro-economic variables is fed into the economy for the next-step iteration: aggregate variables $\underline{X}_t$ are computed by simply summing up or averaging individual characteristics. Once again, the definitions of aggregate variables closely follow those of statistical aggregates (*i.e.*, GDP, unemployment, etc.).

The stochastic components possibly present in decision rules, expectations, and interactions will in turn imply that the dynamics of micro and macro variables can be described by some (Markovian) stochastic processes parameterized by micro- and macro-parameters. Hoverer, non-linearities which are typically present in decision rules and interactions make it hard to analytically derive laws of motion, kernel distributions, time-t probability distributions, etc. for the stochastic processes governing the evolution of micro and macro variables.

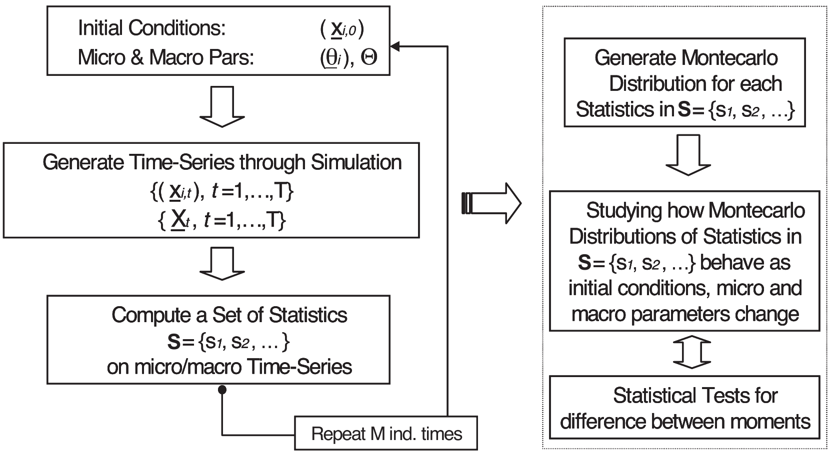
This suggests that the researcher must often resort to computer simulations in order to analyze the behavior of the ABM at hand. Notice that in some simple cases such systems allow for analytical solutions of some kind. Needless to say, the more one injects into the model assumptions sharing the philosophy of the building blocks discussed above (*cf.* Section 3.1), the less tractable turns out to be the model, and the more one needs to resort to computer simulations. Simulations must be intended here in a truly constructive way, e.g. to build and "grow" a society "from the bottom up", in the spirit of object-oriented programming.

3.3. Descriptive analysis of ABMs

When studying the outcomes of ABMs, the researcher often faces the problem that the economy he/she is modeling is by definition out-of-equilibrium. The focus is seldom on static equilibria or steady-state paths. Rather, the researcher must more often look for long-run statistical equilibria and/or emergent properties of aggregate dynamics (that is, transient statistical features that last suffficiently long to be observed and considered stable as compared to the time horizon of the model; see Lane, 1993, for an introduction). Such an exploration is by definition very complicated and it is made even more difficult by the fact that the researcher does not even know in advance whether the stochastic process described by its ABM is ergodic or not and, if it somehow converges, how much time will take for the behavior to become sufficiently stable.

Suppose for a moment that the modeler knows (e.g., from a preliminary simulation study or from some ex-ante knowledge coming from the particular structure of the ABM under study) that the dynamic behavior of the system becomes sufficiently stable after some time horizon T^* for (almost all) points of the parameter space. Then a possible procedure that can be implemented to study the output of the ABM runs as the one synthetically depicted in Figure 1.

Figure 1. A schematic procedure for studying the output of an AB model



Given some choice for initial conditions, micro and macro parameters, assume to run our system until it relaxes to some stable behavior (*i.e.*, for at least $T > T^*$ time steps). Suppose we are interested in a set $S = \{s_1, s_2, \ldots\}$ of statistics to be computed on micro and macro simulated variables. For any given run the program will output a value for each statistic. Given the stochastic nature of the process, each run will output a different value for the statistics. Therefore, after having produced M independent runs, one has a distribution for each statistic containing M observations, which can be summarized by computing its moments.

Recall, however, that moments will depend on the choice made for initial conditions and parameters. By exploring a sufficiently large number of points in the space where initial conditions and parameters are allowed to vary, computing the moments of the statistics of interest at each point, and by assessing how moments do depend on parameters, one might get a quite deep *descriptive* knowledge of the behavior of the system (see Figure 1).

3.4. Model selection and empirical validation

From the foregoing discussion it clearly emerges that in agent-based modeling (as in many other modeling endeavors) one often faces a trade-off between descriptive accuracy and explanatory power of the model. The more one tries to inject into the model "realist" assumptions, the more the system becomes complicated to study and the less clear the causal relations going from assumptions to implications are. ABM researchers are well aware of this problem and have been trying to develop strategies to guide the process of assumption selection. For example, one can try to solve the trade-off between descriptive capability and explanatory power either by beginning with the most simple model and complicate it step-by-step (*i.e.*, the so-called KISS strategy, an acronym standing for "Keep It Simple, Stupid!") or by starting with the most descriptive model and simplify it as much as possible (*i.e.*, the so-called KIDS strategy, "Keep It Descriptive, Stupid!"). A third, alternative strategy prescribes instead to start with an existing model and successively complicate it with incremental additions (this strategy might be labeled TAPAS, which stands for "Take A Previous model and Add Something").

In all these procedures, the extent to which the ABM is able to empirically replicate existing reality should play a crucial role in discriminating the point at which any procedure should stop.[24]

Notice that the very structure of ABMs naturally allows one to take the model to the data and validate it against observed real-world observations. Indeed, an ABM can be thought to provide a DGP, which we think real-world observations being a realization of. More precisely, let us suppose that we believe that observed data are generated by an unknown (to us) colossal DGP, with an almost infinite number of parameters, which we can label as real-world DGP (rwDGP). Suppose further that such rwDGP can be broken in reasonable smaller weakly-exogenous components, each one with a reasonable number of parameters, and each one descri-

24. For a more in-depth discussion of empirical validation in ABMs, we refer the reader to Fagiolo *et al.* (2007a), Pyka and Werker (2009) and papers therein.

bing a small set of variables that we are interested in, on the basis of a small set of other variables. Building an ABM means attempting to approximate one of those small rwDGPs. Due to its stochastic structure, an ABM actually mimics the small rwDGP we are studying by a theoretical DGP that generates the same variables each time we run the model. Of course, we only have one observation generated by the rwDGP, and this makes any inference very difficult (but this has to do with another story, which philosophers call the problem of induction).

Many approaches to empirical validation (and selection) of ABMs can be in principle taken, and the debate is very open here.[25] For example, one might select among ABMs (and within different parameter setups of the same ABM) with respect to the number of stylized facts each of them is able jointly to replicate. A typical procedure to be followed starts with asking whether a particular model can simultaneously reproduce some set of stylized facts for a given parametrization (a sort of "exercise in plausibility"); then explore what happens when the parameter setup changes; finally, investigate if some meaningful causal explanation can be derived out of that step-by-step analysis. Alternatively, one can first select among parameters by calibrating the model (e.g., by directly estimate parameters, when possible, with micro or macro data) and then judge to which extent the calibrated model is able to reproduce the stylized facts of interest. A recent stream of literature tries to recover the parameters of ABMs by indirect estimation (see e.g. Gilli and Winker, 2003; Alfarano et al., 2005; Winker *et al.*, 2007). Notice that, unlike economists supporting the NNS approach— who hold strong theoretical priors rooted in the DSGE model— ACE scholars are more interested in developing plausible theories, which however are not dogmatically deemed to be the "correct" ones (on this point, see also Colander, 2006a).

No matter the empirical validation procedure actually employed, its basic goal is often to restrict the size of the set of free parameters. In fact, over-parameterized models are difficult to interpret and analyze, because no one knows whether the same

25. *Cf.* the special issues edited by Fagiolo *et al.* (2007a) in Computational Economics and by Pyka and Werker (2009) in the Journal of Artificial Socities and Social Simulations.

conclusions could have been obtained in a simpler, less parameterized model. Even if empirical validation allows one to restrict the set of free parameters to a reasonably-sized one, many methodological problems still remain when the model is used to perform policy experiments. If any parametrization represents an alternative world, which one should be employed to assess policy performance? What is the role of initial conditions? We shall briefly come back to these issues in the concluding remarks.

For the moment it is important to notice that the methodological debate within the agent-based community is very lively. Among many interesting lines of methodological research, one of the most crucial ones concerns the issue of realism of the assumptions in economic models (for a more general appraisal, see Schlefer, 2012). Indeed, whereas many ABM scholars argue that their approach allows for more realism in the way individual behaviors and interactions are accounted for in theoretical models (as opposed to neoclassical ones), others have maintained that ABM must as well trade off between successful model building and empirical accuracy of assumptions (Deichsel and Pyka, 2009). Therefore, in order to provide ABMs that deliver meaningful statistical implications, agent-based researchers must often employ assumptions that are not the most descriptively accurate ones.

3.5. Policy experiments in ABMs: Some considerations

ABMs configure themselves as a very powerful device to address policy questions in more realistic, flexible and modular frameworks. Indeed, as far as economic policy is concerned, ABMs have many advantages as compared to neoclassical tools as the DSGE models, which we organize in what follows into two classes: theory and empirics.

Theory. ABMs, contrary to neoclassical ones, do not impose any strong theoretical consistency requirements (e.g., equilibrium, representative individual assumptions, rational expectations). This is because they are not required ex-ante to be analytically solvable. Such no-strait-jacket condition allows for an extremely higher flexibility in model building. If this is coupled with a serious empirical-validation requirement (see below), we are in presence of a semi-instrumentalist approach, where bad (but empirically-plausible) assumptions can be replaced with better (and empirically-plausible) ones if the model does not perform as expected. Notice also that in absence of strong consistency conditions, assumptions can be replaced in a modular way, without impairing the analysis of the model. Indeed, in standard neoclassical models one cannot simply replace the optimization assumption with another one just because the model does not behave well, as that would possibly destroy its analytical solvability. This is not so in ABMs: assumptions—or simply small elements of them—can be taken out of the shelf and easily implemented in the model thanks to the flexibility of computer programming languages.

Empirics. As discussed above, ABMs can be thought as generators of alternative worlds, *i.e.* theoretical DGPs that approximate the unknown one. Contrary to neoclassical models, the structure of ABMs allows to take them to the data more easily. This can be done in two ways. First, one can validate the inputs of ABMs, *i.e.* fine-tune modeling assumptions about individual behaviors and interactions to make them more similar to the observed ones. Second, one can validate the model on the output side, by e.g. restricting the space of parameters, individual behaviors and interactions, and initial conditions to those that allow the model to replicate the stylized facts of interest. This allows for a degree of realism that is much higher than that exhibited by e.g. DSGE models. Furthermore, thanks to the theoretical flexibility discussed above, the set of stylized facts that one can target can include more than one piece of evidence, as instead happens in neoclassical models. In other words, each neoclassical model is typically built—in order to retain analytical solvability—to explain one or two single stylized facts (see the discussion in Aoki, 2006, for more details). On the contrary, each ABM can easily explain a great deal of pieces of empirical evidence at the same time.

But how can one actually conduct policy experiments in ABMs? In a very natural way, indeed. Take again the procedure for ABM descriptive analysis outlined in Figure 1. Recall that micro and macro parameters can be designed in such a way to mimic real-world key policy variables like tax rates, subsidies, interest rates, money, etc. and other key behavioral measures affecting individual incentives in growth, innovation or technologically-related policies. Moreover, initial conditions might play the role of initial endowments and therefore describe different distributional setups.

In addition, interaction and behavioral rules employed by economic agents can be easily devised so as to represent alternative institutional, market or industry setup. Since all these elements can be freely interchanged, one can investigate a huge number of alternative policy experiments and rules, the consequences of which can be assessed either qualitatively or quantitatively (e.g., by running standard statistical tests on the distributions of the statistics in *S*). For example, one might statistically test whether the effect on the moments of the individual consumption distribution (average, etc.) will be changed (and if so by how much) by a percentage change in any given consumption tax rate. Most importantly, all this might be done while preserving the ability of the model to replicate existing macroeconomic stylized facts (e.g. some time-series properties of observed aggregate variables such as persistence of output growth-rate fluctuations, relative standard deviations, cross-correlations, etc.), as well as microeconomic empirical regularities (e.g. firm size distributions, firm productivty dynamics, firm investment patterns, etc.).

4. Macroeconomic policy in ABMs: A survey

Thanks to their flexibility, the number of agent-based models dealing with policy issues is increasing fast over time.[26] This success is partly due to the fact that policy makers appear to be more and more willing to believe in results stemming from detailed simulation models (such as ABMs), where the underlying economic structure can be observed,[27] rather than in general insights produced by quite abstract mathematical models such as DSGE ones.

The number of ABMs addressing policy issues is becoming so large, that a survey of the whole literature would be beyond the scope of this paper. ABMs have indeed been employed in many different policy arenas such as economic growth, industrial dynamics, market design, environmental regulation, traffic management, etc. We then decide to restrict our attention to ABMs

26. See for example the papers contained in the special issue "Agent-Based Models for Economic Policy Design" edited by Dawid and Fagiolo (2008).
27. Moss (2002) discusses the importance of involving the actual decision makers in the process of the generation of agent-based models for policy evaluation.

evaluating the impact of macroeconomic policies in order to assess what the agent-based literature can say on the current Great Recession and to provide a straightforward comparison with DSGE models. More specifically, in what follows we classify agent-based models in four macroeconomic policy areas, namely fiscal policy, monetary policy, bank regulation, and central bank independence.

Fiscal policy. The Great Recession has rewaked interest for employing fiscal policies to tackle economic downturns. An advantage of agent-based models vis-à-vis mainstream ones is the possibility to jointly study the short- and long-run impact of fiscal policies. Dosi *et al.* (2010) try to do so developing an ABM, bridging Keynesian theories of demand-generation and Schumpeterian theories of technology-fueled economic growth (the K+S model). The model is populated by capital-good firms, consumption good-firms, consumers/workers and a public sector. Capital-good firms perform R&D and sell heterogeneous machine tools to consumption-good firms. Consumers supply labor to firms and fully consume the income they receive. The government levies taxes and it provides unemployment benefits. The model is able to endogenously generate growth and business cycles and to replicate an ensemble of stylized facts concerning both macroeconomic dynamics (e.g. cross-correlations, relative volatilities, output distributions) and microeconomic ones (firm size distributions, firm productivity dynamics, firm investment patterns). After having been empirically validated according to the output generated, the K+S model is employed to study the impact of fiscal policies (*i.e.* tax rate and unemployment benefits) on average GDP growth rate, output volatility and unemployment rate. The authors find that Keynesian fiscal policies are a necessary condition for economic growth and they can be successfully employed to dampen economic fluctuations.[28] Moreover, Dosi *et al.* (2012) find a strong interaction between income distribution and fiscal policies: the more income distribution is skewed toward profits, the greater the effects of fiscal policies.

The assessment of alternative uses (demand *vs.* supply policies) of resources collected through taxation is explored in Russo *et al.*

28. More generally, the model of Dosi *et al.* (2010) highlights a strong complementarity between Keynesian policies affecting demand and Schumpeterian policies affecting innovation.

(2007), who develop an ABM where a population of heterogenous, boundedly-rational firms and consumers/workers interact according to random matching protocols. The model delivers sustained growth characterized by fluctuations and reproduce micro and macro regularities such as Beveridge, Phillips and Okun curves, firm growth-rate distributions, etc. On the policy side, they find that average output growth rate is non-monotonically linked to the tax rate levied on corporate profits if revenues are employed to finance R&D investment, whereas growth is negatively affected if the money raised through taxes is employed to provide unemployment benefits.

Finally, the interactions between different expectation-formation mechanisms and fiscal and monetary policies is studied in the ABM developed by Haber (2008). The model is characterized by the presence of households, firms, banks, a government and a central bank, and it is calibrated in order to produce "reasonable" time series for GDP, consumption, unemployment and the inflation rate. The presence of positive fiscal (lower tax rate) and monetary shocks (higher money target) increases GDP growth and inflation and reduce unemployment. The introduction of more sophisticated assumptions about expectations reduce the effects of fiscal policy, whereas it increases the impact of monetary policy.

Monetary policy. DSGE models mostly deal with monetary policy, searching for the best monetary rule. At the same time the current Great Recession has showed that monetary policy alone is not sufficient to put economies back on their steady growth path. Agent-based models can be employed to assess the effects and the limits of monetary policy and to compare the ensuing results with policy prescriptions suggested by DSGE models.

Dosi *et al.* (2012) extend the K+S model introduced above by adding a bank which collects the deposits of firms and provides (costly) loans to financially constrained firms on a pecking-order basis. The model is then employed to assess the effects of monetary policy through changes in the interest rate and the impact of different bank regulatory frameworks (see the section below) preserving its capability to reproduce macro and micro empirical regularities. Simulation results show that higher economic inequalities increase the volatility of output, the unemployment rate and the likelihood of a severe crises, supporting the conjecture advanced by Fitoussi and Saraceno (2010) and Stiglitz (2011) that increasing level of economic inequalities are at the root of the Great Recession. The characteristics of the income distribution also affect the effectiveness of monetary policy. Monetary policy is very effective when inequality is low and interest rates can have significant impact both on output volatility and long-run growth. When income inequality is high, however, the economy is stuck into a liquidity trap where monetary policy is totally ineffective. Similarly, Lengnick (2011) employs a short-run ABM to assess the short- and long-run neutrality of money. More specifically, after checking the capability of the model to reproduce some statistical regularities (e.g. Phillips and Beveridge curves), a series of policy experiments are run by stochastically increasing the money stock. The model shows that money is neutral in the long-run but it affects output in the short-run.

A growing set of agent-based models (Delli Gatti *et al.*, 2005; Oeffner, 2008; Raberto *et al.*, 2008; Mandel *et al.*, 2010) employ Taylor rules to explore the effects of monetary policy on the economy. In this respect, such policy analyses exercises are similar to the ones conducted with DSGE models, but the complexity-rooted approach of ABM can bring new insights. Delli Gatti *et al.* (2005) build an artificial economy populated by firms, workers and a central bank. The latter performs monetary policy employing either a commitment strategy (*i.e.* fixed parameter Taylor rule) or an adaptive, discretionary strategy (*i.e.* the parameters of the Taylor rule change according to a genetic algorithm, mimicking a learning process). Pervasive capital market imperfections imply that monetary policy affects the economy through the credit channel and that money is not neutral in the long-run. Simulation results show that the Taylor principle does not hold and that the adaptive rule outperforms the commitment one according to the standard loss function criterium. Similarly, Raberto *et al.* (2008) compare the effectiveness of a random monetary policy rule vis-á-vis an output gap targeting one, finding that the latter rule can improve social welfare and it outperforms the first one in stabilizing inflation. Oeffner (2008) engages in an accurate input-output empirical validation procedure to study the properties of a monetary ABM which embeds both Keynesian and Wicksellian features (*i.e.* Taylor rule). He finds that monetary policy has real effect also in the medium-run unless the economy is not stuck in a liquidity trap endogenously generated by the model. Mandel *et al.* (2010) develop a multi-sector, heterogenous-agent model, initialized according to input-output tables, and they show that monetary policy performed according to Taylor rule may lead to higher instability in the economy.

Finally, the effects of unconventional monetary policy are explored in Cincotti *et al.* (2010). They develop and ABM based on the EURACE platform to assess the effects of quantitative-easing monetary policy, *i.e.* central bank finances government deficit buying treasury bonds. The EURACE[29] is a large-scale ABM aiming at capturing the main characteristics of the European economy and addressing European policy analyses (Deissenberg *et al.*, 2008; Dawid *et al.*, 2011). Simulation results show that the performance of the economy improves when expansionary fiscal policy and quantitative-easing monetary policy are implemented. However, such expansionary policies raise inflation and lead to higher output volatility in the long-run.

Bank regulation. The flexibility of agent-based models is extremely useful when policy maker want to test the impact of different regulation frameworks on banks' behavior. For instance, one can assess how different regulations affect the liquidity of the interbank payment system or how alternative micro-prudential rules impact on macroeconomic stability. The latter policy question is addressed by Ashraf *et al.* (2011) with an ABM where heterogenous firms interact with banks providing them credit. Banks are subject to various regulations, such as capital-adequacy ratio and limits to loan-to-value ratios. Simulations of the model, calibrated to U.S. data, show that the economy can be hit by "rare disasters", where the behavior of banks strongly affect macroeconomic performance. Banks indeed can be an important "financial stabilizers" of the economy, easing the entry of new firms and avoiding the incumbents to go bankrupt. As a consequence, less strict micro-prudential bank regulation (*i.e.* higher loan-to-value ratios and lower capital-adequacy ratios) allows the economy to recover faster

29. More information are available on http://www.wiwi.uni-bielefeld.de/vpl1/projects/eurace.html. See also http://www.wiwi.uni-bielefeld.de/vpl1/research/eurace-unibi.html for the current Eurace@Unibi model.

from a crisis. Somewhat similarly, Dosi *et al.* (2012) find that in the bank-augmented K+S model, higher loan-to-value ratios positively affect macroeconomic growth when firms can rely less on internal funds. Finally, employing the EURACE model, Raberto *et al.* (2011) find that lower capital-adequacy ratios can spur growth in the short-run, but the higher stock of private debt can lead to higher firm bankruptcies, credit rationing and more serious economic downturns in the long-run.

The modeling of the network structure of an economy has never been embedded in DSGE models. This lack of consideration has prevented these models to explain the emergence, the depth and the diffusion of the current crisis, where the topological properties of the credit market network have a fundamental role. On the contrary, ABMs have started to study the links between alternative network setups and macroeconomic performance. Delli Gatti *et al.* (2010) develop an ABM populated by banks, financially constrained downstream and upstream firms to study the properties of network-based financial accelerator. The topology of the network is continuously evolving because firms can switch their partner trying to finding better credit conditions (*i.e.* lower interest rates). Simulation results show that the interactions of financially constrained agents, occurring through the evolving credit network, give rise to business cycles and to financial crises. Hence, policy makers can try to design a structure for the credit network in order to reduce the magnifying effect of the financial accelerator.

The resilience of the banking system to liquidity shocks is studied by Gai *et al.* (2011) developing an agent-based model of the interbank lending network where heterogenous banks are randomly connected together though unsecured claims and repo activity. The impact of idiosyncratic liquidity shocks are then analized for different network configurations, degrees of connectivity between banks, haircut assumptions, and balance sheet characteristics of financial institutions. The model shows that greater degree of complexity and concentration in the bank network augment the fragility of system, increasing the probability of contagion phenomena and liquidity crises similar to the ones experienced in the Great Recession. Policy experiments show possible ways (e.g. tougher micro-prudential liquidity regulation, countercyclical liquidity requirements) to reduce the network externalities responsible for the emergence of systemic crisis. Similarly, Galbiati and Soramaki (2011) studied the efficiency of the interbank payment system under alternative system configurations. Their model shows that the efficiency of the payment system increases if the number of banks is small and if they are encouraged to provide more liquidity. Moreover, there are strong economies of scale in payment activity (higher volumes reduce total payment costs) calling for higher level of coordination and regulation.

Central bank independence. Agent-based models can be employed to study political economy issues related to the evolution of the institutional role of central banks and to the way monetary policy is announced to the public. Rapaport *et al.* (2009) study why during the nineties many governments decided to delegate authority to their central banks, employing an ABM where heterogenous countries decide whether introducing central bank independence taking into account the behavior of their neighbors. Simulation results, conducted under a Monte Carlo exploration of the parameter space, show that the emergence and the rate of adoption of central bank independence is positively related to the size of the zone of influence of the other countries.

The time-inconsistency problem faced by central banks is analyzed in a more general framework by Arifovic *et al.* (2010) using an ABM where the interaction between a boundedly-rational, evolutionary learning policy maker and a population of heterogenous agents determines the actual inflation rate. The agents can either believe the inflation rate announced by the central bank or employ an adaptive learning scheme to forecast future inflation. The simulations of the calibrated model show that the central bank learns to sustain an equilibrium with a positive, but fluctuating fraction of "believers" and that this outcome is Pareto superior to the equilibrium determined by standard models.

5. Concluding remarks

The subprime mortgage crisis and the ensuing Great Recession have prompted a debate about the state of macroeconomic theory. Certainly, we stand in the camp of those arguing that macroeconomics have entered in a Dark Age (Krugman, 2011). Indeed, as discussed in Section 2, DSGE-based models suffer from a series of dramatic problems and difficulties concerning their inner logic consistency, the way they are taken to the data, the extent to which they are able to replicate existing reality, and the realism of their assumptions. These problems are so deep that impede DSGE models even to conceive the possibility of the current crisis and to propose possible solutions to policymakers. We think that such difficulties are so hard to solve within the neoclassical paradigm that a different research avenue, which attempts to replace the basic pillars of neoclassical economics (rationality, equilibrium, etc.), would be more fruitful.

This alternative paradigm does actually exist and it is called agent-based computational economics (ACE). Section 3 has been devoted to a (necessarily) brief discussion of its philosophical underpinnings, building blocks and policy applications. As our synthetic survey shows (*cf.* Section 4), the number of areas where ACE policy experiments have been already applied with success is rather vast and rapidly increasing. The discussion of Section 4 has also outlined the most prominent values added deriving from performing policy experiments within an ACE approach. These include ACE's extreme modeling flexibility; the friendly relation of agent-based models with empirical data; the easiness of carrying out empirical-validation exercises; the almost infinite possibility of experimentation; and, last but not least, the positive impact that a more realistic and algorithmically-structured model can have on political decision makers—as compared to obscure and un-intuitive mathematical neoclassical models.

Of course, as happens for the New Neoclassical Synthesis, many issues are still far from being settled and the debate is very open. Here, by a way of conclusion, we recall just three of them.

The first issue—which we can label as the problem of over-parametrization—has to do with the role played by micro and macro parameters in ABMs. As mentioned, ABMs are often over-parameterized, for one typically injects in the specification of agents' behavioral rules and interaction patterns many ingredients in order to meet as much as possible what he/she observes in reality. Suppose for simplicity that initial conditions do not matter. Even if empirical validation can provide a way to reduce free parameters, the researchers are almost always left with an ABM whose behavior depends on many free parameters. Many questions naturally arise. How can one interpret these different parameterizations? Which one should be used if one employs the model to deliver policy implications? Should one perfectly calibrate (if possible) the model using the data so that no free parameters are left? Should policy implications be robust to alternative parameterizations instead? Notice that this issue is closely related to a common critique that ABMs usually face: if an ABM contains many free parameters and it is able to reproduce a given set of stylized facts, how can one be sure that it represents the minimal mechanisms capable of reproducing the same set of stylized facts? This point reminds the "unconditional objects" critique in Brock (1999) and it is certainly true for "oversized" ABMs. In practice, however, ACE researchers are well aware of the problem and always try to simplify as much as possible their model by using empirical validation techniques and a KISS or TAPAS approach. Even if it is very difficult to show that a given ABM is the minimal model describing a set of stylized facts, the more stylized facts a model can reproduce, the more one is able to restrict the class of theoretical mechanisms that can do the job.

The second issue concerns the role played by initial conditions. Recall that (if random ingredients are present in the model) any ABM can be considered as an artificial (stochastic) data generation process (*mDGP*) with which we try to approximate the one that generated the data that we observe (*i.e.*, the *rwDGP*). The question is: is the *rwDGP* ergodic or not? If the underlying real-world *rwDGP* is thought to be non-ergodic (as well as the theoretical *mDGP* described in the AB model), then initial conditions matter. This raises a whole host of problems for the modeler. The modeler needs to identify the "true" set of initial conditions in the empirical data, generated by the *rwDGP*, in order to correctly set the initial parameters of the model. Even if *the* "perfect database" would exist, this is a very difficult task. How far in the past does one need to go in order to identify the correct set of initial values for the relevant micro and macro variables? There is a possibility of infinite regress. If this is the case, then one may need data stretching back a very long time, possibly before data started to be collected.

This issue is closely related to a third (and final) one, regarding the relation between simulated and real-world data. While in principle we could generate as many theoretical observations as we like, in practice we may only have a few of such empirical realizations (possibly only one!). If we believe that the empirical observations come from an underlying DGP that could have been "played twice" (*i.e.*, could have generated alternative observations, other than the one we have) the problem of comparing simulated with empirical data becomes very complicated.

It must be said that all three issues above are the subject of never-ending debates among philosophers of science, since they raise fundamental questions related to probability, modeling, inference, etc. (see, e.g., Fagiolo *et al.*, 2007b). As such, they might (and do) affect any stochastic, dynamic (economic) model, DSGE-based ones included. Nevertheless, the large majority of those advocating the New Neoclassical Synthesis approach seems not to care about them. In our view, the fact that they instead occupy center stage in the current ACE debate is another signal of the vitality of this young but promising paradigm.

References

Akerlof G. A., 2002. “Behavioral Macroeconomics and Macroeconomic Behavior." *American Economic Review*, 92: 411–433.

Akerlof G. A., 2007. “The Missing Motivation in Macroeconomics." *American Economic Review*, 97: 5–36.

Akerlof G. A., and R. J. Shiller, 2009. *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism,* Princeton, NJ, Princeton University Press.

Albert R., and A. L. Barabasi, 2002. “Statistical Mechanics of Complex Networks." *Rev. Mod. Phys.*, 4: 47–97.

Alessi L., M. Barigozzi, and M. Capasso, 2007. “A Review of Nonfundamentalness and Identi_cation in Structural VAR models." *Working Paper* 2007/22, Laboratory of Economics and Management (LEM). Sant'Anna School of Advanced Studies, Pisa, Italy.

Alfarano S., T. Lux, and F. Wagner, 2005. “Estimation of Agent-Based Models: The Case of an Asymmetric Herding Model." *Computational Economics*, 26: 19–49.

Aoki M., 2006. “Not More So: Some Concepts Outside the DSGE Framework." in D. Colander, (ed.), *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Arifovic J., H. Dawid, C. Deissenberg and O. Kostyshyna, 2010. "Learning Benevolent Leadership in a Heterogenous Agents Economy." *Journal of Economic Dynamics and Control*, 34: 1768–1790.

Ascari G., G. Fagiolo, and A. Roventini, 2012. "Fat-Tails Distributions and Business-Cycle Models." W*orking Paper Series* 2012/02, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Ascari G., and T. Ropele, 2009. "Trend Ination, Taylor Principle, and Indeterminacy." *Journal of Money, Credit and Banking*, 41: 1557–1584.

Ashraf Q., B. Gershman, and P. Howitt, 2011. "Banks, Market Organization, and Macroeconomic Performance: An Agent-Based Computational Analysis." *Working Paper,* 17102, NBER.

Benhabib J., S. Schmitt-Grohé and M. Uribe 2001. "The Perils of Taylor Rules." *Journal of Economic Theory*, 96: 40–69.

Beyer A., and R. E. A. Farmer, 2004. "On the Indeterminacy of New-Keynesian Economics." *Working Paper Series* 323, European Central Bank, Frankfurt, Germany.

Brock W. A., 1999. "Scaling in Economics: A Reader's Guide." *Industrial and Corporate Change*, 8: 409–46.

Brock W. A., S. Durlauf, J. M. Nason, and G. Rondina, 2007. "Simple Versus Optimal Rules as Guides to Policy." *Journal of Monetary Economics*, 54: 1372–1396.

Brock W. A., and S. N. Durlauf, 2001. "Interactions-Based Models." in J. Heckman and E. Leamer, (eds.), *Handbook of Econometrics* volume 5, Amsterdam, North Holland.

Brunnermeier M. K., T. M. Eisenbach, and Y. Sannikov, 2011. "Macroeconomics with Financial Frictions: A Survey." *mimeo*.

Caballero R. J., 2010. "Macroeconomics after the Crisis: Time to Deal with the Pretense-of-Knowledge SynDrome." *Journal of Economic Perspectives*, 24: 85–102.

Calvo G. A., 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics*, 12: 383–398.

Canova F., 2008. "How Much Structure in Empirical Models?." in T. Mills and K. Patterson, (eds.), *Palgrave Handbook of Econometrics* volume 2, Applied Econometrics, Palgrave Macmillan.

Canova F., and L. Sala, 2005. "Back to Square One: Identiocation Issues in DSGE Models." *Journal of Monetary Economics*, Elsevier, vol. 56(4). 431–449, May.

Chari V. V., P. J. Kehoe, and M. E. R., 2009. "New Keynesian Models Are Not Yet Useful for Policy Analysis." *American Economic Journal: Macroeconomics*, 1: 242–266.

Christiano L. G., M. Eichenbaum, and C. L. Evans, 2005. “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy*, 113: 1–45.

Christiano L. G., R. Motto, and M. Rostagno, 2011. “Financial Factors in Economic Fluctuations." *Working Paper Series,* 1192, European Central Bank.

Christiano L. G., M. Trabandt and K. Walentin, 2010. “DSGE Models for Monetary Policy Analysis", in B. M. Friedman and M. Woodford, (eds.), *Handbook of Monetary Economics,* Volume 3, Elsevier: Amsterdam.

Cincotti S., M. Raberto, and A. Teglio, 2010. “Credit Money and Macroeconomic Instability in the Agent-based Model and Simulator Eurace." *Economics: The Open-Access, Open-Assessment E-Journal*, 4.

Clarida R., J. Gali, and M. Gertler, 1999. “The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature*, 37: 1661–1707.

Cogley T., and J. M. Nason, 1993. “Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model." *Economic Letters*, 43: 77–81.

Coibion O., and Y. Gorodnichenko, 2011. “Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts." *Working Paper Series,* 16537, NBER.

Colander D., 2005. “The Future of Economics: the Appropriately Educated in Pursuit of the Knowable." *Cambridge Journal of Economics*, 29: 927–941.

Colander D., 2006a. “Introduction." in D. Colander, (ed.), *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Colander D., (ed.) 2006b. *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Colander D., 2006c. “Post Walrasian Macroeconomics: Some Historic Links." in D. Colander, (ed.), *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Colander D., H. Folmer, A. Haas, M. D. Goldberg, K. Juselius, A. P. Kirman, T. Lux and B. Sloth, 2009. “The Financial Crisis and the Systemic Failure of Academic Economics." *Technical Report, 98th Dahlem Workshop*.

Curdia V., M. Del Negro, and D. Greenwald, 2011. “Rare Shocks, Great Recessions." *mimeo*.

Curdia V., and M. Woodford, 2010. “Credit Spreads and Monetary Policy." *Journal of Money, Credit and Banking*, 42: 3–35.

Curdia V., and M. Woodford, 2011. “The Central-Bank Balance Sheet as an Instrument of Monetary Policy." *Journal of Monetary Economics*, 58: 54–79.

Dawid H., and G. Fagiolo, (eds.) 2008. “Special Issue on Agent-Based Models for Economic Policy Design." *Journal of Economic Behavior and Organization*, 67.

Dawid H., S. Gemkow, S. van der Hoog, and M. Neugart, 2011. “The Eurace@Unibi Model: An Agent-Based Macroeconomic Model for Economic Policy Analysis." *Working paper*, Bielefeld University.

Debreu G., 1974. “Excess Demand Function." *Journal of Mathematical Economics*, 1: 15–23.

Deichsel S., and A. Pyka, 2009. “A Pragmatic Reading of Friedman's Methodological Essay and What It Tells Us for the Discussion on ABMs." *Journal of Artificial Societies and Social Simulation (JASSS)*, 12: 6.

Deissenberg C., S. van der Hoog, and H. Dawid, 2008. “EURACE: A Massively Parallel Agent-based Model of the European Economy." *Applied Mathematics and Computation*, 204: 541–552.

Delli Gatti D., E. Gaffeo, M. Gallegati, and A. Palestrini, 2005. “The Apprentice Wizard: Monetary Policy, Complexity and Learning." *New Mathematics and Natural Computation*, 1: 109–128.

Delli Gatti D., M. Gallegati, B. Greenwald, A. Russo, and J. Stiglitz, 2010. “The Financial Accelerator in an Evolving Credit Network." *Journal of Economic Dynamics and Control*, 34: 1627–1650.

Delong J. B., 2011. “Economics in Crisis." *The Economists' Voice*, May.

Dixit A., and J. Stiglitz, 1977. “Monopolistic Competition and Optimum Product Diversity.” *American Economic Review*, 67: 297–308.

Dosi G., 2011. “Economic Coordination and Dynamics: Some Elements of an Alternative ’Evolutionary’ Paradigm." *Technical Report*, Institute for New Economic Thinking.

Dosi G., and M. Egidi, 1991. “Substantive and Procedural Uncertainty: An Exploration of Economic Behaviours in Changing Environments." *Journal of Evolutionary Economics*, 1: 145–68.

Dosi G., G. Fagiolo, M. Napoletano, and A. Roventini, 2012. “Income Distribution, Credit and Fiscal Policies in an Agent-Based Keynesian Model." *Working Paper Series* 2012/03, Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Dosi G., G. Fagiolo and A. Roventini, 2010. “Schumpeter Meeting Keynes: A Policy-Friendly Model of Endogenous Growth and Business Cycles." *Journal of Economic Dynamics and Control*, 34: 1748–1767.

Dosi, G., L. Marengo, and G. Fagiolo, 2005. “Learning in Evolutionary Environment." in K. Dopfer, (ed.), *Evolutionary Principles of Economics*, Cambridge, Cambridge University Press.

Dosi G., and R. R. Nelson, 1994. “An Introduction to Evolutionary Theories in Economics." *Journal of Evolutionary Economics*, 4: 153–72.

Fagiolo G., C. Birchenhall, and P. Windrum, (eds.) 2007a. *Special Issue on "Empirical Validation in Agent-Based Models*." In *Computational Economics*, Volume 30, Issue No. 3.

Fagiolo G., A. Moneta, and P. Windrum, 2007b. "A Critical Guide to Empirical Validation of Agent-Based Models in Economics: Methodologies, Procedures, and Open Problems." *Computational Economics*, 30: 195–226.

Fagiolo G., M. Napoletano, and A. Roventini, 2008. "Are Output Growth-Rate Distributions Fat-Tailed? Some Evidence from OECD Countries." *Journal of Applied Econometrics*, 23: 639–669.

Favero C., 2007. "Model Evaluation in Macroeconometrics: From Early Empirical Macroeconomic Models to DSGE Models." *Working Paper,* 327, IGIER, Bocconi University, Milan, Italy.

Fernandez-Villaverde J., J. F. Rubio-Ramirez, and T. J. Sargent, 2005. "A, B, C's, (and D's) for Understanding VARs." *Technical Working Paper,* 308, NBER.

Fitoussi J., and F. Saraceno, 2010. "Inequality and Macroeconomi Performance." *Document de Travail* 2010-13, OFCE-Science Po.

Forni M., and M. Lippi, 1997. *Aggregation and the Microfoundations of Dynamic Macroeconomics*, Oxford, Oxford University Press.

Forni M., and M. Lippi, 1999. "Aggregation of Linear Dynamic Microeconomic Models." *Journal of Mathematical Economics*, 31: 131–158.

Friedman M., 1953. "The Methodology of Positive Economics." in M. Friedman, (ed.), *Essays in Positive Economics*, Chicago, University of Chicago Press.

Fukac M., and A. Pagan, 2006. "Issues in Adopting DSGE Models for Use in the Policy Process." *Working Paper* 10/2006, CAMA.

Fukuyama F., 1992. *The End of History and the Last Man*, London, Penguin.

Gai P., A. Haldane, and S. Kapadia, 2011. "Complexity, Concentration and Contagion." *Journal of Monetary Economics*, 58: 453–470.

Galbiati M., and K. Soramaki, 2011. "An Agent-Based Model of Payment Systems." *Journal of Economic Dynamics and Control*, 35: 859–875.

Galí J., 2008. *Monetary Policy, Ination, and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton, NJ, Princeton University Press.

Galí J., and M. Gertler, 2007. "Macroeconomic Modelling for Monetary Policy Evaluation." *Journal of Economic Perspectives*, 21: 25–46.

Gertler M., and P. Karadi, 2011. "A Model of Unconventional Monetary Policy." *Journal of Monetary Economics*, 58: 17–34.

Gertler M., and N. Kiyotaki, 2010. "Financial Intermediation and Credit Policy in Business Cycle Analysis." in B. M. Friedman and M. Woodford, (eds.), *Handbook of Monetary Economics*, North Holland, Amsterdam.

Gilli M., and P. Winker, 2003. “A Global Optimization Heuristic for Estimating Agent-Based Models." *Computational Statistics and Data Analysis*, 42: 299–312.

Goodfriend M., 2007. “How the World Achieved Consensus on Monetary Policy." *Journal of Economic Perspectives*, 21: 47–68.

Goodfriend M., and R. King, 1997. “The New Neoclassical Synthesis and the Role of Monetary Policy." *NBER Macroeconomics Annual*, : 231–282.

Goodhart C. A. E., 2009. “The Continuing Muddles of Monetary Theory: A Steadfast Refusal to Face Facts." *Economica*, : 821–830.

Guzman G., 2009. “Using Sentiment Surveys to Predict GDP Growth and Stock Returns." in L. R. Klein, (ed.), *The Making of National Economic Forecasts*, Edward Elgar: Cheltenham, 319–351.

Haber G., 2008. “Monetary and Fiscal Policies Analysis with an Agent-Based Macroeconomic Model." *Journal of Economics and Statistics*, 228: 276–295.

Hendry D., and G. Minzon, 2010. “On the Mathematical Basis of Intertemporal Optimization." Economics Series Working Papers 497, University of Oxford.

Howitt P., 2011. “What Have Central Bankers Learned from Modern Macroeconomic Theory?." *Journal of Macroeconomics*.

Johansen S., 2006. “Confronting the Economic Model with the Data." in D. Colander, (ed.), *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Johansen S., and K. Juselius, 2006. “Extracting Information from the Data: a European View on Empirical Macro." in D. Colander, (ed.), *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Juselius K., and M. Franchi, 2007. “Taking a DSGE Model to the Data Meaningfully." *Economics: The Open-Access, Open-Assessment E-Journal*, 1.

Kahneman D., and A. Tversky, (eds.) 2000. *Choices, Values, and Frames*, Cambridge MA, Cambridge University Press.

Kay J., 2011. “The Map is Not the Territory: An Essay on the State of Economics." Technical Report, Institute for New Economic Thinking.

King R., and S. Rebelo, 1999. “Resuscitating Real Business Cycles." in J. Taylor and M. Woodford, (eds.), *Handbook of Macroeconomics*, Elsevier Science: Amsterdam.

Kirman A. P., 1989. “The Intrinsic Limits of Modern Economic Theory: the Emperor Has no Clothes." *Economic Journal*, 99: 126–39.

Kirman A. P., 1992. “Whom or What Does the Representative Individual Represent?."*Journal of Economic Perspectives*, 6: 117–136.

Kirman A. P., 2010. "The Economic Crisis is a Crisis for Economic Theory." *CESifo Economic Studies*, 56:498–535.

Kirman A. P., and K. Koch, 1986. "Market Excess Demand Functions: Identical Preferences and Collinear Endowments." *Review of Economic Studies*, 53: 457–463.

Knight F., 1921. *Risk, Uncertainty, and Profits*, Chicago, Chicago University Press.

Krugman P., 2009. "How did Economics Get it So Wrong?." *New York Times Magazine* : 36–44.

Krugman P., 2011. "The Profession and the Crisis", *Eastern Economic Journal*, 37: 307–312.

Lane D. A., 1993. "Artificial Worlds and Economics, Part I and II", *Journal of Evolutionary Economics*, 3:89–107 and 177–197.

LeBaron B., and L. Tesfatsion, 2008. "Modeling Macroeconomies as Open-Ended Dynamic Systems of Interacting Agents." *American Economic Review*, 98: 246–250.

Leijonhufvud A., 2006. "Episodes in a Century of Macroeconomics." in D. Colander, (ed.), *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Lengnick M., 2011. "Agent-Based Macroeconomics - A Baseline Model." *Working Paper,* 2011/04, Christian-Albrecths-Universitat zu Kiel.

Lindley D., 1994. *The End of Physics*, Basic Books.

Mandel A., C. Jaeger, S. Fuerst, W. Lass, D. Lincke, F. Meissner, F. Pablo-Marti and S. Wolf, 2010. "Agent-Based Dynamics in Disaggregated Growth Models." *Working Paper* 2010.77, CES.

Mankiw G. N., 2006 "The Macroeconomist as Scientist and Engineer." *Journal of Economic Perspectives*, 20:29–46.

Mankiw G. N., and D. Romer, (eds.), *New Keynesian Economics*, Cambridge MA, MIT Press.

Mantel R., 1974. "On the Characterization of Aggregate Excess Demand." *Journal of Economic Theory*, 7:348–353.

Mehrling P., 2006. "The Problem of Time in the DSGE Model and the Post Walrasian Alternative." in D. Colander, (ed.), *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Mishkin F. S., 2007. "Will Monetary Policy Become More of a Science." *Working Paper,* 13566, *NBER*.

Muth J. F., 1961. "Rational Expectations and the Theory of Price Movements", *Econometrica*, 29.

Nelson R. R., and S. G. Winter, 1982. *An Evolutionary Theory of Economic Change*, Cambridge, The Belknap Press of Harvard University Press.

Oeffner M., 2008. "Agent-Based Keynesian Macroeconomics - An Evolutionary Model Embedded in an Agent Based Computer Simulation." *Paper* 18199, MPRA.

Orphanides A., and J. C. Williams, 2008. "Robust Monetary Policy with Imperfect Knowledge." *Journal of Monetary Economics*, 54: 1406–1435.

Pesaran H. M., and A. Chudik, 2011. "Aggregation in Large Dynamic Panels." *Working Paper Series* 3346, CESifo.

Pyka A., and G. Fagiolo, 2007. "Agent-Based Modelling: A Methodology for Neo-Schumpeterian Economics." in H. Hanusch and A. Pyka, (eds.), *The Elgar Companion to Neo-Schumpeterian Economics*, Cheltenham, Edward Elgar Publishers.

Pyka, A. and C. Werker, 2009. "The Methodology of Simulation Models: Chances and Risks." *Journal of Artificial Societies and Social Simulation* (JASSS), 12: 1.

Raberto M., A. Teglio, and S. Cincotti, 2008. "Integrating Real and Financial Markets in an Agent-Based Economic Model: An Application to Monetary Policy Design." *Computational Economics*, 32: 147–162, 10.1007/s10614–008–9138–2.

Raberto M., A. Teglio, and S. Cincotti, 2011. "Debt Deleveraging and Business Cycles. An Agent-Based Perspective", Economics Discussion Paper 2011-31, Kiel Institute for the World Economy.

Rapaport O., D. Levi-Faur, and D. Miodownik, 2009. "The Puzzle of the Discussion of Central-Bank Independence Reforms: Insights from an Agent-Based Simulation", *Policy Studies Journal*, 37: 695–716.

Ravenna F., 2007. "Vector Autoregressions and Reduced Form Representations of DSGE Models." *Journal of Monetary Economics*, 54: 2048–2064.

Rosser B. J., 2011. *Complex Evolutionary Dynamics in Urban-Regional and Ecologic-Economic Systems: From Catastrophe to Chaos and Beyond,* Springer: New York.

Rotemberg J., and M. Woodford, 1999. "Interest Rate Rules in an Estimated Sticky Price Model." in J. Taylor, (ed.), *Monetary Policy Rules,* University of Chicago Press: Chicago.

Russo A., M. Catalano, M. Gallegati, E. Gaffeo, and M. Napoletano, 2007. "Industrial Dynamics, Fiscal Policy and R&D: Evidence from a Computational Experiment." *Journal of Economic Behavior and Organization*, 64: 426–447.

Schlefer J., 2012. *The Assumptions Economists Make*, Harvard, Harvard University Press.

Schmitt-Grohé S., and M. Uribe, 2000. "Price Level Determinacy and Monetary Policy under a Balanced-Budget Requirement." *Journal of Monetary Economics*, 45: 211–246.

Schorfheide F., 2008 “DSGE Model-Based Estimation of the New Keynesian Phillips Curve." *FRB Richmond Economic Quarterly*, Fall Issue: 397–433.

Schorfheide F., 2011 “Estimation and Evaluation of DSGE Models: Progress and Challenges." *Working Paper* 16781, NBER.

Smets F., and R. Wouters, 2003. “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area", *Journal of the European Economic Association*, 1: 1123–1175.

Smets F. and R. Wouters, 2007. “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review*, 97: 586–606.

Sonnenschein H., 1972. “Market Excess Demand Functions." *Econometrica*, 40: 549–556.

Stiglitz J., 2011. “Rethinking Macroeconomics: What Failed, and How to Repair It." *Journal of the European Economic Association*, 9: 591–645.

Summers L., 1991. “The Scientific Illusion in Empirical Macroeconomics." *Scandinavian Journal of Economics*, 93: 129–148.

Taylor J., 1993. “Discretion versus Policy Rules in Practice." *Carnegie-Rochester Series on Public Policy*, 39: 195–214.

Taylor J., 2007. “The Explanatory Power of Monetary Policy Rules." *Working Paper* ,13685, NBER.

Taylor J. B., and J. C. Williams, 2010. “Simple and Robust Rules for Monetary Policy." in B. M. Friedman and Woodford M., (eds.), *Handbook of Monetary Economics*, Volume 3, Elsevier: Amsterdam.

Tesfatsion L., 2006a. “ACE: A Constructive Approach to Economic Theory." in L. Tesfatsion and K. Judd, (eds.), *Handbook of Computational Economics II: Agent-Based Computational Economics*, Amsterdam, North Holland.

Tesfatsion L., 2006b. “Agent-Based Computational Modeling and Macroeconomics." in D. Colander, (ed.), *Post Walrasian Macroeconomics*, Cambridge, Cambridge University Press.

Tesfatsion L., and K. Judd, (eds.) 2006. *Handbook of Computational Economics II: Agent-Based Computational Economics*, North Holland, Amsterdam.

Winker P., M. Gilli, and V. Jeleskovic, 2007. “An Objective Function for Simulation Based Inference on Exchange Rate Data." *Journal of Economic Interaction and Coordination*, 2: 125–145.

Woodford M., 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton, NJ, Princeton University Press.

Woodford M., 2009. “Convergence in Macroeconomics: Elements of the New Synthesis”, *American Economic Journal: Macroeconomics*, 1: 267–279.

Woodford M., 2010. “Optimal Monetary Stabilization Policy." in B. M. Friedman and M. Woodford, (eds.), *Handbook of Monetary Economics* Volume 3, Elsevier: Amsterdam.

Zarnowitz V., 1985. “Recent Works on Business Cycles in Historical Perspectives: A Review of Theories and Evidence", *Journal of Economic Literature*, 23: 523–80.

Zarnowitz V., 1997. “Business Cycles Observed and Assessed: Why and How They Matter." *Working Paper*, 6230, NBER.

RECONSTRUCTING AGGREGATE DYNAMICS IN HETEROGENEOUS AGENTS MODELS
A MARKOVIAN APPROACH[1]

Domenico Delli Gatti
Institute of Economic Theory and Quantitative Methods, Università Cattolica del Sacro Cuore, Milan
Corrado Di Guilmi
School of Finance and Economics, University of Technology, Sydney
Mauro Gallegati
DiSES, Università Politecnica delle Marche
Simone Landini
IRES Piemonte, Turin

The restrictive assumptions imposed by the traditional methods of aggregation prevented so far a sound analysis of complex system of feedback between microeconomic variables and macroeconomic outcomes. This issue seems to be crucial in macroeconomic modelling, in particular for the analysis of financial fragility, as conceived in the Keynesian and New Keynesian literature. In the present paper a statistical mechanics aggregation method is applied to a financial fragility model. The result is a consistent representation of the economic system that considers the heterogeneity of firms, their interactive behaviour and the feedback effects between micro, meso and macro level. In this approach, the impact of micro financial variables can be analytically assessed. The whole dynamics is described by a system of dynamic equations that well mimics the evolution of a numerically solved agent based model with the same features.

Keywords: Financial Fragility, Markov Dynamics, Heterogeneity, Mean-Field Interaction, Master Equation

The Representative Agent (RA) assumption is a methodological shortcut to bypass the problem of dimensionality which arises in heterogeneous agents model. The reasons for dissatisfaction with the RA assumption are well known and have been forcefully

1. The authors gratefully acknowledge the insightful suggestions of an anonymous referee which have improved the quality of the paper and led it to the present version.

discussed in Kirman (1992) and Keen (2011). The efforts to overcome the limits of the exact aggregation (Gorman, 1953) led to methods, such as Lewbel (1992), that are still too restrictive in their basic assumptions to realistically depict an economic system.[2]

As a consequence of the dissatisfaction with the RA approach, a few analytical frameworks have been developed to cope with the dimensionality problem mentioned above. One of the most promising methods has been introduced by Duncan Foley and Masanao Aoki who borrowed from statistical mechanics the concept of mean-field interaction and imported it into economics.[3]

In the mean-field interaction approach, agents are classified into clusters or sub-systems according to their state with respect to one particular feature (the so-called micro-state, e.g. the level of production for a firm on a scale of production levels). This clustering determines the characteristics and the evolution of the aggregate (the macro-state, e.g. the total level of output).[4] The focus is not on the single agent, but on the number or fraction of agents occupying a certain state of a state-space at a certain time. These numbers or fractions are governed by a stochastic law, that also defines the functional of the probability distributions of aggregate variables and, if they exist, their equilibrium distributions. The stochastic aggregation is then implemented through master equation techniques, that allow for a description of the dynamics of probability flows among states on a space. These probability flows are originated by the changes in the conditions of agents and determine the aggregate outcomes.[5]

This paper presents an application of mean-field interaction and master equation on a model in which firms are heterogeneous in terms of financial fragility, along the lines of Di Guilmi *et al.* (2010). The degree of financial fragility, modelled *à la* Greenwald and Stiglitz (1993) (GS henceforth), is the clustering device to classify firms and to develop the analytical solution of the model. The

2. For a review on aggregation methods see Gallegati *et al.* (2006) and Di Guilmi (2008).
3. See Foley (1994); Aoki (1996, 2002); Aoki and Yoshikawa (2006). Further developments of these contributions are: Landini and Uberti (2008), Di Guilmi (2008) and Di Guilmi *et al.* (2011).
4. An early economic application of mean-field theory is Brock and Durlauf (2001).
5. Other applications of master equation in economics, besides the works cited above, can be found in Weidlich and Braun (1992) and Garibaldi and Scalas (2010) among others. Alfarano *et al.* (2008) and Alfarano and Milakovic (2009) offers a further contribution, in particular with reference to agent based pricing models.

analytical approximation mimics well the dynamics of a system with a higher order of heterogeneity and provides insights on the interactions among the micro-units in the system. The analytical solution to agent based models is the result of a functional-inferential method which identifies the most probable path of the system dynamics. The method considers the heterogeneity, representing a large number of agents, and the interaction among them, which originates fluctuations of the macroeconomic variables about a deterministic trend. Individual direct interaction is replaced by indirect mean-field interaction between sub-systems, expressed in terms of the transition rates of the master equations. In particular, according to the local approximation method detailed below, an explicit solution for the master equation is obtained. It yields the analytical identification of an ordinary differential equation, which describes the dynamics of the system trend, and a stochastic differential equation, which quantifies the dynamics of the probability distribution of fluctuations.

The successful application of the aggregation method can be a contribution toward the adoption of a realistic new economic paradigm in the direction suggested by Aoki. As shown in the last section, in fact, the numerical simulation of a similar agent based structure is well reproduced by the stochastic dynamics generated by the master equation.[6]

The structure of the paper is the following: first, we specify the hypotheses for the stochastic structure of the system (section 1) and for the firms that compose it (section 2). In section 3, we develop the framework, setting the dynamical instruments needed for aggregation, and solve the model, determining the two equations that drive production trend and business fluctuations. Section 4 presents a further result coming from the solution of the master equation, stressing the relevance of indirect interaction among agents in shaping macroeconomic outcomes. In section 5, some results of computer simulations are presented. Section 6 concludes.

6. On this point see also Chiarella and Di Guilmi (2011).

1. Stochastic structure

The economy is populated by a fixed number N of firms, each indexed by the subscript i. Agents clusters into micro-states according to a quantifiable individual variable. Two micro-states are defined. State 0 denotes agents characterised by a level of a chosen feature above (or equal to) a certain threshold and 1 labels the state of the rest of the population. In each cluster, therefore, there will be a certain number (the so-called occupation number) of agents. The occupation number of cluster j is N^j, $j = 0,1$. The occupation numbers $(N^0(t), N^1(t))$ define the macro-state of the system. The fraction of firms in micro-state j is $n^j = N^j / N$ where $N^0(t) + N^1(t)$. For the sake of tractability, within each cluster individual levels of a certain variable are approximated by their mean-field values, *i.e.* a specific statistic of the distribution of the variable itself.[7] Therefore, within each cluster heterogeneous agents (characterised by different individual levels) are replaced with an homogeneous agent characterised by this statistic (mean field approximation).

The notation adopted uses a continuous time reference because it is more appropriate for complex systems settings, as remarked among others by Hinich *et al.* (2006). Continuous time functionals are appropriate at system level if we assume that the density of discrete points is large enough within a sufficiently small reference interval of time. This is due to the so-called *principle of limiting density of discrete points*, introduced by Jaynes (1957) to match Shannon's entropy with continuous distributions in information and probability theory.[8] For computational necessity, the numerical simulations must refer to discrete time and, accordingly, occupation numbers, as any other observable, become a discrete time stochastic process.

The probability for a firm of being in micro-state 1 is η: $p(1) = \eta$, hence $p(0) = 1 - \eta$. In order to model the probabilistic flow of firms

7. For example, in our simulations, we adopt the median within each group, as specified in section 5.

8. On this topic other interesting references are Smith (1993) and Milakovic (2001). Besides the *principle of limiting density of discrete points*, modelling discrete time observables with continuous time tools is acceptable when the simulation time, say T, is long enough such that the calendar can be partitioned with sufficiently dense adjacent reference intervals of time of order $o(T)$ w.r.t. the calendar. This conjecture is considered as appropriate due to consistency of analytical trajectories from master equations to experimental simulations.

from one microstate to another, transition probabilities and the transition rates must be defined. The functional specification of transition rates, *i.e.* transition probabilities per vanishing reference unit of time, allows the occupation numbers to be modeled with jump Markov processes.

The transition probability is the probability for a firm to switch from one microstate to the other in a given instant. The transition probability of moving from 0 to 1 is ζ while ι indicates the probability of the opposite transition. The transition rates quantify the probability of observing a jump of one agent from one microstate to another, conditional upon the initial microstate through time. A transition rate is then given by the probability of a firm changing state weighted by the probability of being in one particular starting state. With reference to state 1, the transition rate for entry (from state 0 into state 1) is indicated with λ while the one for exit (from state 1 to state 0) is γ, defined as follows:

$$\begin{aligned} \lambda &= \zeta(1-\eta) \\ \gamma &= \iota\eta \end{aligned} \tag{1}$$

This representation is phenomenological. Indeed, it allows either for λ, γ and η to be constants or functionals of some state variable.[9] In case of only two micro-states, N being constant through time, the attention is focused on only one occupation number (for instance N^1) to characterise the macro state of the entire economy in a given instant, $1 \leq N_k \leq N$: a realisation of the stochastic process $N^1(t)$ on its support is denoted with $N^1(t) = N_k$.

The transition rates determine the probability of observing a certain occupation number at the aggregate level, *i.e.* a certain macrostate of the system. Being $N^1(t) = N_k$, within the length of a vanishing reference unit of time $\Delta \to 0^+$, the expected number of transitions into the macrostate N^1 is $\lambda(N - N_k)$ while the expected number of transitions from macrostate N^1 is γN_k; therefore, the transition rates can be written as follows

$$\begin{aligned} b(N_k) &= P\left(N^1(t+\Delta) = N_{k+1}(t') \,\middle|\, N^1(t) = N_k(t)\right) = \lambda(N - N_k) \\ d(N_k) &= P\left(N^1(t+\Delta) = N_{k-1}(t') \,\middle|\, N^1(t) = N_k(t)\right) = \gamma N_k \end{aligned} \tag{2}$$

9. In Appendix A the stochastic model results are discussed for both cases.

where b and d indicate, respectively, "births" ($N_k \rightarrow N_{k+1}$) and "deaths" ($N_{k-1} \leftarrow N_k$) rate functions of the stochastic process and $t' - t = \Delta$.

2. Firms

This section presents the assumptions for the microeconomic units of the system. The approach is the one pioneered by GS, and implemented in a heterogeneous agents framework by Delli Gatti *et al.* (2005). If not otherwise specified, variables indicated by small letters refer to single firms while symbols in capital letters stand for aggregate quantities, within the state if followed by the superscripts 0 or 1 and economywide otherwise.

2.1. Financial fragility as a clustering device

We assume that financially constrained firms are subject to iid shocks to revenue and, therefore, they run the risk of bankruptcy if revenue fall short of pre-incurred costs. In this setting the optimal scale of activity for the firm is constrained by its net worth due to bankruptcy risk. The firm's probability of bankruptcy depends upon its equity ratio, *i.e.* the ratio of net worth to assets.

In the present paper this approach has been followed in a somewhat stylised way. The economy is populated by a fixed number N of firms which agglomerate into clusters depending on the level of individual equity ratio $\alpha_i = a_i / k_i$, $i = 1, 2, ... N$, where a_i is net worth[10] and k_i total assets (physical capital). The threshold $\overline{\alpha}$ divides the populations of firms in two clusters: firms in state 0 (whose occupation number is N^0), characterised by $\alpha_i \geq \overline{\alpha}$, are financially robust while firms in state (whose occupation number is N^1). characterised by $\alpha_i < \overline{\alpha}$, are financially fragile and exposed to the risk of bankruptcy. Within each cluster, individual levels of the equity ratio are approximated by their mean-field values α^0 and α^1 respectively.

In order to keep the number of firms N constant, each bankrupted firm is replaced by a new one which, by assumption, enters the system in state 1. The probability of being fragile is η

10. Equity or own capital are assumed synonyms of net worth.

while μ denotes the probability of bankruptcy—*i.e.* of exiting from the economy. Hence the rate of exit from the system is $\mu\eta$. Of course, due to the one-to-one replacement assumption, $\mu\eta$ represents also the rate of entry into the system.

2.2. Technology, costs and prices

Each firm employs physical capital as the only input in production. Therefore, the production function of the i-th firm is:

$$q_i(t) = (2k_i(t))^{1/2} \tag{3}$$

and the capital requirement function is:

$$k_i(t) = 1/2\,(q_i(t)^{/2} \tag{4}$$

Firms can finance capital with previously retained profits (net worth). When internal funds are not sufficient, firms resort to loans: $b_i(t) = k_i(t) - a_i(t)$. Debt commitments in real terms are $rb_i(t)$, where r is the real interest rate.[11] For the sake of simplicity the interest rate is constant and uniform across firms.

The firm has no market power (it is a price taker) but is operating in an uncertain environment. The price $P_i(t')$ of the t'-th firm at time t'—*i.e.* when the output is actually sold—is equal to the average or market price $P(t)$ at time t—i.e. when the output is produced and ready for sale—subject to an idiosyncratic multiplicative shock $\tilde{u}_i(t')$:

$$P_i(t') = \tilde{u}_i(t')P(t) \tag{6}$$

The random variable $\tilde{u}_i(t') \to U(u_0, u_1)$ s.t. $u_1 > u_0 > 0$ and $\mathbb{E}(\tilde{u}_i) = 1$. Its support can be any positive neighbourhood of 1: in this paper it has been chosen to set $\tilde{u}_i(t')$ within $[u_0 = 0.75; u_1 = 1.25]$.[12]

2.3. Profit, net worth and bankruptcy

The law of motion of net worth (in real terms) is:

$$a_i(t') = a_i(t) + \pi_i(t') \tag{7}$$

11. By hypothesis, the return on own capital is equal to the interest rate r, so that the firm's financing costs are:

$$r(b_i(t) + a_i(t)) = rk_i(t) \tag{5}$$

12. Due to the normalisation procedure detailed below, the choice of the support for $\tilde{u}_i$ does not affect probabilities.

where $\pi_i\,(t')$ is profit (in real terms):

$$\pi_i(t') = \tilde{u}_i(t')q_i(t) - rk_i(t) \tag{8}$$

A firm goes bankrupt when $a_i\,(t')$ reaches the zero threshold, *i.e.* when

$$\pi_i\,(t') = -a_i\,(t) \tag{9}$$

Substituting (8) into (9), and solving for $\tilde{u}_i(t')$, the bankruptcy threshold level of the shock is

$$\bar{u}_i(t') \equiv \frac{rk_i(t) - a_i(t)}{q_i(t)} \tag{10}$$

Notice that, by construction, the threshold level of the shock occurring at t' is a function of variables defined at time t'. If the shock $\tilde{u}_i(t') \leq \bar{u}_i(t')$, then equity becomes negative (or zero) and the firm goes bankrupt.

Since $\dfrac{a_i(t)}{q_i(t)} = \alpha_i(t)\dfrac{k_i(t)}{q_i(t)}$, and recalling (4), Equation (10) reads as:

$$\bar{u}_i(t') = \frac{q_i(t)}{2}\left[r - \alpha_i(t)\right] \tag{11}$$

The random variable $\tilde{u}_i$ has support [0.75;1.25], therefore, denoting with F the cdf of $\tilde{u}_i(t')$, the probability of bankruptcy μ_i for firm i is

$$\mu_i(t) = F(\bar{u}_i(t)) = \frac{\bar{u}_i(t) - 0.75}{0.5} = 2\bar{u}_i - 1.5 \tag{12}$$

Every firm which goes bankrupt has to bear bankruptcy costs $C_i\,(t)$, non-linearly increasing with firm size,

$$C_i(t) = c(q_i(t))^2 \qquad 0 < c < 1 \tag{13}$$

As discussed by Greenwald and Stiglitz (1990), bankruptcy costs hold to the borrower. They are due to legal and administrative costs incurred during the bankruptcy procedure and to the reputational costs for the managers of a firm which goes bankrupt. These reputational costs are assumed to be increasing with the scale of production.

2.4. Output

Following GS, we assume that at time t the firm (optimally) decides the quantity to produce which will be sold at t' in conditions of uncertainty concerning the sale price. Therefore, the problem of the firm at t consists in maximising an objective function V which is equal to expected profits at t', net of bankruptcy costs, subject to the production function (3):

$$\max V = \mathbb{E}\left\{\tilde{u}_i(t')q_i(t) - rk_i(t) - C_i(t)\mu_i\right\} \tag{14}$$

$$s.t. \quad k_i(t) = \frac{1}{2}(q_i(t))^2 \tag{15}$$

Since $\mathbb{E}(\tilde{u}_i(t')) = 1$, assuming that agents consider the expected probability of bankruptcy at time t' equal to the one at time t, the problem above boils down to the following:

$$\max V = q_i - rk_i - C_i\mathbb{E}(\mu_i) \tag{16}$$

$$s.t. \quad k_i = \frac{1}{2}(q_i)^2 \tag{17}$$

where the time index has been removed to simplify notation.

Firms in state 0 know that their probability of bankruptcy is $\mathbb{E}(\mu_0) = 0$. Hence, for financially robust firms, the problem is:

$$\max V_0 = q_0 - rk_0 \tag{18}$$

$$s.t. \quad k_0 = \frac{1}{2}(q_0)^2 \tag{19}$$

which solves with $q^0 = r^{-1}$ being r given and q^0 constant through time. Financially fragile firms know that they run the risk of bankruptcy. Due to the mean-field approximation, the probability of bankruptcy for firms in state 1 is constant across agents. Hence the optimisation problem becomes:

$$\max V_1 = q_1 - rk_1 - c(q_1)^2\mu \tag{20}$$

$$s.t. \quad k_1 = \frac{1}{2}(q_1)^2 \tag{21}$$

and the solution is

$$q^1 = (r + 2c\mu)^{-1} \tag{22}$$

Note that μ is indeed defined at time t and time dependent so that also q^1 is time dependent. Aggregate production is:

$$Y = Y^1 q^1 + N^0 q^0 \tag{23}$$

where N^1 and N^0 are the previously introduced occupation numbers. Plugging the above obtained results into this definition, aggregate output can be expressed as

$$Y = \frac{N^1}{r + 2c\mu} + \frac{N^0}{r} \tag{24}$$

From Equation (24), it is clear that business fluctuations are driven by (i) the probability of bankruptcy μ and (ii) the dynamics of the occupation numbers. The impact of financial fragility on the aggregate may be better appreciated by reformulating Equation (24) as follows

$$Y = \frac{N}{r} - \xi N^1 \tag{25}$$

where $\xi = \frac{1}{r}\left(1 + \frac{r}{2c\mu}\right)^{-1}$.

While N and r are given, each factor in the product $\xi \mathrm{V}^1$ is defined at time t and time dependent. N^1 can be considered a macroeconomic indicator of the financial fragility of the system; in (25) it is weighted by ξ, which is a function of the probability μ. Therefore, the dynamics of aggregate production appears to be determined by the micro and macro level of financial distress of the economy.

2.5. Transition probabilities

The probability of bankruptcy μ_i can be expressed as

$$\mu_i = F(\bar{u}_i) = \frac{\bar{u}_i - 0.75}{0.5} = 2\bar{u}_i - 1.5 \tag{26}$$

By assumption, only firms in state 1 are exposed to the risk of bankruptcy. It is expected firms lumped in cluster 1, the group of financially fragile firms, have the following bankruptcy threshold:

$$\bar{u}_1 \equiv \frac{q^1}{2}(r - \alpha^1) \tag{27}$$

Hence $\mu = F(\bar{u}_1) = 2\bar{u}_1 - 1.5 = q^1(r - \alpha^1) - 1.5$

According to equation (26), it is possible to quantify the equity ratio threshold $\bar{\alpha}$, which is the minimum level of the equity ratio that ensures the firm's survival (*i.e.* $\mu = 0$),[13] and can be expressed as

$$\overline{\alpha} = r - \frac{1.5}{q^1} \tag{28}$$

Since q^1 is time dependent, the threshold $\overline{\alpha}$ also evolves over time.

The transition probabilities ζ (*i.e.* the probability of moving from 0 to 1) and ι (*i.e.* the probability of moving from 1to 0) can be expressed as variables depending on the price shock $\widetilde{u}_i(t)$, with the appropriate critical values $\overline{u}_\zeta(t)$ and $\overline{u}_\iota(t)$

$$\begin{aligned} \widetilde{u}_i(t) &\leq \frac{q^0(t)}{2}\left[r + \overline{\alpha}(t) - \alpha^0(t)\right] \equiv \overline{u}_\zeta(t) \\ \widetilde{u}_i(t) &> \frac{q^1(t)}{2}\left[r + \overline{\alpha}(t) - \alpha^1(t)\right] \equiv \overline{u}_\iota(t) \end{aligned} \tag{29}$$

The explicit formulation for transition probabilities is therefore

$$\zeta(t) = p(\widetilde{u}(t) \leq \overline{u}_\zeta(t)) = 2\overline{u}_\zeta(t) - 1.5 \tag{30}$$

$$\iota(t) = 1 - p(\widetilde{u}(t) \leq \overline{u}_\iota(t)) = -2\overline{u}_\iota(t) + 2.5 \tag{31}$$

3. Dynamic analysis and solution

This section introduces the master equation, which is the fundamental tool in the analytical solution process, and the main result of its asymptotic solution.

13. It is now straightforward identifying an upper bound for the total credit demand $B = B^0 + B^1$, where B^0 and B^1 are the total demands for each group of firms. Given the optimal levels of capital for each cluster of firms, namely k^1 and k^0, the quantity of credit demanded reaches its maximum when α^1 and α^0 reach their minimum. Note that α^1 cannot go below $r - 2.5/q^1$, at which value μ becomes equal 1. By definition, the minimum level for α^0 is $\overline{\alpha} = r - 1.5/q^1$. For these values it follows that:

$$\lim_{\alpha^0 \to \overline{\alpha}} B^0 = \frac{(q^0)^2}{2}\left(1 - r + 1.5/q^1\right)$$
$$\lim_{\alpha^1 \to \min(\alpha^1)} B^1 = \frac{(q^1)^2}{2}\left(1 - r + 2.5/q^1\right)$$

Consequently, the demand of credit must be smaller than or equal to:

$$\max(B) = N^0\left[\frac{(q^0)^2}{2}\left(1 - r + \frac{1.5}{q^1}\right)\right] + N^1\left[\frac{(q^1)^2}{2}\left(1 - r + \frac{2.5}{q^1}\right)\right]$$

that cannot grow indefinitely since $q^0 = 1/r$ and $q^1 < q^0$ as shown below.

3.1. Aggregate dynamics

The solution of the model requires the specification of aggregate output dynamics. As shown by equations (24) and (25), aggregate output depends on a stochastic process, whose outcome is given by the occupation numbers N^0 and N^1. It is assumed that the stochastic process is a jump Markov process and its macro dynamics is analytically explored by means of the master equation, *i.e.* a differential equation that describes the dynamics of the probability distribution of a system of agents over its state space through time. The master equation can be primarily specified as a balance flow equation between probability inflows and outflows in and from a generic macro-state.

The state variable $N^1(t) = N_k$ is the number of fragile firms, those in state 1. The variation of probability in a vanishing reference unit of time is

$$\begin{aligned}\frac{dP(N_k,t)}{dt} &= b(N_{k-1},t)P(N_{k-1},t)+d(N_{k+1},t)P(N_{k+1},t)+ \\ &-\{[(b(N_k,t)+d(N_k,t))P(N_k,t)]\}\end{aligned} \tag{32}$$

with boundary conditions:

$$\begin{cases} P(N,t) = b(N^1)P(N^1-1,t)+d(N)P(N,t) \\ P(0,t) = b(1)P(1,t)+d(0)P(0,t) \end{cases} \tag{33}$$

The variation of probability defined in the equation above is defined as the sum of inflow-births from N_{k-1} and inflow-deaths from N_{k+1} less outflows from N_k due to births and deaths. Finally, the boundary conditions ensure a consistent value for the probability $P(N_k)$. Therefore, in order to identify the dynamics of firms and production, Equation (32) must be solved.

3.2. Master equation's solution: stochastic dynamics of trend and fluctuations

As shown by (Gardiner, 1985; Risken, 1989), a direct solution of the master equation is possible only under restrictive assumptions. Inspired by van Kampen (2007), Aoki (1996, 2002) and Aoki and Yoshikawa (2006) suggest a method to overcome this problem which consists in splitting the control variable into the drift and diffusion components of the underlying process.[14] The application of this method appears of particular interest in this context as it allows to analytically identify both the trend and the fluctuations distribution of aggregate production. More precisely, the fraction of firms in state 1 in a given instant is assumed to be determined by its expected value (m), the drift, and an additive fluctuations component of order $N^{1/2}$ around this value, that is the spread:

$$N_k = Nm + \sqrt{N}s \tag{35}$$

Once the master equation has been modified accordingly, that is in terms of s rather than of N_k, it can be solved using approximation methods. As shown in Appendix A, the asymptotically approximated solution of the master equation is given by the following system of coupled equations:

$$\frac{dm}{dt} = \zeta m - (\zeta + \iota)m^2 \tag{36}$$

$$\frac{\partial Q(s)}{\partial t} = \left[2(\zeta + \iota)m - \zeta\right]\frac{\partial}{\partial s}(sQ(s)) + \frac{\left[\zeta m(1-m) + \iota m^2\right]}{2}\left(\frac{\partial}{\partial s}\right)^2 Q(s) \tag{37}$$

where $Q(s(t),t) = P(N^1(t),t)$ is substituted into (32) to reformulate the master equation as a function of the spread s. Equation (35) is an ordinary differential equation which displays a logistic dyna-

14. The authors are aware that this method presents some drawbacks. First of all, van Kampen's method develops a local approximation suitable to be applied only when the underlying observable has a unimodal distribution, as the case under study in this paper. Secondly, by allowing for a second order approximation it ends up with a Fokker-Planck equation which solves into a Gaussian distribution for fluctuations: in the present paper it is shown that this is not the distribution for the state variable but only for its spreading fluctuations about the drift. Thirdly, when fluctuations are not of the order of the square root of N, higher order moments might not vanish asymptotically and thus leading to non-Gaussian distributions. This last aspect can be found when dealing with microscopic models grounded on global interactions, which is not the case under study, or when mean-field approached are not very suitable, see Castello *et al.* (2006) and Stauffer *et al.* (2006) on this issue. Despite these limits, this method is adopted for different reasons. First of all it is relatively easy to handle, as shown in the present paper. In particular, if one is interested in macroscopic dynamics of a given quantity or an aggregation procedure, it does not require to provide a solution to the master equation in terms of the probability distribution of the state variable. The aim is to find the equations for the drift and spread only. Secondly, it allows for a complete description of the stochastic aggregate dynamics in terms of transition rates and related parameters, such as transition probabilities at micro-level, which can be analytically obtained from the underlying agent based model. Third: if one is allowed to assume the van Kampen's ansatz (34) and mean-field interaction, rather than global, is considered, then the needed condition for a suitable second order approximation are met. The method is here developed following Landini and Uberti (2008) and Di Guilmi *et al.* (2011).

mics for the drifting component. Equation (36) is a second order partial differential equation, known as the Fokker-Planck equation, that drives the density of the spreading component s. The dynamics converge to the steady state values: setting the l.h.s. of (35) to 0, the stable steady-state value for m is

$$m^* = \frac{\zeta}{\zeta + \iota} \tag{37}$$

Then, by integration of (35) with an initial condition $m(0) = m_0$ we get:

$$m(t) = \frac{m_0 \zeta e^{\zeta t}}{\zeta + m_0 (\zeta + \iota)\left(e^{\zeta t} - 1\right)} \tag{38}$$

This equation describes the dynamics of the fraction m of firms occupying state 1 at each point in time. It is fully dependent on transition rates. The stationary solution of the equation for the spread component (see Appendix B) yields the distribution function $\overline{Q}$ for the spread s, thus determining the probability distribution of fluctuations:

$$\overline{Q}(s) = C \exp\left(-\frac{s^2}{2\sigma^2}\right) \; : \; \sigma^2 = \frac{\zeta\iota}{(\zeta + \iota)^2} \tag{39}$$

which looks like a Gaussian density, dependent only on transition probabilities. Given the relationship among m and total production, the dynamics of our economy is now fully described by having at hands a differential equation for output dynamics, its equilibrium value, and a probability function for business fluctuations around the trend.

4. Interaction and output dynamics: the stochastic financial contagion

This section shows how the transition rates provide a functional representation of the interaction of firms within each cluster and of the feedback effects between the macro and the micro-level of our stylised economy. The first subsection proposes and endogenous formulation for the probability η, which makes possible a reinterpretation of the formula for aggregate output and the transition rates, as illustrated in the second subsection.

4.1. Stationary points and equilibrium probability

An important result derived by the asymptotic solution of the master equation concerns the theoretical probability η of being in state 1. In particular it is possible to identify a functional form that quantifies the impact of indirect interaction among firms. By definition, the steady-state condition implies that the probability of in-flows is equal to the probability of out-flows for all possible states. Analytically it means a null value for the r.h.s. of the master equation. This condition is defined as detailed balance.[15] Provided that detailed balance holds for each pair of macro-states, Appendix C shows that the stationary probability for a given macro state $P^e(N_k)$ is

$$P^e(N_k) = P^e(N(0)) \left(\frac{\iota}{\zeta}\right)^{N_k} \genfrac{}{}{0pt}{}{N}{N_k} \prod_{h=1}^{H} \frac{\eta(N-N_h)}{(1-\eta(N_h))} \tag{40}$$

The probability $P^e(N_k)$ can be also expressed in Gibbs form,[16] and a Gibbs functional form for the probability η is

$$\eta(N^1) = N^{-1} e^{\beta g(N^1)} \tag{41}$$

where:

$$\begin{aligned} \beta(t) &= \ln\left(-\frac{y^1(t)-\bar{y}(t)}{y^0(t)-\bar{y}(t)}\right)\left(y^1(t)-y^0(t)\right)^{-1} \\ g(N^1) &= -\frac{1}{2\beta}\ln\left(\frac{N^1}{N-N^1}\right) = \frac{y^0-y^1}{2} \end{aligned} \tag{41'}$$

The symbol $\bar{y}(t)$ stands for the average production $Y(t)/N$. The probability of being in state 1 in a given instant depends on three factors: the number of firms already occupying the state, N^1; the parameter β, which measures the impact on total output of the relative financial distress of firms; the function $g(N^1)$, which quantifies the average difference in the optimal levels of production. The circular feedback effects are displayed by Equation (41): the macro-to-micro effect captures the link of the behaviour of a firm

15. It is worth stressing that the detailed balance does not imply that agents do not switch between the micro-states, but that inflows and outflows for each micro-state balance out.
16. The equivalence is demonstrated by the Hammersley and Clifford theorem which states that for each Markov random field there exists one and one only Gibbs random field (Clifford, 1990).

to the state of the economy; the bottom-up or micro-to-macro effect, on the other hand, determines the aggregate performance by the number of firms with lower output and by the relative difference in optimal outputs, captured by $g(N^1)$ and β. Therefore, by means of equation (41), the whole dynamics of the system can be interpreted as the result of indirect interaction among firms and of the feedback effects between macro, meso and micro level.

4.2. Output dynamics

Making use of equations (25) and (37), the steady-state value of aggregate production, Y^e, can be expressed as:

$$Y^e = N\left[\frac{1}{r} - \frac{\zeta}{\zeta+\iota}\frac{2c\mu}{r(r+2c\mu)}\right] = N\left[\frac{1}{r} - \frac{\zeta}{\zeta+\iota}\left(y^0 - y^1\right)\right] \tag{42}$$

Equation (42) highlights two factors that influence production dynamics: the difference between firms' optimal production levels and the transition rates. The first component is determined by the exogenous parameter c, that reflects institutional conditions, and by the probability of bankruptcy μ, that is the result of the relative financial condition of financially distressed firms, being a function of the difference between their equity ratio α^1 and the "safety" level $\overline{\alpha}$.

The transition rates component is the result of a micro factor (the relative financial conditions of the two types of firms) and of a macro factor (the general financial situation of the system, revealed by the number of firms in each state), as shown by equations (1) and (30). The formulation of λ and γ under detailed balance condition helps in clarifying further the point. Substituting equations (30) and (41) in equations (1), we obtain:

$$\lambda = \left(N^{-1}e^{\beta g(N^1)}\right)\left(2\overline{u}_\zeta - 1.5\right) \tag{43}$$

$$\gamma = \left[1 - \left(N^{-1}e^{\beta g(N^1)}\right)\right]\left(-2.5\overline{u}_\iota + 1.5\right) \tag{44}$$

The micro factor is quantified by $\overline{u}_\zeta$ and $\overline{u}_\iota$, that, as shown in equations (29), reflect the difference between $\overline{\alpha}$ and the mean-field variables α^1 and α^0. This effect is amplified by the macro factor that, in turn, is dependent on the occupation numbers and on the relative difference in optimal levels of production. The joint effect of these variables gives rise to a mechanism that can be defined as stochastic financial contagion, given that a worsening in micro financial conditions raises both the probability of bankruptcy and the probability of entering state 1.

The solution reveals that the dynamics of the economy is dependent on the distribution of agents and on its evolution. Given the inherent uncertainty of these dynamics, all the functional relationship are expressed as probability functions. Therefore, the dynamics of the system appears to be fully stochastic, and the steady-state level of production cannot be considered as a natural equilibrium.

5. Simulations

In order to visualize the actual dynamics of the system and check the reliability of the stochastic approximation, Monte Carlo simulations have been performed. The agent based model have been simulated with fully heterogeneous firms according to the hypothesis detailed in section 2. Then mean-field variables α^1 and α^0 as the medians of the equity ratios within each cluster have been calculated. These values are the input of the stochastic dynamics procedure, performed according to the structure of section 1. The simulation has been repeated 1000 times, drawing a new set of random numbers for each replication. The number of firms is $N = 300$ and the parameter c is set equal to 1.

To appreciate the volatility endogenously generated by the system, figure 1 displays the symmetric dynamics of low-equity firms and aggregate production for a single replication. The convergent evolution of n^1 is driven by equation (38) with fluctuations around the trend distributed according to (39). Its dynamics fully explains the growth of aggregate production and the business fluctuations. The higher volatility in the series of output is due to the shocks in price.

Figure 2 compares the agent based results with its stochastic approximation in the initial stages of the adjustment process. Agent based trend dynamics are well mimicked by the stochastic approximation. The fluctuations generated by the two procedures cannot match, as in the latter they are the outcome of a random variable. Nevertheless, the amplitudes of volatility are comparable.

The result is satisfactory as the average variance of the time series is .0068 for agent based and .0063 for the stochastic approximation.

Figure 1. Trends and fluctuations for value of aggregate production (left scale) and n^1 (right scale). Single replication

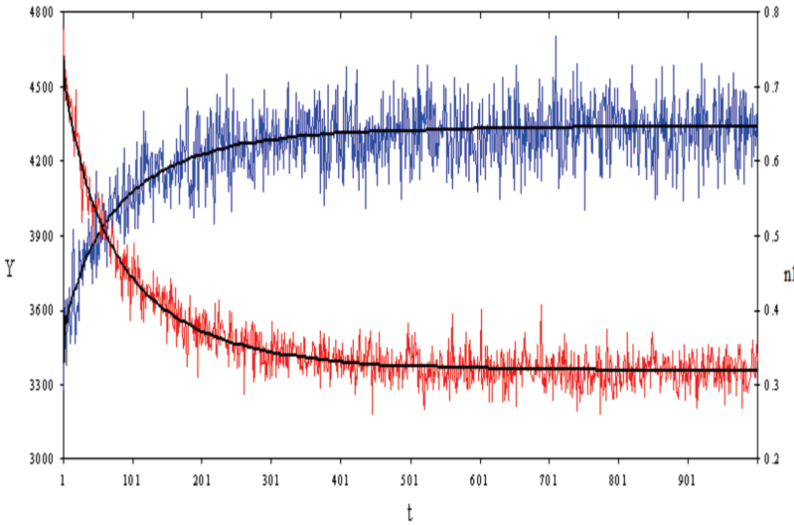
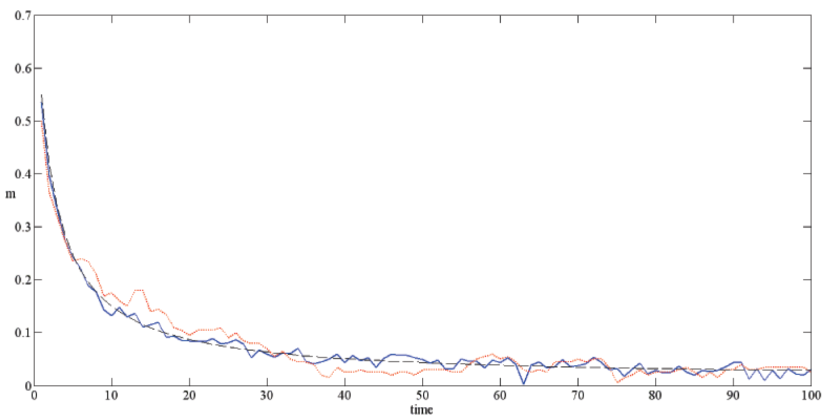


Figure 2. Dynamics for n^1 for agent based simulation (red dotted line) and stochastic approximation (black dashed line for trend and continuous blue line for fluctuations). Single replication



The dynamics of the two series obtained by the Monte Carlo simulation over 1000 replications are displayed in figure 3. They overlap for almost all the periods, although the adjustment process to the steady state is shorter for the stochastic approximation. Their significant correlation is .96. Thus, the stochastic dynamics proves reliable for an analytical representation of more complex and diversified structure. The simplification to the two states approximation does not seem to reduce the accuracy of the solution.

Figure 3. Dynamics for n^1 for agent based simulation (red dotted line) and stochastic approximation (black continuous line). Monte Carlo simulation with 1000 replications

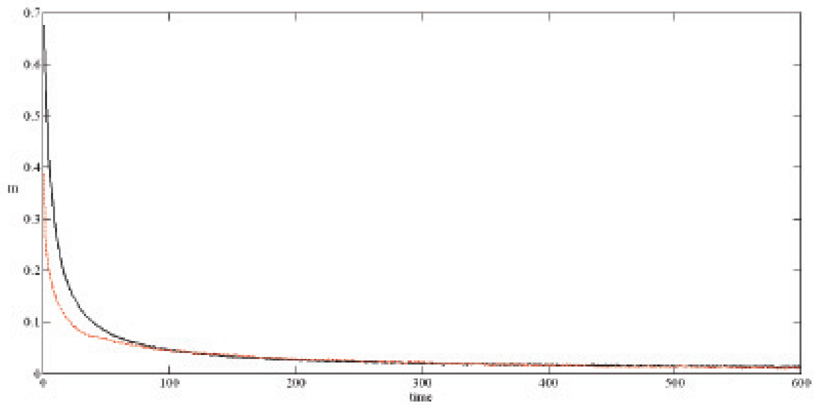
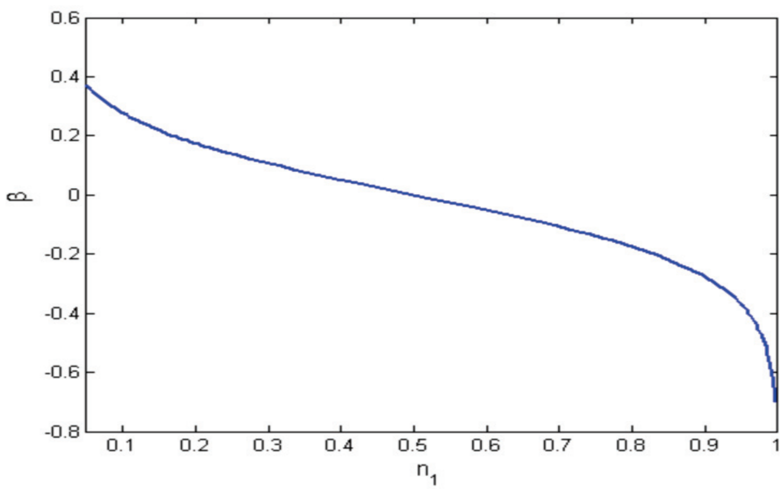
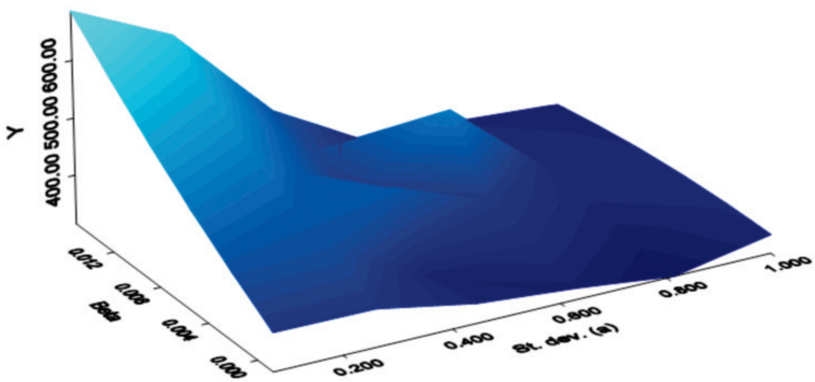


Figure 4. Relationship between β and n^1



The variable β, which enters the definition of transition rates, is inversely related to the number of financially fragile firms, as shown by figure 4. Hence, it represents an inverse index of the systemic financial fragility. According to Equation (42), the aggregate output is expected to be higher for lower level of N^1 and, thus, higher β. This result is confirmed by the simulation and illustrated in figure 5. The same graph reveals that the performance of the economy is also dependent on the shape of the distribution of the net worth, as lower levels of standard deviation for a appear to be associated to a larger aggregate production.

Figure 5. Countour plot of the aggregate output as a function of the standard deviation of net worth distribution and β



6. Concluding remarks

This work proposes a solution to the problem of the aggregation of heterogeneous agents in a dynamical context by applying a method which analytically identifies the components of macroeconomic dynamics, namely, trend and fluctuations. It is worth stressing that the long run steady-state of production cannot properly be defined as natural equilibrium. From the methodological point of view, the main contribution of the present work is the identification of a differential equation for trend and a probability distribution function for the fluctuations of the aggregate production by means of the asymptotic solution of the master equation. All the variables that appear in these two formulations are endogenous and provide an analytical representation of the interaction among agents and the feedback effects that arise among the different levels of aggregation within the system. In particular, both the probability for a firm to reduce its production as a consequence of the risk of failure and the actual probability of bankruptcy are dependent on the financial distress of the other firms in the system, measured by the number of firms with low equity ratio and by the mean-field approximations of the equity ratios. Aggregate production is itself dependent on the ratio among debt and equities of each firm, and this gives rise to feedback effects between micro and macro levels of the system. The overall effect can be defined as stochastic financial contagion.

This methodology appears as particularly suitable for models where the micro financial variables have a relevant impact on the macroeconomy. In such a way, the modelling of the links among financial fragility, business cycles and growth dynamics can be consistently microfounded, taking into account the heterogeneity of firms' financial variables and the interaction among agents and between agents and the macro level of the system. However, the actual range of application of this body of tools extends to all the contexts in which the heterogeneity of agents and their interaction cannot be neglected or reduced in order to represent, *e. g.*, the efficacy of an economic policy measure or the transmission mechanism of a shock. All in all, the whole of macroeconomics.

The limitation to the heterogeneity does not seem to impact on the performance of the model that proves capable to replicate the behaviour of an analogous agent based model, with no restrictions on the heterogeneity of firms.

References

Alfarano S., T. Lux T., and Wagner F., 2008. Time variation of higher moments in a financial market with heterogeneous agents: An analytical approach, *Journal of Economic Dynamics and Control*, 32(1): pp. 101–136.

Alfarano S., and M. Milakovic, 2009. Network structure and N-dependence in agent-based herding models, *Journal of Economic Dynamics and Control*, 33(1): pp. 78–92, URL http://ideas.repec.org/a/eee/dyncon/v33y2009i1p78-92.html

Aoki M., 1996. *New approaches to macroeconomic modeling*. Cambridge University Press.

Aoki M., 2002. *Modeling aggregate behaviour and Fuctuations in economics*. Cambridge University Press.

Aoki M., and H. Yoshikawa, 2006. *Reconstructing Macroeconomics*. Cambridge University Press.

Balian R., 1991. *From Microphysics to Macrophysics. Volume I*, Berlin/Heidelberg/New York: Springer-Verlag.

Brock W. A., and S. N. Durlauf, 2001. Discrete Choice with Social Interactions, *The Review of Economic Studies*, 68(2): pp. 235–260.

Brook D., 1964. On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems. *Biometrica*, 51: pp. 481–483.

Castello X., V. M. Eguiluz, and M. S. Miguel, 2006. Ordering dynamics with two non-excluding options: bilingualism in language competition. *New Journal of Physics*, 8(12): p. 308, doi:10.1088/1367-2630/8/12/308, URL http://stacks.iop.org 1367-2630/8/308.

Chiarella C., and C. Di Guilmi, 2011. The Financial Instability Hypothesis:

A Stochastic Microfoundation Framework. *Journal of Economic Dynamics and Control*, 35(8): pp. 1151–1171.

Clifford P., 1990. Markov random fields in statistics, in: Grimmett, G. R. and Welsh, D. J. A. (eds.), *Disorder in Physical Systems. A Volume in Honour of John M. Hammersley*, pp. 19–32, Oxford: Clarendon Press, URL citeseer.ist.psu.edu/clifford90markov.html

Delli Gatti D., C. Di Guilmi, E. Gaffeo, G. Giulioni, M. Gallegati, and A. Palestrini, 2005. A new approach to business fluctuations: heterogeneous interacting agents, scaling laws and financial fragility. *Journal of Economic Behavior and Organization*, 56(4).

Di Guilmi C., 2008. *The generation of business fluctuations: financial fragility and mean-field interaction*. Peter Lang Publishing Group: Frankfurt/M.

Di Guilmi C., M. Gallegati, and S. Landini, 2010. Financial fragility, mean-field interaction and macroeconomic dynamics: A stochastic model, in: N. Salvadori, (ed.). *Institutional and social dynamics of growth and distribution*, pp. 323–351, Edward Elgar, UK.

Di Guilmi C., M. Gallegati, and S. Landini, 2011. *Interactive Macroeconomics*, URL http://www.bus.uts.edu.au/fin&econ/staff/Corrado/Interactive_Macroeconomics.pdf, forthcoming.

Foley D., 1994. A Statistical Equilibrium Theory of Markets. *Journal of Economic Theory*, 62: pp. 321–345.

Gallegati M., A. Palestrini, D. Delli Gatti, and E. Scalas, 2006. Aggregation of heterogeneous interacting agent: the variant representative agent framework. *Journal of Economic Interaction and Coordination*, 1(1): pp. 5–19.

Gardiner C. W., 1985. *Handbook of stochastic methods*. Springer-Verlag: New York.

Garibaldi U., and E. Scalas, 2010. *Finitary probabilistic methods in Econophysics*. Cambridge University Press: Cambridge.

Gorman W. M., 1953. Community Preference Fields. *Econometrica*, (21): pp. 63–80.

Greenwald B., and J. E. Stiglitz, 1990. Macroeconomic models with equity and credit rationing. in: Hubbard, R. (ed.), *Information, Capital Markets and Investment,* Chicago University Press, Chicago.

Greenwald B., and J. E. Stiglitz, 1993. Financial markets imperfections and business cycles. *Quarterly journal of Economics*, (108).

Hinich M. J., J. Foster, and P. Wild, 2006. Structural change in macroeconomic time series: A complex systems perspective. *Journal of Macroeconomics*, 28(1): pp. 136–150, available at http://ideas.repec.org/a/eee/jmacro/v28y2006i1p136-150.html

Jaynes E. T., 1957. Information Theory and Statistical Mechanics. *Phys Rev*, 106: pp. 620–630, doi:10.1103/PhysRev.106.620, URL http://link.aps.org/doi/10.1103/PhysRev.106.620.

Keen S., 2011. *Debunking Economics: The Naked Emperor Dethroned?*. London: Zed Books, first edition: 2001.

Kirman A. P., 1992. Whom or What Does the Representative Individual Represent?. *Journal of Economic Perspectives*, 6(2): pp. 117–36, available at http://ideas.repec.org/a/aea/jecper/v6y1992i2p117–36.html

Landini S., and Uberti M., 2008. A Statistical Mechanic View of Macrodynamics in Economics. *Computational Economics*, 32(1): pp. 121–146.

Lewbel A., 1992. Aggregation with Log-Linear Models. *Review of Economic Studies*, 59: pp. 635–42.

Milakovic M., 2001. A Statistical Equilibrium Model of Wealth Distribution. Computing in Economics and Finance 2001 214, Society for Computational Economics, URL http://ideas.repec.org/p/sce/scecf1/214.html

Risken H., 1989. *Fokker-Planck equation. Method of solutions and applications*. Berlin: Springer Verlag.

Smith J., 1993. Moment Methods for Decision Analysis. *Managment Science*, 39(3): pp. 340–358.

Stauffer D., X. Castello, V. M. Eguiluz, and M. S. Miguel, 2006. Microscopic Abrams-Strogatz model of language competition. *Physica A*, 374(2): pp. 835–842.

van Kampen N., 2007. *Stochastic processes in physics and chemistry*. Elsevier: Amsterdam.

Weidlich W., and M. Braun, 1992. The Master Equation Approach to Nonlinear Economics. *Journal of Evolutionary Economics*, 2(3): pp. 233–65.

Appendix A

This appendix develops the method involved to reach a mean-filed system of coupled equations for the drift and the spread of the state variable in the master Equation (32). According to (34) for a fixed it follows that $N^1(t) = N_k \in [1, N]$

$$s(t) = (N_k - Nm(t))N^{-1/2} \tag{A.1}$$

Accordingly, the master equation (ME) (32) can be rewritten a function of the state variable s. The fact that N_k is fixed does not mean that it is constant but just that we focus out attention on it as a specific realization of $N^1(t)$; Accordingly, from (A.1) it follows that

$$\frac{ds}{dt} = -\sqrt{N}\frac{dm}{dt} \tag{A.2}$$

hence the l.h.s. of (32) reads as

$$\frac{dP(N_k)}{dt} = \frac{\partial Q(s)}{\partial t} - \sqrt{N}\frac{dm}{dt}\frac{\partial Q(s)}{\partial s} = \frac{dQ(s)}{dt} \tag{A.3}$$

being $P(N^1(t) = N_k) = Q(s(t))$. In order to find a suitable expression for the r.h.s. of (32), transition rates are written as follows

$$\begin{aligned} b(N_{k-\theta}) &= \lambda[N - (N_k - \theta)] = \lambda N\left\{1 - \left[m + \frac{1}{\sqrt{N}}\left(s - \frac{\theta}{N}\right)\right]\right\} \\ &= \lambda N\left[1 - \left(m + \frac{1}{\sqrt{N}}s_{(-\theta)}\right)\right] = b(s_{(-\theta)}) \end{aligned} \tag{A.4}$$

$$\begin{aligned} d(N_{k+\theta}) &= \gamma[N_k + \theta] = \gamma N\left[m + \frac{1}{\sqrt{N}}\left(s + \frac{\theta}{\sqrt{n}}\right)\right] \\ &= \gamma N\left(m + \frac{1}{\sqrt{N}}s_{(+\theta)}\right) = d(s_{(+\theta)}) \end{aligned} \tag{A.5}$$

where $\theta = 0$ means outflow and $\theta = 1$ inflow, consistently with the phenomenological ME (32).

The lead (+) and lag (–) operators are defined as

$$f_{(\pm)}(N_{k\pm\theta})) = L^{\pm}[f_{(\pm)}(N_{k\pm\theta})]_{\theta=0} = \sum_{z=1}^{\infty}\frac{(\pm N^{-1/2}\partial_s)^z}{z!}[g_{(\pm)}(s_{(\pm\theta)})]_{\theta=0} \tag{A.6}$$

where

$$f_+(N_{k+\theta}) = d(N_{k+\theta})P(N_{k+\theta}) = d(s_{(+\theta)})Q(s_{(+\theta)}) = g_+(s_{(+\theta)}) \tag{A.7}$$

$$f_-(N_{k-\theta}) = b(N_{k-\theta})P(N_{k-\theta}) = b(s_{(-\theta)})Q(s_{(-\theta)}) = g_-(s_{(+\theta)}) \quad \text{(A.8)}$$

Hence it follows that

$$d(N_{k+1})P(N_{k+1}) = L^+[d(N_k)P(N_k)] = d(s_{(+1)})Q(s_{(+1)}) = L^+[d(s)Q(s)] \quad \text{(A.9)}$$

$$b(N_{k-1})P(N_{k-1}) = L^-[b(N_k)P(N_k)] = b(s_{(-1)})Q(s_{(-1)}) = L^-[b(s)Q(s)] \quad \text{(A.10)}$$

Therefore, by using (A.9) and (A.10) into the r.h.s. of (32) it follows that

$$\frac{dQ}{dt} = [L^+ - 1](d(s)Q(s)) + [L^- - 1](b(s)Q(s)) \quad \text{(A.11)}$$

which is the ME to be solved. The solution is approximated as (A.6) involves Taylor's polynomials to approximate probability flows about N_k.

Rescaling time as $t = N\tau : [t] \neq [\tau]$, with a second order approximation it follows

$$\partial_t Q - \sqrt{N}\dot{m}\partial_s Q = N^{-1/2}\partial_s[\rho_-(s)Q(s)] + N^{-1}\frac{1}{2}\partial_s^2[\rho_+(s)Q(s)] \quad \text{(A.12)}$$

where, for notation convenience, t stands for r. Expression (A.12) is a Fokker-Planck equation, equivalent to the approximation one gets with the Kramers-Moyal expansion if Pawulas' theorem does not allow for a closed form solution (see Risken, 1989; Gardiner, 1985; Di Guilmi *et al.*, 2011), and coefficients are given by

$$\rho_\pm(s) = d(s) \pm b(s) \quad \text{(A.13)}$$

Case 1. If transition rates in (2) have birth (λ) and death (γ) rates constant through time, by substituting (A.13) into (A.12) according to (A.4) and (A.5) with $\theta = 0$, after having computed derivatives and collected terms with powers of N, it happens that as $N^{-p/2} \to 0$ as $N \to \infty$ $\forall p \geq 2$ hence, by using the polynomial identity principle, it can be found that

$$\begin{cases} (i) \quad \dot{m} = \lambda - (\lambda + \gamma)m \\ (ii) \quad \partial_t Q(s) = [\lambda + \gamma]\partial_s(sQ(s)) + [\lambda - (\lambda + \gamma)m]\frac{1}{2}\partial_s^2 Q(s) \end{cases} \quad \text{(A.14)}$$

The (A.14–*i*) in the mean-field system gives the so called macroscopic equation: its solution provides the most probable drifting path trajectory for $\langle N^1(t)/N \rangle$. The (A.14–*ii*) is the Fokker-Planck equation for the probability distribution of spreading fluctuations about the drift. Both admit a closed form solution allowing for a solution of the ME (A.11) equivalent to (32).

Case 2. In the case of this model, the transition rates have birth and death rates which change over time. Therefore, the following externality functions are introduced in order to model their evolution

$$\lambda(t) = \Lambda(N^1(t)) = \zeta(t)\left[\frac{N^1(t)}{N}\right] \tag{A.15}$$

$$\gamma(t) = \Gamma(N^1(t)) = \iota(t)\left[\frac{N^1(t)-1}{N}\right] \tag{A.16}$$

Therefore, since $N^1(t) = N_k$, (A.15) and (A.16) can be substituted into (A.4) and (A.5) with $\theta = 0$ to get an expression for (A.13) with the modification of transition rates just highlighted. Subsequently, after the derivatives have been computed, the terms with the same order of powers for N are collected such that $N^{-p/2} \to 0$ as $N \to \infty$ $\forall p \geq 2$. By applying the polynomial identity principle it then follows that

$$\begin{cases} (i) & \dot{m} = \zeta m - (\zeta + \iota) m^2 \\ (ii) & \partial_t Q(s) = [2(\zeta + \iota)m - \zeta]\partial_s(sQ(s)) + [\zeta m(1-m) + \iota m^2]\frac{1}{2}\partial_s^2 Q(s) \end{cases} \tag{A.17}$$

where the macroscopic equation (A.17–*i*) gives a logistic dynamics. The non linearity is due to the rate functions (A.15) and (A.16) which account for external field effects on transition rates (A.4) and (A.5).

The macroscopic equation (A.17–*i*) is an ODE, hence with an initial condition $m(0) = m_0 = N_1(0) = N$ it allows for a logistic dynamics with multiple equilibria: $\in \{0, \zeta/(\zeta + \iota)\}$. The stable equilibrium is $m^* = \zeta/(\zeta + \iota)$ and the general solution is

$$m(t) = \frac{m_0 \zeta e^{\zeta t}}{\zeta + m_0(\zeta + \iota)(e^{\zeta t} - 1)} \tag{A.18}$$

which describes the evolution of $\langle N^1(t)\rangle = Nm(t)$ as the expected, *i.e.* most probable, drifting path.

Appendix B

Herein a solution for the Fokker-Planck equation (A.17–*ii*) is found in terms of $Q(s)$. Using $\bar{Q}(s)$ to indicate the stationary probability for , by setting $\dot{Q} = 0$ it follows that

$$[2(\zeta + \iota)m^* - \zeta](s\bar{Q}(s)) = \frac{\zeta m^*(1-m^*) + \iota m^{*2}}{2}\partial_s^2 \bar{Q}(s)$$

By direct integration it gives

$$\bar{Q}(s) = C\exp\left[\frac{\zeta - 2(\zeta+\iota)m^*}{\zeta m^* + (\iota - \zeta)m^{*2}} s^2\right] \tag{B.1}$$

By substituting for $m^* = \zeta / \zeta + \iota)$ it can be found that

$$\bar{Q}(s) = C\exp\left(-\frac{s^2}{2\sigma^2}\right) : \sigma^2 = \frac{\zeta\iota}{(\zeta+\iota)^2} \tag{B.2}$$

being $C = 1/\int \exp\left(-\frac{s^2}{2\sigma^2}\right)ds$ a normalization constant.

Since the variance is a function of ζ and ι, which are time dependent, this representation allows for *stochastic determinism*. That is, the stationary solution for the distribution of spreading fluctuations still performs some *vibrating* volatility due to the exchange of agents between the two states. These exchanges let the volumes almost constant on expectation through time when approaching the stable equilibrium, but fluctuations depend on *who* is jumping because agents jump from one state to another carrying their own characteristics and endowments. Unfortunately these individual jumps are unobservable, and agents are indistinguishable, from a macroscopic point of view. Nevertheless, it known it happens and this let macroscopic observables to *vibrate* about some equilibrium path. On the other hand, equilibrium itself is a state of nature for the system as a whole, it is not a property of its elementary constituents; equilibrium is a probability distribution for agents over a space of states and not a point of balance of two forces.

Appendix C

The basic steps for deriving of the steady state probability are here sketched, referring the interested reader to the cited references. Stationary probability can be obtained by applying Brook's lemma (Brook, 1964) which defines local characteristic of continuous Markov chains. Hammersley and Clifford demonstrate that, under opportune conditions, for each Markov random field there is one and only one Gibbs random field, and define the functional form for the conjunct probability structure once the neighbourhood relations have been identified (Clifford, 1990). The expected stationary probability (40) of the Markovian process for N^1, when detailed balance holds, can be expressed by:

$$P^e(N_k) \propto Z^{-1} e^{-\beta N U(N_k)} \tag{C.1}$$

where $U(x)$ is the Gibbs potential and can be defined as a functional of the local dynamic characteristics of the state variable N_k. In particular:

$$e^{\beta g(N^1)} + e^{-\beta g(N^1)} = N \tag{C.2}$$

The above formulation leads (Aoki, 2002) to an explicit formulation for the probability η as a function of the state variable N^1:

$$\eta(N^1) = N^{-1} e^{\beta g(N^1)} \tag{C.3}$$

where $g(N^1)$ is a function that evaluates the relative difference in the outcome as a function of N^1. β may be interpreted as an inverse measure of the system uncertainty. The uncertainty among the different possible configurations in a stochastic system can be evaluated through a statistical entropy measure (Balian, 1991). The quantification of the parameter β can be obtained by maximising the statistical entropy of the system (Jaynes, 1957). In the present case the problem is configured as follows:

$$\begin{cases} \max H(N^1, N^0) = -N^1(t)\ln(N^1(t)) - N^0(t)\ln(N^0(t)) \quad s.t. \\ N^1(t) + N^0(t) = N \\ N^1(t)y^1(t) + N^0(t)y^0(t) = Y(t) \\ \end{cases} \tag{C.4}$$

The first of the two constraints ensures the normalization of the probability function. The second ensures that all the wealth in the system is generated by firms in the two kind of states. The solution of the maximisation problem (for details see Di Guilmi, 2008) yields

$$\beta(t) = \ln\left(-\frac{y^1(t) - \bar{y}(t)}{y^0(t) - \bar{y}(t)} \right) \left(y^1(t) - y^0(t) \right)^{-1} \tag{C.5}$$

Large values of β associated with positive values of $g(N^1)$ cause $\eta(N^1)$ to be larger than $1 - \eta(N^1)$, making the transition from state 0 to state 1 more likely to occur than the opposite one. In binary models and for great N, the equation of the potential is:

$$U(N^j) = -2\int_0^{N^j} g(z)dz - \frac{1}{\beta} H(\underline{N})$$

where $H(\underline{N})$ is the Shannon entropy with $\underline{N} = (N^1, N^0)$. In order to find the stationary points of probability dynamics we need to individuate its peak (if it exists). β is an inverse multiplicative factor for entropy: a relative high value of β means that the uncertainty in the system is low, with few firms exposed at bankruptcy risk. For values of β around 0, and a more relevant volatility in the system, in order to find the peak of probability dynamics we need to find the local minimum of the potential. Aoki (2002) shows that the points in which the potential is minimized are also the critical point of the aggregate dynamics of $P^e(N_k)$. Deriving the potential with respect to N^1 and then setting $U' = 0$:

$$g(N^1) = -\frac{1}{2\beta}\frac{dH}{dN^1} = -\frac{1}{2\beta}\ln\left(\frac{N^1}{N - N^1} \right) \tag{C.6}$$

and using equation (C.5),an explicit formulation for $g(N^1)$ is found in stationary conditions:

$$g(N^1) = \frac{y^0 - y^1}{2}$$

that quantifies the mean difference (for states) of the output.

OF ANTS AND VOTERS
MAXIMUM ENTROPY PREDICTION OF AGENT-BASED MODELS WITH RECRUITMENT[1]

Sylvain Barde
School of Economics, Keynes College, University of Kent, Canterbury
OFCE

Maximum entropy predictions are made for the Kirman ant model as well as the Abrams-Strogatz model of language competition, also known as the voter model. In both cases the maximum entropy methodology provides good predictions of the limiting distribution of states, as was already the case for the Schelling model of segregation. As an additional contribution, the analysis of the models reveals the key role played by relative entropy and the model in controlling the time horizon of the prediction.

Keywords: Information entropy, Agent-based models, Voter models.

The maximum entropy (MaxEnt) methodology was first introduced as a general method of statistical prediction by (Jaynes 1957a,b), who showed that its use in predicting the dynamic evolution of an unobserved system could be extended beyond its initial use in physics. This insight was incorporated into the bayesian image reconstruction framework of Cornwell & Evans (1985), Narayan & Nityananda (1986) and Skilling & Gull (1991). Assuming that the received data d about a signal is noisy or distorted, the observer is interested in obtaining a reconstruction μ of the original clean signal. If *p(x)* is the probability measure for *x*, the

1. The author is grateful for the suggestions received at an OFCE seminar in June 2011, relating to applications of the methodology to the two models investigated here. Any errors are of course the author's.

best reconstruction satisfies the maximum a posteriori criteria max $p(\mu \mid d)$. Bayes' theorem states that this should be proportional to a prior probability on the reconstruction, $p(\mu)$, multiplied by the likelihood $p(d|\mu)$ that the observed data originated from the reconstructed signal.

The MaxEnt methodology assumes an entropic prior of the form $p(\mu) \propto \exp(S(\mu \mid m))$, where $S(\mu \mid m)$ is the relative Shannon (1948) entropy between the reconstruction μ and a model m, which is the observer's ex ante guess of the reconstruction, based on the data d. This choice of prior in the image restoration literature is underpinned by the rigorous bayesian formulation of Shore and Johnson (1980), who provide an axiomatic proof that the entropy measure S is the only prior that does not introduce biases into the reconstruction. As a result, the reconstruction can be identified as the one that maximises the following expression, where $\ell(d \mid \mu)$ is the log-likelihood $\log(p(d \mid \mu))$:[2]

$$p(\mu \mid d) \propto \exp(\alpha S(\mu \mid m) + \ell(d \mid \mu)) \tag{1}$$

The MaxEnt methodology was initially introduced in economics by Foley (1994) and extended by Toda (2010) as a way of deriving the statistical equilibrium of a market, *i.e.* the equilibrium distribution of endowments over agents. In a companion paper to the present study, Barde (2012) shows that the problem of allocating goods between rational agents can be modeled as a congestion game that possesses the finite improvement property. This means that any initial condition is linked to a Nash equilibrium by a finite path. Because each step on this path is the result of agents performing welfare-increasing trades, the reversed improvement path (which starts at the Nash equilibrium and ends at the initial condition) can be interpreted as a noise process, where agents make systematic mistakes. This is shown to imply that the problem of predicting the Nash equilibrium from the initial condition is formally equivalent to the problem of retrieving an image that has been corrupted by noise.

2. The multiplicative α term allows for the fact that the entropic prior $p(\mu)$ is defined only up to a multiplicative constant. α therefore effectively serves as a lagrangian parameter for the maximisation.

The image reconstruction interpretation of MaxEnt, which rests on the existence of a finite improvement path linking initial and final states, suggests that the use of relative entropy in the prior increases the flexibility of the methodology compared to Shannon (1948) entropy suggested in Foley (1994). This is because prior knowledge of the initial condition and of the fact that the initial and final states are linked by a finite path reduces the uncertainty of the observer with respect to the final state. This should be reflected in the entropy measure uncertainty by the inclusion of a correction term for this prior knowledge, embodied in m. A specific aspect of this, raised in Barde (2012), is that the model term m should reflect the length of the finite improvement path. If the improvement path is known to be short, the model should be strongly peaked around the initial condition. Conversely, if the path is long, the model should be flatter, reflecting the fact that the initial condition is no longer informative as to the final equilibrium.

In Barde (2012) the use of MaxEnt image reconstruction as a prediction methodology is investigated by applying it to the Schelling (1969, 1971) model of segregation. This empirical application was chosen specifically because for a given set of parametrisations the Schelling model is known to possess the finite improvement property, where every initial condition leads to a Nash equilibrium in a finite number of steps. The Schelling model thus provides an ideal setting for illustrating the image reconstruction interpretation detailed above.

The purpose of this companion paper is to investigate the MaxEnt methodology further, by attempting to predict the outcome of two agent-based models with recruitment, the Kirman (1993) model of ants and the Abrams and Strogatz (2003) model of language competition, a type of voter model. In both of these models there exists different populations, and agents within them can be recruited, *i.e.* convinced to switch group, by social pressure from members of other groups. As a result, the growth of one type of population depends on the size of the other populations. This setting is more complicated to predict than that of the Schelling model, as population sizes are not constant and a final absorbing state may not even exist. Nevertheless, our first central finding is that MaxEnt can predict the evolution of these models. The second important finding, which stems from the image reconstruction interpretation of MaxEnt presented above, is the confirmation that the width of the underlying model m does indeed play a role in controlling the time horizon of the MaxEnt prediction, which strongly supports the use of relative entropy rather than absolute Shannon (1948) entropy.

The remainder of the paper is organised as follows. Section 1 first presents the Kirman model of ant recruitment and then investigates the effectiveness of the MaxEnt methodology. Section 2 does the same on the voter model, and finally section 3 concludes.

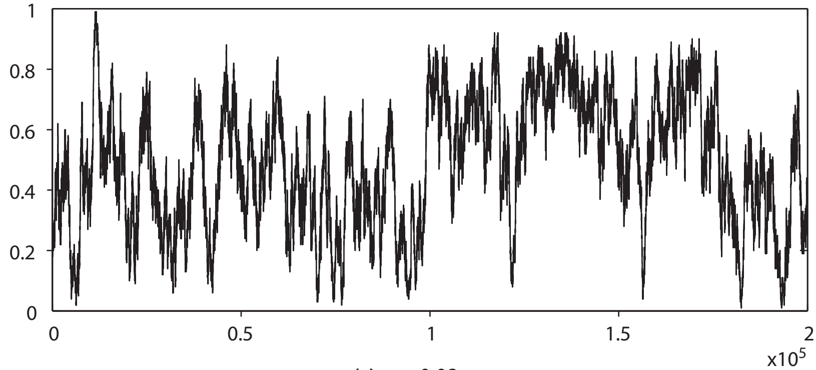
1. The one-dimensional problem: Kirman's model of ants

The Kirman (1993) model of ant recruitment was initially developed to provide a theoretical explanation for a curious empirical puzzle in an experiment involving ants feeding from two different food sources. In the wild, ants that encounter a food source recruit other ants, quickly causing a large amount of ants to feed from that source. The experimental puzzle came from the "cascading" behaviour exhibited by the ants, where most ants used a single source of food for a period of time, and then suddenly switched to the other in a very short period. The central advantage of starting with this model is that because of the simple recruitment process that governs the evolution of the system, the limit distribution of the system is well known, which facilitates the process of verifying the improvement in prediction brought by the use of the MaxEnt methodology.

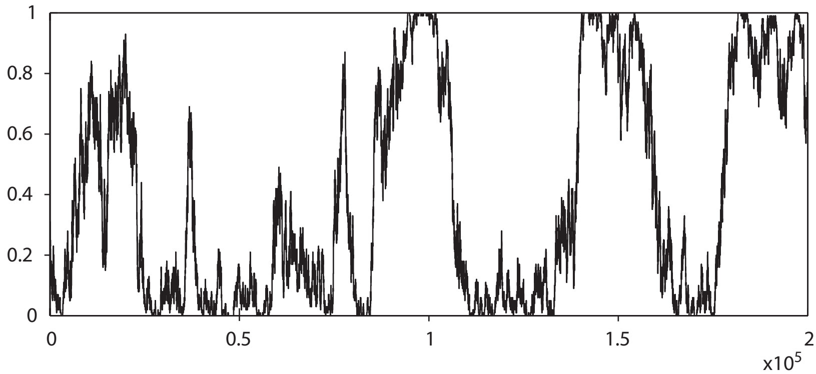
1.1. Kirman's model of ants

In this model, two sources of food are available to a group of N ants, which are denoted "black" and "white". Describing the state of the system is simple: at any point in time, let $k \in \mathbb{N}$ be the number of ants feeding from the black source, with the remaining $N - k$ ants feeding from the white source. In the following discussion, it will be convenient to refer to $x = k/N$ as the share of the ant population feeding at the black source, with $1 - x = (N - k)/N$ the share feeding from the white source. By extension, we will refer directly to the color of the ant as identifying the food source it uses.

Figure 1. Time evolution of the share of black ants x



(a) ε = 0.02



(b) ε = 0.005

As pointed out in Kirman (1993), ants can change color over time, either spontaneously (by making a mistake, for instance), or because they are recruited by an ant of the other color. Because of this, the system will evolve over time. If ε is the probability of an ant spontaneously changing color and $1 - \delta$ the probability of an encounter between two differently coloured ants leading to a successful recruitment, then the dynamic evolution of the system is governed by the following probabilities:

$$\begin{cases} p_{w \to b}(k) = \left(1 - \dfrac{k}{N}\right)\left(\varepsilon + (1-\delta)\dfrac{k}{N-1}\right) \\ p_{b \to w}(k) = \dfrac{k}{N}\left(\varepsilon + (1-\delta)\dfrac{N-k}{N-1}\right) \end{cases} \tag{2}$$

At every point in time, the number of black ants k can therefore jump, either $k \to k+1$ with probability $p_{w\to b}$, or conversely $k \to k-1$ with probability $p_{b\to w}$. One can see that for large values of the ant population N, the share of black ants $x = k/N$ can be approximated by a continuous interval [0,1], which allows us to rewrite the transition probabilities (2):

$$\begin{cases} p_{w\to b}(x) = \varepsilon(1-x) + (1-\delta)x(1-x) \\ p_{b\to w}(x) = \varepsilon x + (1-\delta)x(1-x) \end{cases} \tag{3}$$

As is the case in Kirman (1993), we assume as a simplification that $\delta = 2\varepsilon$, which allows the simulations to depend on a single parameter. Figure 1 shows the evolution of the state of $N = 100$ ants over time for the two main parametrisations of ε used by Kirman (1993). Of the two, the second case, where $\varepsilon = 0.005$ is the most interesting, as it displays the cascading transitions mentioned previously.

The research question initially addressed by Kirman (1993) was to find the limit distribution $\mu(k)$ of the proportion of time the system spends in a state k. We show below that the MaxEnt methodology can not only replicate this finding, but in fact provide a more general prediction of the time-density of the system for any number of steps τ.

1.2. Prior model and likelihood

We start by specifying a model $m(x, \tau)$ for the relative entropy term in (1). As stated in the introduction, one would intuitively expect this to change depending on the desired time-horizon of the prediction. For low values of τ (short horizons), one would expect the model to be peaked around the initial condition x_0. Conversely, for large values of τ (long horizons) the system will be able to explore large areas of the state space, and the model should be flatter.[3] This movement away from the initial condition x_0 after τ steps is modeled by the diffusion process of 1-dimensional stopped random walk. Given the transition probabilities (3), at any point in time the probability that a jump occurs is $p_j(x) = p_{w\to b}(x) + p_{b\to w}(x)$, while the system remains unchanged with probability $1 - p_j(x)$.

3. As pointed out in Barde (2012), if the system is ergodic, then in the limit $\tau \to \infty$, the model should be a uniform distribution, as all states become accessible.

$$p_j(x) = \varepsilon + 2(1-\delta)\, x\, (1-x) \tag{4}$$

Comparing equations (4) and (3), one can see that in the limit $\varepsilon \to 0$, we have $p_{w\to b}(x) = p_{b\to w}(x) = p_j(x)/2$. This also holds for all values of ε if $x = 0.5$. As a result, the diffusion away from a known initial state x_0 after τ units of time have elapsed can be approximated by a one-dimensional random walk where $p_j(x_0)\tau$ jumps occur, each of which takes values $\{+1/N, -1/N\}$ with equal probability. The standard result for such a process is that the probability at time τ of having moved a specified distance is given by a binomial distribution with probability parameter 1/2. Because the prediction $\mu(x, \tau)$ relates to the predicted share of time system will spend in each state, the model is obtained by averaging the binomial density over the expected $p_j(x_0)\tau$ jumps for each value of x[4]

$$m(x,\tau) = \frac{\displaystyle\sum_{T=0}^{T=\tau p_j(x_0)} 2^{-T} \binom{T}{\dfrac{T+N|x-x_0|}{2}}}{\tau p_j(x_0)+1} \tag{5}$$

Clearly this is an imperfect representation of the diffusion generated by the recruitment process (3), as one can immediately see from Table 3 in appendix A that after $\min(k_0, N-k_0)$ jumps, there is a non-zero probability that the random walk process has gone beyond the [0,1] bound for x. This is because the model (5) assumes that probability of a jump $p_j(x_0)$ is constant, and the probability of a positive and negative jump is always equal to 1/2. This is not the case in the actual process (3), as the transition probabilities adjust to guarantee the process remains within the bounds. It will be shown, nevertheless, that this simple random walk diffusion (5) provides a reliable model for the MaxEnt prediction.

The likelihood term for the MaxEnt program (1) can be obtained from the net transition probabilities. The intuition is that the transition probabilities (3) provide a stochastic growth process for each population, which can be integrated to provide an expected time path. However, given the fixed overall number of ants, it must be that these expected time paths of both populations

4. The specification used for calculating the time-average of the binomial density is explained in appendix A.

cancel out for the limiting distribution. As mentioned previously, an attractive aspect of the ants model is that the transition probabilities in fact allow for direct derivation of the limit distribution as $\tau \to \infty$. A particularly elegant derivation is provided in the appendix of Alfarano and Milaković (2009), which uses a model of herding that is very similar to Kirman (1993).

As shown by Alfarano and Milaković (2009), deriving the limit distribution directly requires a second order approximation of the transition process, using a drift and a diffusion term obtained through a Taylor expansion of the transition process.[5] The likelihood obtained below uses only the drift term, *i.e.* a first-order approximation of the process. Given the number of black ants k, the share of black ants x, the transition probabilities (3) lead to the following expected state after a jump:

$$(k+1)\, p_{w\to b}(x) + (k-1)\, p_{b\to w}(x) + k\left(1 - p_j(x)\right) = k + \varepsilon(1-2x) \tag{6}$$

Assuming that an interval of time $[t, t+1]$ is short enough that a only single jump is expected to occur, this expected jump directly determines the expected change in the share of black agents x:

$$E\left[x_{t+1}\right] - x_t = \frac{\varepsilon}{N}\left(1 - 2x_t\right) \tag{7}$$

One can see from this expression that assuming a minority of black ants ($x < 1/2$) the expected change in the share of black ants is positive. Conversely, if $x > 1/2$ and a majority of ants are black, one would expect to see the share of black ants k fall. Thus, the expectation is that the transition probabilities will bring the state towards $x = 1/2$ over time.[6] Dividing on both sides by x gives the expected growth rate of the black ant population during over the time interval:

$$\frac{E\left[x_{t+1}\right] - x_t}{x_t} = \frac{\varepsilon}{N}\left(\frac{1}{x_t} - 2\right) \tag{8}$$

5. This is outlined in appendix B.

6. Expression (7) helps to clarify the simplifying assumption that $\delta = 2\varepsilon$. Because in (3) the probability of a black ant recruiting a white ant is equal to the probability of a white ant recruiting a black ant, the expected effects cancel out and δ does not enter the expected change in population (6).

As we are assuming a large ants population N, the share of black ants $x \in [0, 1]$ can be treated as a continuous variable. As a result, the left hand side of this expression can be expressed as the time-derivative of a logarithm, $\partial \ln x / \partial t$, leading to an ordinary differential equation. For τ units of time, the expected time path of the black ant population is approximated following expression, where $\mu(x, \tau)$ is the share of time spent in state x and $\ln x$ is the expected time path in that state:

$$\sum_x \mu(x,\tau) \ln x \tag{9}$$

Given the inherent symmetry of the system, it is possible to obtain a similar expression involving $\mu(x, \tau)\ln(1 - x)$ for the expected time path of the white ant population.[7] Furthermore, because the overall number of ants is fixed at N and growth of one population implies an equivalent reduction in the other, the sum of the two time paths should cancel out. This is used for formulate the following likelihood for a candidate prediction μ:[8]

$$\ell(\mu(x,\tau)) = \sum_x \mu(x,\tau)(\ln x + \ln(1-x)) \tag{10}$$

Given the model (5) and likelihood (10), the MaxEnt program for the share of time τ spent in state x in the ants recruitment model is given by:

$$\max_{\mu} \left(\alpha S(\mu(x,\tau) | m(x,\tau)) + \ell(\mu(x,\tau))\right) \tag{11}$$

The first order condition of (11) provides the predicted value of $\mu(x, \tau)$. As explained in Barde (2012), one can see that this is effectively a mixture density between the model (5) and the limit distribution, which in this case is a symmetric beta distribution:

$$\mu(x,\tau) \propto m(x,\tau) x^{-\frac{1}{\alpha}} (1-x)^{-\frac{1}{\alpha}} \tag{12}$$

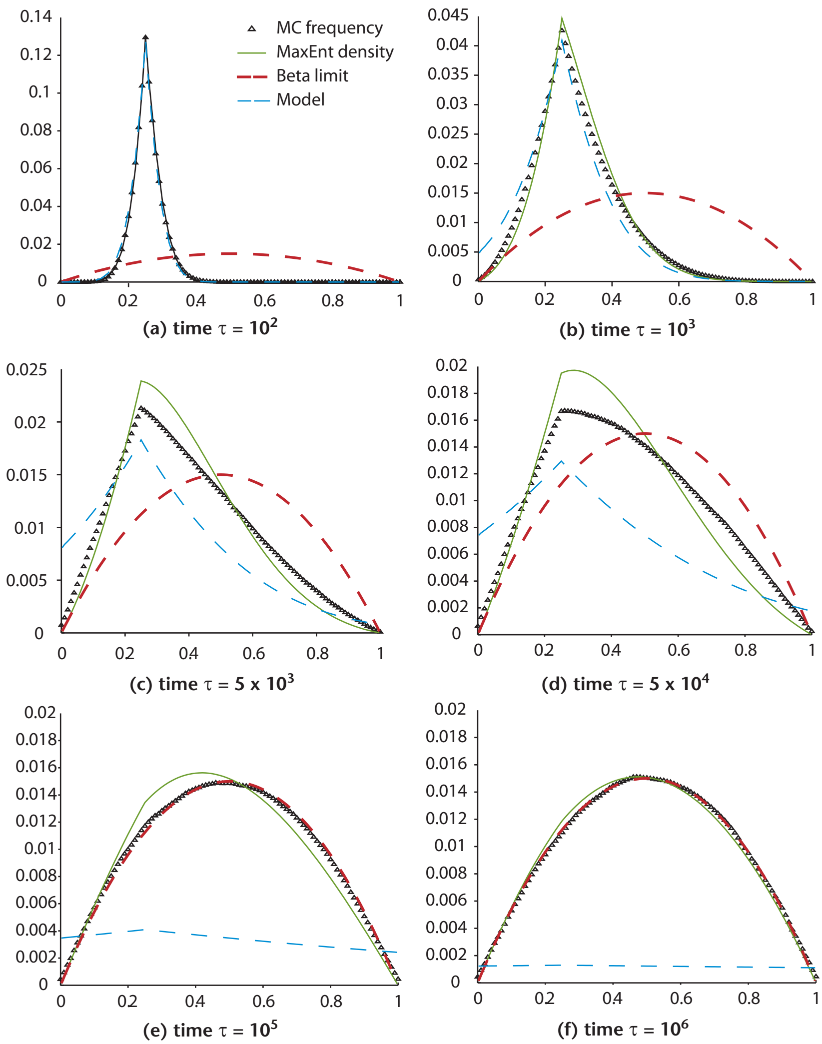
The alpha parameter in (12) is effectively the Lagrange parameter from the maximisation problem (11), and controls the relative weight of the entropy and likelihood terms. In this case,

7. In fact, there is no a priori reason for the state of the system to be measured using the number of black ants k. Intuitively, using instead the number of white ants $N - k$ as the state variable should not change the predictions that an observer can make about system.

8. One can see that this expression exhibits the important properties of a log-likelihood: given a candidate distribution μ and $x \in [0, 1]$ it will be negatively valued, and for the limit beta distribution obtained below, it reaches a maximum value of zero.

one can see that as $\tau \to \infty$ and the model term $m(x, \tau)$ becomes a uniform distribution, the value of α also controls the exponent of the limit beta distribution. As shown in appendix B, the limit distribution is is simply $x^{N\varepsilon-1}(1-x)^{N\varepsilon-1}$, which corresponds with the one identified in Kirman (1993). We therefore set $\alpha = -1/(N\varepsilon - 1)$.

Figure 2. Time evolution of the MC frequencies vs. MaxEnt prediction, $x_0 = 0.25$, $\varepsilon = 0.02$



1.3. Maximum entropy prediction of the ants model

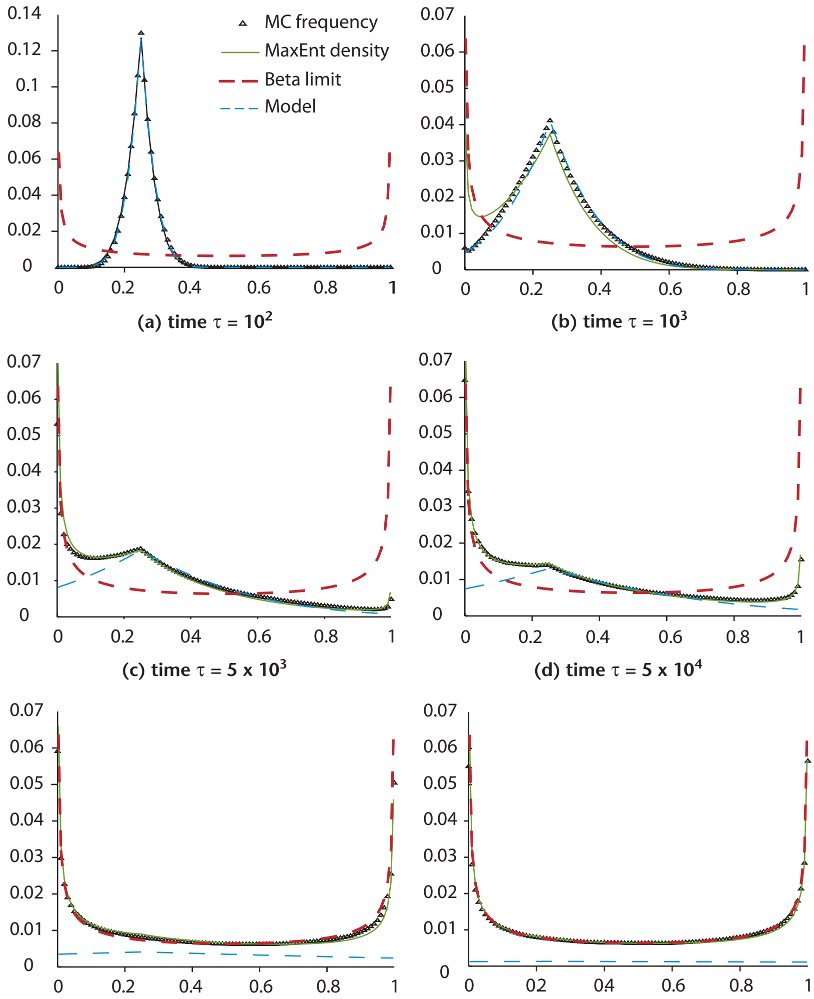
The reliability of the mixture density (12) in predicting the proportion of time τ the system spends in each state x is assessed by comparing the MaxEnt prediction to a Monte-Carlo simulation of the system defined by the transition probabilities (3), using the same values of ε as for figure 1. In both cases, $N = 100$ and the initial condition was set to $x_0 = 0.25$, in order to explicitly examine how the system moves from an asymmetric initial condition to its symmetric limit distribution.[9]

In both figures 2 and 3, the Monte-Carlo frequencies are represented by the sequence of triangular markers, while the solid line represents the MaxEnt prediction (12). As an illustration of how the mixture density is reached, the dashed lines represent the components of this prediction, with the thin dash representing the model (5) and the thicker dash showing the limit beta distribution. Goodness-of-fit statistics are displayed in Table 1 for both parametrisations, and report the Spearman rank correlation ρ and mean square error of the prediction (12) relative to the variance of the Monte-Carlo frequency of the for each of the time steps in the figures.

The first case, shown in figure 2, uses the setting that produced the path shown in figure 1a. As expected, in the early stages of the process, for low values of τ, the prediction is dominated by the model term (5), and over time it gradually converges to the limit beta distribution. For intermediate values of τ (particularly $\tau = 5 \times 10^3$ and $\tau = 10^4$), one can see that the empirical frequencies are converging towards the limit distribution faster than suggested by the prediction. Nevertheless, the mean square error of the prediction relative to the variance of the Monte-Carlo frequencies remains low even for these intermediate values of τ. Furthermore, an important aspect is that the prediction successfully captures the asymmetry in the empirical frequencies about x_0 whenever the initial condition is not located at 1/2.

9. The number of Monte-Carlo iterations R carried out for a given time-horizon τ is $R = 10^8/\tau$. The implication is that the resulting time-averages in figures 2 and 3 are all calculated over 10^8 time steps, the only difference between sub-figures being that the system is essentially reset to the initial condition x_0 every τ steps. This this is done to ensure that the goodness-of-fit statistics, which are based on the variance of the Monte-Carlo frequencies, are comparable across time horizons τ.

Figure 3. Time evolution of the MC frequencies *vs.* MaxEnt prediction, $x_0 = 0.25$, $\varepsilon = 0.005$



(a) time $\tau = 10^2$

(b) time $\tau = 10^3$

(c) time $\tau = 5 \times 10^3$

(d) time $\tau = 5 \times 10^4$

(e) time $\tau = 10e^5$

(f) time $\tau = 10^6$

Table 1. Goodness of fit, MaxEnt prediction *vs.* MC frequencies

		$\tau = 10^2$	$\tau = 10^3$	$\tau = 5\text{x}10^3$	$\tau = 10^4$	$\tau = 10^5$	$\tau = 10^6$
	Spearman ρ	0.9418	0.9985	0.9975	0.9969	0.9537	0.9917
ε = **0.02**	p-value	$5.74\text{x}10^{-49}$	0	0	0	0	0
	MSE/var	0.0011	0.0154	0.0669	0.1183	0.0553	0.0092
	Spearman ρ	0.9333	0.9716	0.9923	0.9989	0.8816	0.9523
ε = **0.005**	p-value	$3.96\text{x}10^{-46}$	0	0	0	0	0
	MSE/var	0.0015	0.1378	0.0664	0.0098	0.04	0.0224

The second case, in figure 3, corresponds to the cascading path in figure 1b. As for the previous setting of ε, the goodness-of-fit is shown in the second set of rows in Table 1. The qualitative behaviour is similar, with the model component of the prediction dominating in the early stages and a convergence to the limit distribution for large values of τ. Again, the predicted and Monte-Carlo frequencies deviate slightly for intermediate values of τ ($\tau = 10^3$ in this case), but the relative mean square error is low and as for the previous case the methodology correctly predicts the asymmetry involved in shifting from an early distribution that is practically symmetric about x_0 to a limit distribution that is symmetric about 1/2.

Two important observations stem from these results. The first is that the MaxEnt methodology produces a good prediction of the proportion of elapsed time τ the system will spend in each state x, for all values of τ, and not just in the limit $\tau \rightarrow \infty$. This provides an improvement with respect to Kirman (1993), as it allows a description of all the phases of the adjustment from an initial condition to the limit distribution. The MaxEnt prediction is even able to capture the asymmetry in the distribution caused by the adjustment from an arbitrary distribution to the symmetric limit. While this adjustment to the limit distribution can also be obtained using a traditional Monte-Carlo approach, the MaxEnt result can be obtained at a greatly reduced computational cost.

The second observation, which results from this ability to predict over a wider range of time horizons, is that when observed over short horizons the ant recruitment process (3) behaves very much like a simple random walk (5). Indeed the early stages of the adjustment, the model term dominates the mixture distribution completely. It is only over longer time horizons that it converges to the limit distribution identified in Kirman (1993). This suggests that in practical terms it might be very difficult to distinguish recruitment processes from a random walks over short time horizons by looking only at the time density of states.

2. The two-dimensional problem: Voter models

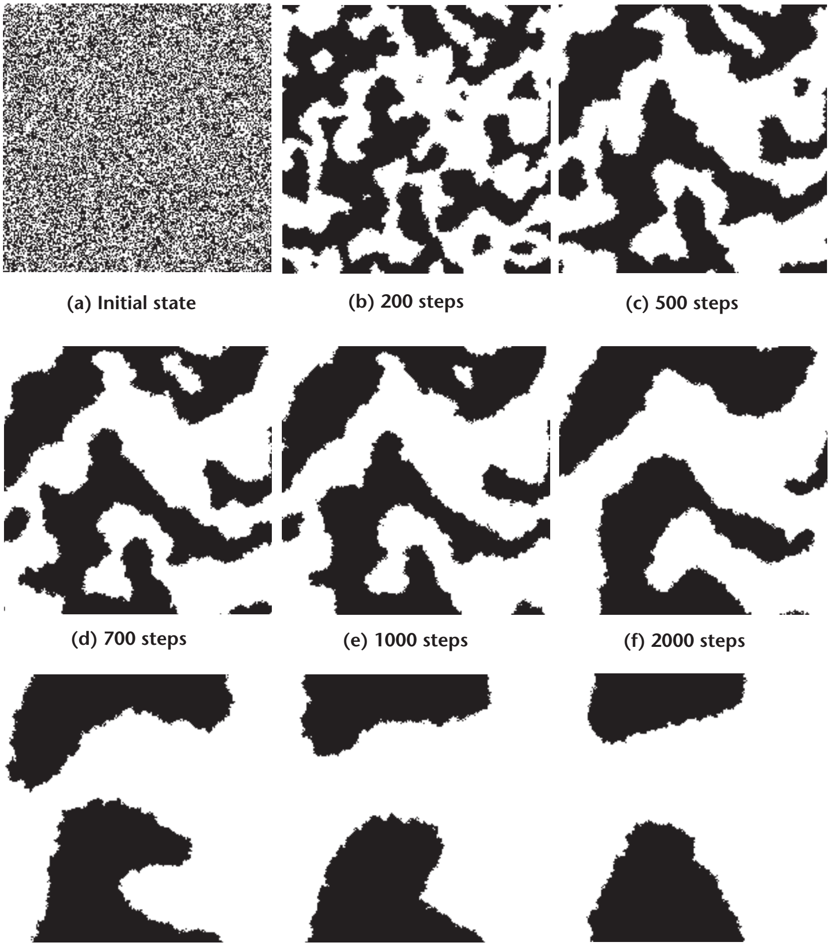
Voter models, also known as consensus models, are similar in spirit to the ants model seen in the previous section. Typically, several populations coexist in the same space and members of each group attempt to convince members of competing groups to switch over, much like the recruitment process described above. One of the main attractions of these models is that it is straightforward to integrate localised spatial effects that are similar to those in the Schelling (1971) model, which provides a further setting for investigating where the predictive power of the MaxEnt methodology.

2.1. The Abrams-Strogatz model of language competition

In the Abrams and Strogatz (2003) model of language competition, two languages, W and B are spoken within a population, and individuals switch from one language to the other according to its attractiveness. Assuming, as was the case for the ants recruitment model, that x is the share of individuals speaking language B and $1 - x$ is the share speaking W, the transition probabilities (13) are determined by three elements. The first is the intrinsic prestige of the languages, controlled by a parameter $s \in [0, 1]$ for B and $1 - s$ for W. The second element is the effect of social pressure, as measured by the shares x and $1 - x$ of individuals speaking the language, and the third is a volatility parameter a which increases or reduces the effect of social pressure through exponentiation of this social pressure term:

$$\begin{cases} p_{w \to b}(x) = sx^a \\ p_{b \to w}(x) = (1-s)(1-x)^a \end{cases} \tag{13}$$

Figure 4. Simulated time evolution of a voter model, $s = 0.5$, $a = 2$



(a) Initial state (b) 200 steps (c) 500 steps (d) 700 steps (e) 1000 steps (f) 2000 steps (g) 5000 steps (h) 7500 steps (i) 10000 steps

As a result of these transition probabilities, the dynamic evolution of the system is given by:

$$\frac{dx}{dt} = (1-x)p_{w\to b}(x) - xp_{b\to w}(x) = x(1-x)\left(sx^{a-1} - (1-s)(1-x)^{a-1}\right) \quad (14)$$

The equilibrium predictions are qualitatively similar to the ones obtained in Kirman (1993) and shown in figures 2 and 3. The system displays one interior equilibrium $0 < x < 1$ and two corner solutions at $x = 0$ and $x = 1$.[10] In the low volatility case, where $a > 1$, the corner solutions are stable and the interior solution is unstable, while the opposite is true of the high volatility case $a < 1$.

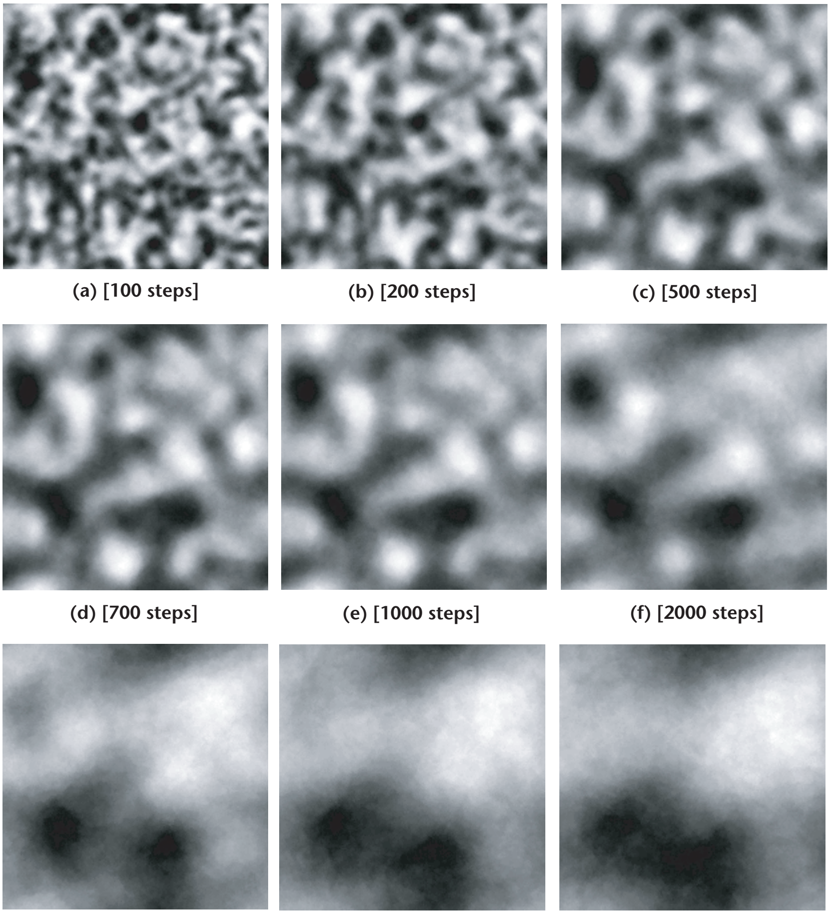
There are several key differences, however, compared to the ant recruitment model in section 1. These are due to the assumption that individuals cannot accidentally switch languages, as was the case in the ants model through the ε term in the transition probabilities (3). As a consequence, once a corner equilibrium $x = 0$ or $x = 1$ is reached in the voter model, the system will remain in that state, which is not the case in the ant model, as is visible in Figure 1b. This clarifies why in the voter model the discussion centers on the stability or instability of the three defined equilibria, while the ant model focuses instead on the share of total time spent in each state. The absence of a ε term allowing accidental individual switching also explains the difference in the parameter that controls the amount of social interaction and thus the stability of equilibrium or shape of the distribution. In both models, the probability of an interaction between agents of different colour is given by $x(1 - x)$. In the ants model, this is translated into a switching probability through the additive ε term, while in the voter model this is done by exponentiating the interaction probability with $a - 1$.

In an important analysis of this model of language competition, Stauffer *et al.* (2007) show that simulations carried out with a finite number of individuals produces different results compared to the continuous equations shown above.[11] Furthermore, they show that local interaction matters for understanding the dynamics of the system and the time until a stable equilibrium is reached. Local interaction is defined as a situation where the social pressure on any given agent to switch from language B to W comes from the share of the agent's direct neighbours that already speak W rather than the share of the overall population that speaks W, as is the case in (14).

---

10. The interior equilibrium is located at 1/2 in the case of two equivalent languages, with prestige $s = 1 - s = 1/2$. This is the parameter value that was used in the simulation reported below.
11. In Stauffer *et al.* (2007) the continuous version is referred to as a "mean-field approximation".

Figure 5. **Time evolution of voter model state frequencies, 200 MC iterations**



(a) [100 steps] (b) [200 steps] (c) [500 steps]

(d) [700 steps] (e) [1000 steps] (f) [2000 steps]

(g) [5000 steps] (h) [7500 steps] (i) [10000 steps]

Stauffer *et al.* (2007) focus on the case where the languages have the same prestige ($s = 1 - s = 1/2$) and the volatility is low ($a > 1$), therefore this is the parameter setting that will be used below to investigate the effectiveness of the MaxEnt predictions. In the following simulation there are $N = 40000$ agents arranged in a 200 x 200 lattice. The neighbourhood which determines the local social pressure to switch language is a 3 x 3 square centered on the agent of interest. As was the case in the analysis of the Schelling model in Barde (2012), the space occupied by the agents is assumed to be toroidal, which means that localised neighbourhood effects can be calculated directly by applying a $N \times N$ circulant matrix A to the state vector, a $N \times 1$ vector recording the language spoken by each agent. This implies that x, the share of agents speaking B and $1 - x$, the share speaking W in the transition probability (13) is a local variable that is determined by the 8 closest neighbours of an agent rather than a global variable, averaged over the overall population.

Figure 4 presents the simulated state of such a system at several points in time, starting from a random initial condition where half the agents speak B and half speak and W. One can see that starting from a dispersed state in the initial condition, relatively few steps suffice for distinct clusters to appear.[12] As the number of steps is increased, the smaller clusters tend to disappear and the interfaces between the large clusters of different colours tend to smooth out, a process which Stauffer *et al.* (2007) refer to as an increase in the surface tension of the system. Another characteristic outlined by their analysis is the existence of long-lived meta-stable equilibria. This is visible in the last few panels of figure 4, where the two clusters remain similar over a large number of steps. In the limit, however, the system always ends up in one of two absorbing states, $x = 0$ or $x = 1$.

2.2. Maximum entropy prediction of the voter model

The MaxEnt methodology used to predict the evolution of the state of the voter model is broadly similar to the one used for the analysis of the Schelling (1971) model of segregation carried out in Barde (2012). First of all, a Monte-Carlo analysis was run in order to obtain a point of comparison for the MaxEnt prediction. 400 random initial conditions were drawn and for each of these 200 separate simulations were run, replicating the process shown in in figure 4. As an illustration, figure 5 shows the result of running 200 such simulations on the initial condition provided in figure 4a. Each sub-figure shows the share of runs in which an agent is in state B after the specified number of steps.

The setting of the voter model implies that agents do not have ex-ante preferences for a language, and only determine their state

12. As is the case in Stauffer *et al.* (2007) one step in time corresponds to one update opportunity for all N agents to update their state.

relative to the language spoken by neighbouring agents. As a result, assuming that languages *B* and *W* have equal status and initial populations have equal size, the initial state of any given agent *i* does not provide any information that can be used to directly provide a model for the relative entropy term (1). The best guess an observer might make is that $m_i^w = m_i^B = 0.5$, which is the uninformative uniform distribution. This situation is effectively the same as in the Schelling (1971) model, in which agents do not have any intrinsic preference for a particular location, and the attractiveness of a location to an agent is only determined by the state of the agents around that location. As was the case in Barde (2012), this is dealt with by using the following double-space entropy, which measures the information entropy of a message revealing the state of two randomly picked agents *i* and *j*:

$$S(\mu_i, \mu_j | m_{i,j}) = \frac{1}{N}\left(-2\sum_i \sum_L \mu_i^L \ln \mu_i^L + \frac{1}{N}\sum_{i,j}\sum_L \mu_i^L \mu_j^L \ln m_{i,j}^L\right) \quad (15)$$

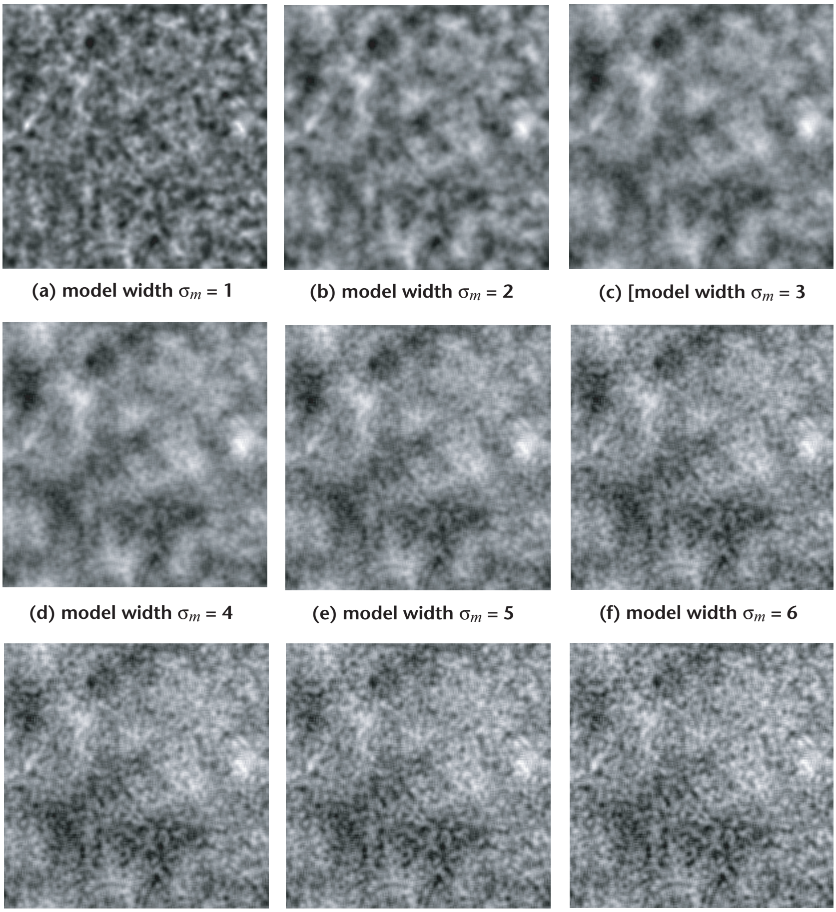
The use of a double-space entropy (15) allows the model to encode correlations across agents: if two agents *i* and *j* are located close to each other, the probability that they both speak a given language *L* is higher than if they are far from each other.[13] The model $m^L_{i,j}$, which models the probability that *i* and *j* both speak language *L* is assumed to be a normal distribution over the distance between agents, as was the case in the analysis of the Schelling model in Barde (2012). Because the width of this normal distribution is determined by its standard deviation σ_m, this parameter directly controls the distance over which agent decisions are likely to be correlated.

The likelihood component of (1) that is used to generate the prediction is assumed to be Gaussian, following again the methodology used in Barde (2012). This effectively measures the similarity be the data available in the initial condition and the MaxEnt prediction for each of the two languages *L*:

$$\ell\left(d^L \mid \mu^L\right) = -\sum_i \frac{\left(\left(A\mu^L\right)_i - d_i^L\right)^2}{(\sigma)^2} = \frac{-\chi^2\left(\mu^L\right)}{2} \quad (16)$$

13. The reader is referred to the appendix of Barde (2012) for a derivation of double space entropy.

Figure 6. MaxEnt state density prediction of voter model



(a) model width $\sigma_m = 1$ (b) model width $\sigma_m = 2$ (c) [model width $\sigma_m = 3$

(d) model width $\sigma_m = 4$ (e) model width $\sigma_m = 5$ (f) model width $\sigma_m = 6$

(g) model width $\sigma_m = 7$ (h) model width $\sigma_m = 8$ (i) model width $\sigma_m = 9$

Here A is the $N \times N$ symmetric adjacency matrix, with entries 0 or 1 in the i^{th} row indicating which J agents are neighbours to i. The data d_i^L used in the comparison is given by the initial social pressure $d^L = A\mu_0{}^L$, where $\mu_0{}^L \in \{0, 1\}$ is the vector indicating the language spoken by each agent in the initial condition, taken from figure 4a. It is important to point out that given the choice of parameterisation $a = 2$, the local interaction term $A\mu^L$ in the likelihood is in fact a linear approximation of the social pressure term in the transition probabilities (13).[14] It will be shown that this simplification nevertheless produces good predictions of the Monte-Carlo frequencies.

Given the relative entropy (15) and likelihood (16), the MaxEnt program is given below. Its solution, displayed in figure 6 for increasing values of σ_m (and therefore increasing model widths) is obtained numerically using the image reconstruction algorithm of Skilling and Gull (1991), modified in Barde (2012) to predict the outcome of the Schelling model.

$$\max_{\mu}\left(\alpha S\left(\mu_i,\mu_j \middle| m_{i,j}\right)+\ell\left(d^L \mid \mu^L\right)\right) \tag{17}$$

Figure 6 illustrates the MaxEnt predictions obtained with the initial condition provided in figure 4a for various values of σ_m. These are visually comparable with the Monte-Carlo frequencies in figure 5, also generated from the same initial condition. In addition, the Spearman rank correlation and mean-square-error relative to the variance of the Monte-Carlo frequencies were calculated for each of the 400 sets of predictions and frequencies. The resulting means and standard deviations are presented in Table 2. The bold entries in each column indicate identify the value of σ_m (row) that best fits the Monte-Carlo frequencies for the relevant number of steps. The diagonal pattern made up by these bold entries indicates that the Monte-Carlo frequencies of the system at successively higher time steps are, up to a point, better predicted by successively wider models. As was the case for the MaxEnt prediction of the ants recruitment model in section 1.3, this supports the suggestion that the width of the model in the relative entropy term, determined in this case by σ_m, controls the time-horizon of the MaxEnt prediction.

A further observation that can be made from Table 2, however, is that the predictive power of the MaxEnt methodology falls as the width of the model is increased. Indeed, the bold entries in each column show a gradual reduction in the correlation coefficient ρ and an increase in the size of the mean-square error as the number of steps is increased, coupled with a widening of the standard deviations around the means of the two statistics.

14. This is intended as a simplification: using $d^L = (A\mu_0{}^L)^2$ and $(A\mu^L)^2$ in (16) produces a Hessian matrix for the MaxEnt methodology where all $N \times N$ entries are non-zero, requiring an intractable amount of storage and computation time. Using instead the linear approximation produces a Hessian matrix that is basically $A \times A$, and therefore can be stored and manipulated efficiently using sparse matrix algorithms.

Table 2. Goodness of fit, predicted vs. MC voter model state densities (standard deviations in parenthesis)

		Steps								
σ_m		100	200	500	700	1000	2000	5000	7500	10000
1	ρ	0.859	0.740	0.562	0.498	0.434	0.325	0.215	0.179	0.157
		(0.006)	(0.011)	(0.017)	(0.019)	(0.021)	(0.024)	(0.026)	(0.026)	(0.027)
	MSE	0.291	0.518	0.865	0.992	1.120	1.340	1.561	1.635	1.678
		(0.011)	(0.020)	(0.033)	(0.037)	(0.042)	(0.049)	(0.052)	(0.052)	(0.054)
2	ρ	**0.950**	**0.921**	0.810	0.752	0.685	0.548	0.383	0.323	0.287
		(0.004)	**(0.007)**	(0.016)	(0.020)	(0.025)	(0.034)	(0.042)	(0.045)	(0.047)
	MSE	**0.114**	**0.167**	0.376	0.487	0.618	0.888	1.217	1.337	1.411
		(0.007)	**(0.012)**	(0.028)	(0.036)	(0.046)	(0.065)	(0.082)	(0.088)	(0.094)
3	ρ	0.887	0.918	**0.898**	0.869	0.826	0.713	0.538	0.465	0.417
		(0.009)	(0.008)	**(0.014)**	(0.018)	(0.024)	(0.036)	(0.052)	(0.059)	(0.064)
	MSE	0.228	0.170	**0.205**	0.259	0.340	0.557	0.901	1.047	1.142
		(0.016)	(0.013)	**(0.023)**	(0.031)	(0.041)	(0.066)	(0.101)	(0.115)	(0.127)
4	ρ	0.813	0.861	0.890	**0.884**	**0.865**	0.791	0.638	0.564	0.513
		(0.015)	(0.014)	(0.016)	**(0.019)**	**(0.023)**	(0.035)	(0.056)	(0.066)	(0.074)
	MSE	0.366	0.274	0.217	**0.228**	**0.263**	0.403	0.698	0.845	0.945
		(0.026)	(0.022)	(0.026)	**(0.031)**	**(0.040)**	(0.064)	(0.106)	(0.128)	(0.145)
5	ρ	0.753	0.801	0.850	0.857	0.855	**0.815**	0.694	0.626	0.577
		(0.019)	(0.018)	(0.020)	(0.022)	(0.026)	**(0.037)**	(0.058)	(0.070)	(0.080)
	MSE	0.480	0.387	0.290	0.276	0.280	**0.355**	0.586	0.719	0.815
		(0.033)	(0.030)	(0.033)	(0.038)	(0.045)	**(0.066)**	(0.109)	(0.134)	(0.153)
6	ρ	0.708	0.747	0.801	0.814	0.821	0.806	0.717	0.659	0.615
		(0.022)	(0.022)	(0.025)	(0.027)	(0.031)	(0.041)	(0.061)	(0.073)	(0.083)
	MSE	0.566	0.489	0.383	0.356	0.341	0.370	0.539	0.653	0.740
		(0.038)	(0.036)	(0.042)	(0.047)	(0.055)	(0.075)	(0.114)	(0.139)	(0.159)
7	ρ	0.675	0.701	0.750	0.766	0.779	0.779	**0.718**	**0.669**	**0.630**
		(0.023)	(0.023)	(0.029)	(0.032)	(0.037)	(0.047)	**(0.065)**	**(0.076)**	**(0.086)**
	MSE	0.630	0.577	0.479	0.447	0.422	0.420	**0.538**	**0.632**	**0.708**
		(0.040)	(0.039)	(0.050)	(0.057)	(0.066)	(0.087)	**(0.123)**	**(0.145)**	**(0.164)**
8	ρ	0.652	0.663	0.702	0.718	0.732	0.742	0.700	0.660	0.627
		(0.022)	(0.023)	(0.033)	(0.037)	(0.042)	(0.054)	(0.071)	(0.081)	(0.090)
	MSE	0.675	0.651	0.571	0.540	0.512	0.491	0.571	0.649	0.714
		(0.040)	(0.041)	(0.059)	(0.068)	(0.078)	(0.100)	(0.134)	(0.154)	(0.172)
9	ρ	0.635	0.632	0.658	0.672	0.686	0.700	0.673	0.641	0.613
		(0.021)	(0.023)	(0.036)	(0.042)	(0.048)	(0.060)	(0.076)	(0.084)	(0.092)
	MSE	0.709	0.712	0.657	0.629	0.601	0.572	0.624	0.687	0.743
		(0.037)	(0.042)	(0.066)	(0.077)	(0.089)	(0.113)	(0.144)	(0.162)	(0.177)

The visual comparison of figures 5 and 6 also supports this: As the model with is increased the prediction gradually becomes more "grainy", to the point where it becomes difficult to distinguish an image. The intuitive conclusion that can be drawn from this is that there is a limit to the time-horizon over which reliable a prediction can be made.

3. Discussion and Conclusion

The analysis of the Kirman model of ant recruitment and the locally-interacting Abrams-Strogatz model of language competition both show that the maximum entropy methodology can be used to predict the state distributions of agent-based models with recruitment, where agents can switch groups based on a measure of social pressure. This provides support for the use of the MaxEnt image reconstruction methodology as a prediction methodology in economics. A first aspect is that MaxEnt can reliably predict the state space of these agent-based models, even in the case where there is no defined final state, as in the Kirman (1993) model of ant recruitment, or in the case where the transition of a given agent is probabilistic rather than a best response, as was the case with the initial MaxEnt analysis of the Shelling model carried out in Barde (2012). A second important aspect is the confirmation of the suggestion made in Barde (2012) that the width of the model term controls the time-horizon of the prediction, by serving as a proxy for the length of improvement path between initial and final state.

In methodological terms, the maximum entropy methodology used here and in Barde (2012) therefore provides a generalisation of the existing applications of MaxEnt in economics, mentioned in the introduction. These typically rely on Shannon (1948) entropy in their analysis, the justification being that this measures the absolute uncertainty of an observer as to the state of the system. As was suggested in Barde (2012) and demonstrated here, this implicity corresponds to using relative entropy with respect to a uniform model *m*. Given the link between model width and time horizon, this implies that Shannon MaxEnt predicts over a large time horizon only. The predictions obtained here, using relative entropy, carry over a much larger range of time. This is potentially relevant, as dynamic systems may behave differently over different time horizons.

This last point has potential implications given the suggestion made in Kirman (1993) as to the relevance of recruitment models in economics. Indeed, Kirman suggests that recruitment is pervasive in many markets, in particular financial markets, where individual agents make decisions based not only on objective information, but also based on imitation of surrounding agents. This point is reinforced by the use Alfarano and Milaković (2009) make of a very similar model to analyse herding behaviour in agent-based finance. Importantly, the results in section 1.3 reveal that even when recruitment is present, such that herding occurs over long horizons of time, it may be nevertheless very difficult to detect this process over short time horizons, as the system will be difficult to distinguish from a standard random walk.

References

Abrams Daniel M. and Steven H. Strogatz. Modelling the dynamics of language death. *Nature*, 424:900, 2003.

Alfarano Simone and Mishael Milaković, 2009. Network structure and n-dependence in agent-based herding models. *Journal of Economic Dynamics and Control*, 33:78–92.

Barde Sylvain, 2012. Back to the future: economic rationality and maximum entropy prediction. *University of Kent School of Economics Discussion Paper*, 12(02).

Cornwell T.J. and K.F. Evans. A simple maximum entropy deconvolution algorithm, 1985. *Astronomy and Astrophysics*, 143:77–83.

Foley Duncan K., 1994. A statistical equilibrium theory of markets. *Journal of Economic Theory*, 62:321–345.

Jaynes Edwin T. Information theory and statistical mechanics i, 1957a. The Physical Review, 106:620–630.

Jaynes Edwin T. Information theory and statistical mechanics ii, 1957b. *The Physical Review*, 108:171–190.

Kirman Alan, 1993. Ants rationality and recruitment. *Quarterly Journal of Economics*, 108:137–156.

Narayan Ramesh and Rajaram Nityananda, 1986. Maximum entropy image restoration in astronomy. *Annual review of astronomy and astrophysics*, 24:127–170.

Schelling Thomas C. Dynamic models of segregation, 1971. *Journal of Mathematical Sociology*, 1:143–186.

Schelling Thomas C. Models of segregation, 1969. *American Economic Review*, 59:488–493.

Shannon Claude E. A mathematical theory of communication, 1948. *The Bell System Technical Journal*, 27:379–423.

Shore John E. and Rodney Johnson. Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, 1980. *IEEE Transactions on Information Theory*, 26:26–37.

Skilling John and Stephen F. Gull. Bayesian maximum-entropy image reconstruction, 1991. *Spatial Statistics and Imaging*, 20:341–367.

Stauffer Dietrich, Xavier Castelló, Victor M. Eguíluz, and Maxi San Miguel, 2007. Microscopic Abrams-Strogatz model of language competition. *Physica A*, 374:835–842.

Toda Alexis A., 2010. Existence of a statistical equilibrium for an economy with endogenous offer sets. *Economic Theory*, 45:379–415.

Appendix

A. Time density of states in a stopped random walk

The diffusion model used in section 1.2 is a stopped one-dimensional random walk, with $T = \tau p_j(x_0)$ expected jumps of equally probably size $\pm\, 1/N$. The probability of having moved distance $\pm\, k/N$ for the first six steps is shown in Table 3. One can see that these are simply the relevant binomial coefficient divided by two to the power of the number of steps. As a result, the general probability of having moved distance k/N after T steps is given by:

$$p\left(d_T = \frac{k}{N}\right) = \begin{pmatrix} T \\ \frac{T+k}{2} \end{pmatrix} 2^{-T} \qquad \text{(A-1)}$$

Table 3. Diffusion from initial condition in a random walk model

	Distance traveled from x_0												
Steps	-6/N	-5/N	-4/N	-3/N	-2/N	-1/N	0	1/N	2/N	3/N	4/N	5/N	6/N
0							1						
1						1/2	0	1/2					
2					1/4	0	2/4	0	1/4				
3				1/8	0	3/8	0	3/8	0	1/8			
4			1/16	0	4/16	0	6/16	0	4/16	0	1/16		
5		1/32	0	5/32	0	10/32	0	10/32	0	5/32	0	1/32	
6	1/64	0	6/64	0	15/64	0	20/64	0	15/64	0	6/64	0	1/64

The proportion of the T steps spent at a given distance k/N, needed for the model (5) is then simply the average over the relevant column in Table 3. The major difference from the standard "Pascal triangle" visible in this table is that given the transition probabilities, even distances can only be reached with an even number of steps, and conversely, odd distances require an odd number of steps. As a result, in order to simplify the calculation of the average over the number of steps, a recurrence rule is developed that links every other entry in a column. This uses the two central recurrence rules for binomial coefficients:

$$\begin{pmatrix} n+1 \\ k \end{pmatrix} = \begin{pmatrix} n \\ k \end{pmatrix} + \begin{pmatrix} n \\ k-1 \end{pmatrix} \qquad \text{(A-2)}$$

$$\binom{n}{k+1} = \binom{n}{k}\frac{n-k}{k+1} \tag{A-3}$$

Combining the two and rearranging the indexes provides the following recurrence rule:

$$\binom{n+2}{k+1} = \binom{n}{k}\left(2 + \frac{n-k}{k+1} + \frac{k}{n-k+1}\right) \tag{A-4}$$

This rule can be used to link directly adjacent non-zero entries in a given column of Table 3. By using the formula recursively, one can eliminate the binomial coefficient from the right hand side, and express the binomial coefficient for any T and k as a product of terms generated in the same column for lower values of k, where $k = N\,|\,x - x_0|$ represents the absolute number of steps the system has traveled away from the initial condition x_0. This allows us to specify (A-1) as follows:

$$2^{-T}\binom{T}{\frac{T+k}{2}} = \begin{cases} 2^{-k}\displaystyle\prod_{i=0}^{i=\left\lfloor\frac{T-k}{2}\right\rfloor-1}\frac{1}{4}\left(2+\frac{i}{k+i+1}+\frac{k+i}{i+1}\right) & \text{if} \quad \left\lfloor\frac{T-k}{2}\right\rfloor \geq 1 \\ 2^{-k} & \text{if} \quad 1 > \left\lfloor\frac{T-k}{2}\right\rfloor \geq 0 \\ 0 & \text{if} \quad T-k<0 \end{cases} \tag{A-5}$$

The first right hand-side element forms the core of the expression. The third element simply states that the number of time steps T is also a strict upper bound on the distance than can be traveled in that time, while the second element states that the probability of being on this upper bound is given by a negative power of 2.[15]

In practical terms one starts by computing for each value of k a vector containing the argument in brackets for all values of $i \in \{0,1,2,\ldots,\lfloor(\tau p_j(x_0) - k/2)\rfloor - 1\}$, using the argument of the product term in (A-5). The cumulative product of this vector provides all the non-zero probabilities (A-1) in the k^{th} column in Table 3. The sum of these vector entries, divided by $\tau p_j(x_0) + 1$, then provides the required model:

15. These two expressions can be seen directly in Table 3: The top sides of the triangle are simply formed by increasing powers of 1/2. Above these, the distribution is not defined.

$$m(x,\tau)=\frac{\displaystyle\sum_{T=0}^{T=\tau p_j(x_0)} 2^{-T}\begin{pmatrix} T \\ \dfrac{T+N|x-x_0|}{2}\end{pmatrix}}{\tau p_j(x_0)+1} \tag{A-6}$$

B. Fokker-Planck derivation of the limit distribution of the ant model

Alfarano and Milaković (2009) show, using a Taylor expansion of the step operator formed by the Markov transition matrix of transition probabilities (3) that the following Fokker-Planck equation describes the evolution of the distribution over states:

$$\frac{\partial\mu(x)}{\partial\tau}=-\frac{\partial}{\partial x}A(x)\mu(x)+\frac{1}{2}\frac{\partial^2}{\partial x^2}B(x)\mu(x) \tag{A-7}$$

The drift term $A(x)$ of the equation corresponds to the expected jump size (7):

$$A(x)=N\left(p_{w\to b}-p_{b\to w}\right)=N\varepsilon(1-2x) \tag{A-8}$$

The diffusion term $B(x)$ corresponds to the probability of a jump $p_j(x)$ given by (4) and used to model the diffusion process away from the initial condition (5).

$$B(x)=p_{w\to b}+p_{b\to w}=\varepsilon+2(1-\delta)x(1-x) \tag{A-9}$$

The Fokker-Planck equation (A-7) is a second order differential equation with variable coefficients, and as shown by Alfarano and Milaković (2009), the general solution is of the following form, where c is a constant of integration that can serve to normalise the probability distribution:

$$\mu(x)=\frac{c}{B(x)}\exp\left(2\int_{x'=0}^{x}\frac{A(x')}{B(x')}dx'\right) \tag{A-10}$$

Replacing $\delta=2\varepsilon$ in (A-9) and assuming, as is done both in (Kirman, 1993; and Alfarano Milaković, 2009) that $\varepsilon\to 0$ and $N\to\infty$ in such a way that that $N\varepsilon$ remains constant, the diffusion term can be simplified to $B(x) = 2x(1 - x)$. Replacing the drift term $A(x)$ and diffusion term $B(x)$ in the general solution gives:

$$\mu(x)=\frac{c}{2x(1-x)}\exp\left(N\varepsilon\int\limits_{x'=0}^{x}\frac{(1-2x')}{x'(1-x')}dx'\right) \qquad \text{(A-11)}$$

The integral term is equal to ln $x(1-x)$, which leads to the limit distribution identified in Kirman (1993):

$$\mu(x)\propto x^{N\varepsilon-1}(1-x)^{N\varepsilon-1} \qquad \text{(A-12)}$$

ASYMMETRIC (S,s) PRICING:
IMPLICATIONS FOR MONETARY POLICY[1]

Zakaria Babutsidze
SKEMA Business School, Sophia Antipolis
OFCE

This paper presents a model of asymmetric (S,s) pricing. We investigate implications of such a behavior for the effectiveness of the monetary policy. We discuss two types of asymmetric responses to monetary interventions. One is the symmetry in the responses to positive and negative monetary shocks. The other is the variance in responses to monetary shocks during booms and recessions. The conclusion is that first type of asymmetry can be attributed to the asymmetry in adjustment bands, while the second kind of asymmetry is a result of firm heterogeneity, and asymmetry of (S,s) bands does not contribute to it.

Keywords: (S,s) Pricing, Monetary policy, Heterogeneity, Asymmetry.

Pricing behavior of individual firms has implications for the aggregate price and output movements. The propagation of money supply shocks crucially depends on pricing patterns. If firms in every moment in time charge the optimal price and there are no imperfections on financial markets, it is easy to show that money supply shocks have no real effects (Akerlof and Yellen, 1985). If financial markets are imperfect, for example there are information asymmetries (Greenwald and Stiglitz, 2001) or the credit market is characterized by financial accelerator (Bernanke *et al.,* 1999), money neutrality disappears. In this paper we abstract from the possibility of financial market imperfections and concentrate on the possibility

1. The author is grateful to Ricardo Caballero, Maurizio Iacopetta, Batlome Janjgava, Attila Rátfai, Vladimir Yankov, an anonymous referee and participants of several meeting in Bologna, Budapest and Maastricht for helpful comments and suggestions.

of individual price deviations from optimum. We believe that in real life prices are rarely at the optimum. There are two reasons for this. One is that there exists costs to price adjustment. The other one is that firms do not reconsider their prices as regularly as it would be necessary for keeping them at optimum permanently.

This paper does not go into the discussion of which of these arguments is more plausible. Instead we assume a specific type of firm pricing behavior which has been empirically well documented and try to understand the implications of this behavior for the effectiveness of monetary policy during equilibrium as well as during different phases of business cycle. More precisely we assess the effectiveness of monetary policy during booms and recessions.

We adopt the framework of (S,s) pricing (Caplin and Spulber, 1987) which introduces the inaction interval around the optimal price. As long as price is within the interval it is optimal for the seller not to adjust the price. In this type of models money has been found to be neutral (Caplin and Spulber, 1987). However, this finding is not robust to asymmetry of inaction bands above and below optimal price. Asymmetry creates some room for monetary policy. In asymmetric setup money is not neutral. We build on empirical finding pointing to the possibility of asymmetry in adjustment bands around the optimal price and analyze a simple model. We do not model neither the fine-grained micro behavior nor non-market interaction among firms. We simply assume asymmetry of price adjustment bands.[2] We also assume that each firm is

2. Although the present work does not concentrate on the derivation of the optimality of asymmetric bands, here we provide further possible explanations and a sketch of possible modeling technique. As we argued before, menu costs and adjustment costs are not exactly the same. So, adjustment costs can be different for movements of price in different directions. For example there are some psychological factors that can be at work making adjustment costs different (Greenslade and Parker, 2012). Then optimality of asymmetric bands can be derived from the usual monopolist profit maximization problem (Babutsidze, 2006). In principle, the asymmetric adjustment cost is not the only way to get asymmetric bands of adjustment. Similar results can be obtained by assuming the asymmetric profit function. Namely, profit function that is steeper before optimal price and flatter after it. This assumption makes not adjustment, but rather deviation costs asymmetric. To see this define deviation costs as

$$C_{dev} = \frac{1}{2}\frac{\partial^2 \pi}{\partial p^2}(p - p^*)^2.$$

Then if a profit function is flatter when $p > p^*$ for the same absolute value of deviation

$$C_{dev}^{p<p^*} > C_{dev}^{p>p^*}.$$

Thus, even with symmetric adjustment costs firm's pricing behavior will feature a longer right tail and a shorter left one.

hit by an idiosyncratic shock at any point in time that might push its price outside the inaction interval and induce it to adjust its price to the optimal level. Besides, government can conduct active monetary policy that would equally affect all the firms. Firms' responses to these policies are analyzed in order to assess the effectiveness of the monetary interventions.

Under the assumption of infinitesimally small idiosyncratic shocks, or alternatively very wide inaction bands, the model can be solved analytically. We derive a long-run density function of price level distributions in absence of monetary shocks. This is interpreted as equilibrium distribution. The effects of monetary policy in equilibrium can be also assessed analytically. However, there are two interesting departures from equilibrium that are worthy of analysis. One is a moderate size of adjustment bands. This is because wide adjustment bands imply excessively large adjustment costs that are not in line with empirical findings. The other departure is related to the cyclicality of the economy. Any external aggregate shock that hits the economy may knock it out of the equilibrium state. As price adjustments are not instantaneous, it takes a while until the system converges back to the ergodic price distribution. We try to assess the powers of monetary policy during this transitional dynamics.

These two exercises cannot be performed using analytic tools. This is where Agent Based Modeling (AMB) comes in handy. ABM is a flexible framework that does not require analytical tractability, which simplifies the task in the present case. It is a bottom-up modeling framework, which means that modeler can specify behavior of individual agents at the microscopic level and explore its implications for macroscopic outcomes. Merits of ABM are extensively discussed in few of the articles in this special issue (e.g. Fagiolo and Roventini, 2012; Napoletano and Gaffard, 2012). Using computational tools we set up an ABM equivalent of the model and explore the behavior of the system in simulated environments. Using ABM methodology we analyze the effects of the monetary intervention in presence of non-trivial idiosyncratic shocks and during booms and recessions.

There are two major findings. One is that in presence of sufficiently large shocks the model is able to reproduce significant asymmetry in output's reaction to positive and negative macroeconomic shocks. Asymmetry of adjustment bands plays the crucial role in this. The second major finding is that model is characterized by asymmetry in responses to similar shocks across different phases of business cycle. However, the asymmetry on micro level is not necessary for this. The difference in responses to similar shocks across booms and recessions seems to be the result of simple existence of inaction interval, rather than its asymmetry.

The rest of the work is organized as follows. Section 1 reviews related strands of literature. Section 2 lays out the model. Section 3 presents the results. Section 4 concludes.

1. Related literature

Our work is closely related to two large strands of literature. One strand is concerned with the pricing behavior of firms and implications of this behavior for macroeconomics. The second one discusses the empirical findings about asymmetries on micro and macro levels.

From wide range of models concerning firms' pricing behavior most closely related to the work presented in this paper are sticky price models. During the last few decades sticky price models have proved to be of great importance. The empirical findings illustrate that prices are not flexible enough to always be at the optimum. The evidence of price stickiness is found in many markets. For example, Stigler and Kindahl (1970) and Carlton (1986) find evidence of price stickiness for various industrial goods, Cecchetti (1986) for magazine prices.[3]

Sticky price models can be divided into two parts: in one class of models firms follow a time-dependent policy of price adjustment; in the other one they follow a state dependent policy. Time-dependent pricing models assume that a firm's decisions of revising and modifying the existing price are constrained by some time limits. For example, in Fisher (1977) and Taylor (1980) models of staggered pricing firms are allowed to set their prices every other period. In Calvo (1983) the information about the changes in market conjuncture arrives randomly in time. So, decisions about

3. More recent documentation of price stickiness is due to Levy *et al.* (1997), Blinder *et al* (1998), Wolman (2000), etc.

the price changes also follow a random process. These models imply forward-looking price setting and result into the phenomenon called New Keynesian Philips Curve, which differs from the classical Philips curve that is constructed using backward-looking pricing. Time dependent price setting models feature money non-neutrality and some empirical support has been found for them (Gali and Gertler, 1999; Fabiani *et al.,* 2006). However, they have also raised some criticism because they do not match wide range of macroeconomic regularities (see for example Fuhrer and Moore, 1995; Mankiw, 2001) and prompted researchers to propose alternative models (such as one due Mankiw and Reis, 2002).

State-dependent pricing models are more intuitive. The baseline logic here is that firms change prices depending on the state of economy. In this setup firms may change the price every period or leave it unchanged for a number of periods. The best representation of state-dependence is (S,s) pricing (Caplin and Spulber, 1987; Caplin and Leahy, 1991). The (S,s) rule was first introduced by Arrow et al. (1951) for inventory management purposes. Later, Barro (1977) and Sheshinski and Weiss (1977; 1983) also applied it to pricing models. In these models, due to the existence of adjustment costs, the zone of inaction is created around the optimal price for the firm. As long as the price is inside of the band, it is optimal not to adjust it. When the price crosses any of the inaction bands the adjustment to optimal price is observed. More recently (S,s) pricing models have been used to gain insights into the effects of monopolistic competition (see for example Caplin and Leahy, 1997). They have also been used successfully in multi-sector general equilibrium models (e.g. Damjanovic and Nolan, 2007).

All these pricing models allow for agent heterogeneity despite the fact that the strategies and the incentives of all of them are usually assumed to be identical. Heterogeneity comes with the different prices of the producers that are due to the frictions to the price adjustment. If there were no frictions, all the prices would coincide and the behavior of the aggregate variables would be the same as the individual ones (scale adjusted).

Recent years have seen a development of mixed models, or so called generalized state dependent pricing models (Devereux and Siu, 2007; Woodford, 2009; Costain and Nakov, 2011). In this paper we present a model with asymmetric (S,s) bands which belongs to this later class of general models. We take the asymmetry of inaction bands as given, based on a well-documented empirical findings (e.g. Tobin, 1972; Ball and Mankiw, 1994).

The literature on empirics about asymmetries can be divided in two parts. One part documents asymmetries on micro level, the other—on macro level. The fact that prices do not change very often is a well documented fact (Klenow and Malin, 2011; Greenslade and Parker, 2012). The present work is based on a more fine-grained finding which is that individual prices are more rigid downwards than upwards, but if they decline, they decline by a higher magnitude relative to price increases. This means that firms' adjustment policies are asymmetric on microeconomic level. There are two types of asymmetries observed on aggregate level also. One is that the aggregate output has low and high response regimes to the monetary policy (Lo and Piger, 2005; Peersman and Smets, 2001). Namely, the output responds to a somewhat lesser extent to positive monetary shocks during the recession than during the normal periods and even lesser than during the booms. Second, the output response is smaller in magnitude when we have positive money supply shocks rather than when we have negative ones (Cover, 1992).

The asymmetry of microeconomic adjustment policies has been documented long ago. In the 70's, economists were talking about the downward rigidity of prices (Tobin, 1972). More resent research also shows the overwhelming evidence on more frequent price increases than decreases. For example, Borenstein *et al.* (1997) find the microeconomic asymmetry on gasoline and agricultural products' markets, Jackson (1997) finds it on bank deposits. To this Chen *et al.* (2004) add the documentation of the asymmetry in price changes in American supermarket chains.

The asymmetry in the frequency and the magnitude of adjustment is better documented for European countries. Loupias and Ricart (2004) investigate the pricing behavior of over 1600 French manufacturing firms and find that positive price changes are more frequent than negative ones. They also find that the magnitude of up- and downward price changes are different: they report an average of 3% for price upgrades in contrast with an average of -5% for price downgrades. Their findings are supported by another study of French manufacturing firms' behavior by Baudry *et al.*

(2004), who find no evidence of nominal downward rigidity but support the asymmetry in magnitude of changes, although less pronounced (+4% versus -5%). Similar relation between frequency of price change and its magnitude has also been found recently for the UK (Bunn and Ellis, 2012).

A similar picture emerges in other European countries. In Belgium, Aucremanne and Dhyne (2005) find no differences in the frequency, but in the magnitude of price changes: +6.8% versus -8.7%. For Spain, Alvarez and Hernando (2004) find that the ratio of price increases to price decreases is 1.6. With regard to the asymmetry in the magnitude of price changes they report +8.2% for price increases versus 10.3% of price decreases. For Portugal, Dias *et al.* (2004) find no difference in magnitude of changes but a huge contrast in the frequency of price changes in different directions; they report the ration of positive to negative price changes equal to 2.34.[4]

Lach and Tsiddon (1992) also find the asymmetry in magnitudes of price deviations for Israel. They examine disaggregated price data of foodstuffs in Israel during 1978-1984. Their main conclusion is that the asymmetry is more pronounced during high inflation periods, more precisely when the annual inflation goes above 130%.

Of course, these findings are not left without attention. Ball and Mankiw (1994) incorporate the difference in frequency into their model. They do this by introducing the positive drift in inflation process justifying this with some kind of Harrod-Balassa-Samuelson effect due to the faster economic integration and the development of countries. This introduces the asymmetry in price distribution. Although Ball and Mankiw's (1994) model is able to feature more frequent price upgrades than downgrades, still the magnitudes of changes on the firms level are equal. Thus, anticipated positive drift in inflation explains only half of the story.

Tsiddon (1991) presents a simple menu cost model for high inflationary environment. He introduces the costs for adjustment that are proportional to the deviation from the optimal price and

4. Further evidence on asymmetry for all EU15 countries is provided by Lunnemann and Matha (2004).

derives the optimal pricing policy for the representative firm. The author distinguishes between price stickiness and downward rigidity and concludes that the model features the latter. The model exhibits an asymmetry in the following way. According to the optimal pricing policy, during the low inflation periods firms adjust their prices more frequently than during the high inflation periods. This is due to the fact that high inflation increases the uncertainty in future optimal price movements and the optimality is achieved by waiting. A similar result is obtained by Hansen (1999) who derives the dependence of the "first passage time" function on the degree of uncertainty. So, in a sense, Tsiddon's (1991) model features the difference in the magnitudes of the price adjustment as well as the difference in the frequency of price adjustment.[5]

Although the inflation trend assumed in these models is an intuitive device for introducing asymmetry, as it aggravates the effect of a positive shock and mitigates the effect of a negative one, it is not well matched with the empirical findings. For example, Peltzman (2000) shows that asymmetry is very pronounced in the United States in the period 1982-1996, when the positive drift in inflation was measured to be less than 2%. DeLong and Summers (1988) find an asymmetry during the Great Depression period when the price trend was deflationary. All this points to the fact that trend inflation can not explain even the different frequency of price up- and downgrades. Some other factors seem to be in work.

The overwhelming majority of sticky price models (e.g. Tsiddon, 1991; Ball and Mankiw, 1994) take the inaction bands lying on an equal distance from the optimal price. If we take the adjustment cost to be a menu cost[6] type, the symmetry is justified: there is no reason why the menu costs can be different for changing the prices in different directions. But the problem is that the adjustment cost is a much wider notion than the menu cost. There are many other factors that can be regarded as the ingredients of the cost of changing price. For example, the psychological factor as seeing the

5. There are also the examples of the other kinds of asymmetry in price adjustment derived in different setups. See for example Danziger (1988) where asymmetry is due to the discounting of future profits in inflationary environment. There every price spends most of the time being below the optimal one.
6. See for example Mankiw (1985).

product's price raising with large jumps can result in loss of consumers and decreasing profits. This can further propagate to firm's large negative jump in purchases of inputs offending the suppliers. Large discrete downward jumps are better justified: this will probably result in "stealing" the buyers from competitors and also hoping to bargain a good discount with a supplier on a larger order due to the increased output. Recent empirical support for this view is due Greenslade and Parker (2012) who analyze large sample of UK firms.

The importance of these considerations is outlined in Bowman (2002). The author presents a model of sticky prices without any menu costs. In this model for firms it is optimal not to change prices in response to nominal shocks because doing so increases their profits by expanding the customer base. Then the non-neutrality of the money is obtained without any kind of menu costs. Some other kinds of cost seem to deter firms from adjusting prices.

Also, as documented by Kwapil *et al.* (2005), firm's decisions about price upgrades and downgrades depend on different factors. Research on Austrian manufacturing firms shows that changes in wage and intermediate goods' costs are two of the most important factors for price increases, while changes in competitors' prices and technological improvements are the main driving factors for price reductions. Furthermore, Loupias and Ricart (2004) conclude that menu costs are absolutely not important for price changes of manufacturing products. Then, from this point of view, there is absolutely no reason why the costs of price changes in different directions have to be the same.

The literature on asymmetries on macro level concentrates on two major asymmetries. The first one is the asymmetry in responses of output to the expansionary and contractionary shocks of the same size. This is a well documented empirical finding for developed economies. For example, Cover (1992) exploits the quarterly data spanning 1951:1-1987:4 and finds a very high degree of asymmetry. He uses three model specifications for the identification of the asymmetry: the one proposed by Barro and Rush (1980), modified specification of Mishkin (1982) and his own. Asymmetry is pronounced in all three models. In Barro-Rush model 73% of a negative monetary shock is passed to output, while the same indicator for positive shocks is only 1% and it is not significant. In the modified Mishkin model the same indicator is 66% versus 6% (the latter again not significant). In Cover's original model 96% of negative monetary shock is passed to output, while, although not significant, the passthrough from positive shocks has the wrong sign. From these considerations one can conclude that positive monetary shocks do not have any effect on output and they basically pass to prices while negative shocks are passed to output to a larger extent. The more recent study of Ravn and Sola (2004) confirms the basic conclusions of Cover (1992) about the existence of asymmetry, but in their case the asymmetry is less pronounced.[7]

The second type of macro asymmetry is in reaction of output to monetary shocks during different phases of business cycle. Lo and Piger (2005) employ a Markov regime-switching model to investigate the asymmetry in output movements after monetary shocks to different directions. Their finding is that there is a very well pronounced time variation in output responses that can be explained by the time varying transition probability model. Basically, they find that the variation can be explained by inclusion in the model of a simple dummy variable indicating whether the economy is in a recession or in a boom. This confirms the authors' hypothesis that output reaction has two regimes: "low response" and "high response." In particular, policy actions taken during recessions seem to have larger effects on output than those taken during expansions.

Similar two-regime character of output responses has been found for number of economies. For example, Garcia and Schaller (2002) found asymmetry in US output response a bit earlier than Lo and Piger (2005). Peersman and Smets (2001) find the same type of asymmetry for the whole set of European countries. Furthermore, Kaufmann (2002) and Kwapil *et al.* (2005) document two regimes of output reaction for Austria.

All in all there is an asymmetry on macro as well as on micro levels. However, the link between micro- and macroeconomic asymmetries is complicated. In fact, microeconomic asymmetry in price adjustment can totally cancel out at the aggregate level, or macroeconomic asymmetry can be introduced by aggregation of

---

7. The asymmetry to positive and negative monetary shock responses is also found in other parts of the world. Karras (1996) finds asymmetry in 18 European countries. Chu and Ratti (1997) find asymmetry in the Japanese economy.

the firms with absolutely symmetric microeconomic pricing properties. A simple model presented by Caballero (1992) is an excellent demonstration of this point. Caballero (1992) demonstrates the the link between micro and macro asymmetries has to be analyzed very carefully. There is no distinct link identified between these two phenomena. The motivation of the present work is to contribute to this line of research with aspiration of gaining further insight into the functionality of monetary policy. In the next section we provide the baseline model of the present paper.

2. The model

2.1. Setup of the model

We model Chamberlinian monopolistic competition following Dixit and Stiglitz (1977). The economy consists of a continuum of monopolistically competitive firms indexed on [0;1] interval that produce close (but not perfect) substitutes. This form is chosen because in a perfect competition setup a positive deviation from the optimal price results in large losses due to the loss of the entire market share. This is because, in the case of perfect competition, the profit function of the firm is not continuous in own price: it has a discrete jump immediately after the optimal price (Akerlof and Yellen, 1985). This makes competitive environment useless for the purposes of this paper.

Consider a monopolistic firm that faces downward sloping demand of a form

$$Y = \left(\frac{P}{\overline{P}}\right)^{-\eta} \frac{M}{\overline{P}}, \tag{1}$$

where P is the own price of firm's product, M is the money supply per firm, $\overline{P}$ is the aggregate price. The positivity of monopolistic markup gives the condition $\eta > 1$. The firm operates at a constant real marginal costs $C = \beta Y^{\alpha}$, where β can be interpreted as the real wage per unit of effort (in equilibrium it is constant), α is the inverse of productivity parameter. Then, the monopolistic profit maximization problem is

$$\max_{P} \pi = PY - PC \tag{2}$$

with respect to the demand on Y. Assuming symmetry, that the prices of all the goods are equal, the problem results in $P = \overline{P}$ and gives $P = GM$, where G is constant and is equal to:

$$G = \left(\frac{\eta - 1}{\beta(\eta\alpha - 1)} \right)^{\frac{1}{1-\alpha}}.[8]$$

Notice that in this (no adjustment costs) setup the output of a single firm, and as a consequence of the whole economy, is constant at a value G.

Taking the natural logarithms of the price-money supply relationship, denoting the logarithms by lower case letters, we get

$$p^* = g + m. \tag{3}$$

Then, it is apparent that $dp^* = dm$. Thus, the idiosyncratic, mean-zero shocks in money supply would call for no aggregate price changes.

Let's introduce a variable x that is the deviation of firm's actual price from its optimal one, defined as $x = p - p^*$. Note that unlike other papers (e.g. Hansen, 1999) the negative value of x means that the actual price is lower and the positive value—that the actual price is higher than the desired price. We make this assumption because of simpler tractability of results of the density function of x derived in the next sub-section.

We also assume that there is a fixed cost of adjustment that is not necessarily equal for up- and downgrading the price. And there is a cost of being apart from the optimal price. Following Hansen (1999) we assume that this cost is incurred at every moment when $p \neq p^*$ and can be measured as accumulated flow costs. Note that due to the concavity of the profit function, the cost of being at non-optimum is the second order. Then an entrepreneur makes a decision by comparing the two costs. As long as the deviation cost is sufficiently lower prices do not change. This behavior creates the zone of inaction that is *not necessarily symmetric* around the optimal price.

---

8. Note that the solution puts stricter requirement on η. It requires $\eta > 1/\alpha$ for the positivity of G.

2.2. Deriving the long-run density

In this framework we can derive the long-run density of price deviations. Define *f(x)* as the long-run, time-invariant density function of price deviations. This function can also be interpreted as the likelihood of having a price deviation equal to *x* at any particular moment. For the derivation of the density function we assume that Brownian motion in money supply has very simple properties: it is a mean zero process and at every instant *dt* it can change *x* by *dx* with equal probabilities going up and down. This means that if we are now at *x* after one period *(dt)* we will be at $x + dx$ with probability 0.5 and at $x - dx$ with probability 0.5. Then,

$$f(x) = \frac{1}{2} f(x + dx) + \frac{1}{2} f(x - dx), \tag{4}$$

as being today at *x* means being either at $x - dx$ or at $x + dx$ a moment ago. This is a very convenient property. We can rewrite (4) as

$$\big(f(x+dx) - f(x)\big) - \big(f(x) - f(x-dx)\big) = 0.$$

Then, division by *dx* gives

$$\frac{f(x+dx) - f(x)}{dx} - \frac{f(x) - f(x-dx)}{dx} = 0. \tag{5}$$

Notice that as $dx \to 0$ two parts of left hand side of expression (5) converge to derivatives of *f(x)* and then whole left hand side is something like the change in the derivative from point $x + dx$ to point x[9]. Then the whole expression (5) is equivalent to the second derivative of *f(x)* being zero

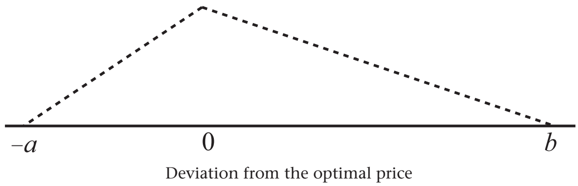
$$\frac{d^2 f(x)}{dx^2} = 0\,.$$

Now, as *f(x)* is a density function, we know that

$$\int_{-a}^{b} f(x)dx = 1, \tag{6}$$

9. In real life this would mean to assume that inaction bands on both sides of the optimal price are wide in comparison to the average size of an idiosyncratic shock. This assumption is necessary for deriving analytical results and is relaxed in coming sections when we employ ABM techniques.

Figure 1. The long-run density funtion



Now, as *f(x)* is a density function, we know that

$$\int_{-a}^{b} f(x)dx = 1, \tag{6}$$

where –*a* and *b* are optimal bands of price adjustment. Thus, price deviation *(x)* is distributed between –*a* and *b*. We also have two boundary conditions $f(-a) = f(b) = 0$, by assumption that prices are adjusted immediately as they reach any of the boundaries, thus none of them, in principle, are reached. Then we can split the integral (6) into two parts

$$\int_{-a}^{0} f(x)dx + \int_{0}^{b} f(x)dx = 1. \tag{7}$$

From the second derivative of $f(x)$ being zero we know that both of these parts are linear. From the boundary conditions we know their crossing points with x axis are $x = -a$ and $x = b$. Also, note that $f(b)$ has to reach maximum at $x = 0$, because has the highest probability equal to

$$\frac{1}{2}(f(-dx)+f(b-dx))+\frac{1}{2}(f(dx)+f(-a+dx))=f(0) \tag{8}$$

This is the probability of being either at $-dx$ or at $b-dx$ and getting a positive shock plus the probability of being either at dx or at $-a+dx$ and getting a negative shock. Then, two strait lines have to cross at $x = 0$, otherwise the density function will not be continuous.

All these conditions together imply that $f(x)$ has a triangular shape with the base $a + b$ and the height $2/(a + b)$ (and it reaches maximum at $x = 0$). This gives us the solution to the problem

$$f(x)=\begin{cases}\dfrac{2}{a+b}\left(1+\dfrac{x}{a}\right) & if\, x<0\\[2ex] \dfrac{2}{a+b}\left(1-\dfrac{x}{b}\right) & if\, x\geq 0\end{cases} \tag{9}$$

Thus, the resulting density function looks like the one shown on Figure 1.[10]

The shape of the resulting density function has an interesting implication. The figure is drawn for the case when $a < b$ which seems to be a realistic scenario given the empirical findings summarized in section 1 of this paper. This implies difference in an intensive margin (Klenow and Kryvtsov 2008). From the figure we can infer that near the upgrading band (near $-a$) there are relatively more firms than near the downgrading band (near b). This demonstrates the difference in an extensive margin. This result emphasizes the obscurity of the link between micro- and macro-asymmetry: although price downgrades are higher in magnitude there are fewer firms who want to reduce their prices as a result of a shock. Consequently, it is not obvious that the positive shock in price deviations[11] will induce the aggregate price level to reduce with higher magnitude than the rise caused by the negative shock of the same magnitude. In fact, there is a chance that these two factors completely cancel out each other and we get the same result as Caballero (1992).

The long-run density (9) has few interesting characteristics. The share of firms that hold price under their optimal price is $a/(a + b)$. Consequently the share of firms holding the price over the optimal one is $b/(a + b)$. In fact this average price deviation can be calculated as

$$\bar{x} = \frac{a}{a+b}\int_{-a}^{0}\frac{2}{a+b}\left(1+\frac{x}{a}\right)xdx + \frac{b}{a+b}\int_{0}^{b}\frac{2}{a+b}\left(1-\frac{x}{b}\right)xdx, \qquad (10)$$

which results into

$$\bar{x} = \frac{b^3 - a^3}{3(a+b)^2}. \qquad (11)$$

This is interesting as it implies that in case of asymmetry (when $a < b$) the average price deviation will be positive. In other words

10. Notice that the original assumption of discretization of a continuous process, mainly that x can go to only two states, either $x + dx$ or $x - dx$ is not crucial for the form of the density function. If one assumes many different type of idiosyncratic shock distribution it is easy to show that the same shape results. A crucial assumption for the shape is that the distribution is symmetric and centered around zero, which is maintained throughout the whole paper.

11. As shown in the next section a positive shock in price deviations is equivalent to a negative monetary shock.

an "average" firm will be holding the price above the optimal level. This happens without assuming any inflationary expectations.

3. Response to monetary policy

For the analysis of the responses of the system to monetary policy we setup an ABM equivalent to the model described in previous section. There are two reasons for this. One is that we want to depart from the unrealistic assumption of infinitesimally small idiosyncratic shocks. Recall that this was a necessary assumption for derivation of the long-run density. If idiosyncratic shocks are not of a negligible size compared to the adjustment bands, the price deviation density will depart from the one described by equation (4). In this case larger share of firms will hold prices close (or equal to) the optimal price.

The second reason for using ABM is that we want to analyze the implications of the model for the effectiveness of monetary policy during turbulent periods. We want to check how system responds to monetary shocks during booms and recessions. Recall one of the empirical findings regarding marco asymmetry has been that expansionary monetary policy is more effective during recessions than during booms. We want to check the implications of our model in this respect.

3.1. Methodology

In this sub-section we provide essential details of the simulation methodology. Of course, we can not work with the continuum of firms any longer. As we work with price deviations we have to transform the results in terms of price and output responses. Let x_0 be an initial price deviation for a single firm $x_0 = p_0 - p_0^*$. Then money supply shock of a magnitude ε is also an optimal price shock of the same magnitude $p_1^* = p_0^* + \varepsilon$. This gives $x_1 = p_1 - p_1^*$. From these identities we get $x_1 = p_1 - p_0^* - \varepsilon$. Then it is apparent that a positive shock in money supply transforms into a negative shock in price deviations and vice versa. Intuitively, the immediate rise in optimal price for the firm means that its relative price has lowered. Finally, one can express the evolution of the price of a single firm as

$$p_1 - p_0 = \epsilon + x_1 - x_0 \tag{12}$$

We track the evolution of every single price in the economy. Then, the evolution of the aggregate price is derived by simply averaging all the prices in the economy.

For output changes, we proceed with demand functions. Taking natural logarithms of the original demand function and totally differentiating gives

$$dy = \eta(d\overline{p} - dp) + (dm - d\overline{p}) \tag{13}$$

From here it is obvious that the output changes for every single firm depend on the parameter η. But on the aggregate level, note that by definition

$$\sum d\overline{p} = \sum dp \text{, as } nd\overline{p} = n\frac{\sum dp}{n}.$$

So, the first summand in (13) disappears on the aggregate level and we are left with

$$dm = d\overline{p} + d\overline{y} \tag{14}$$

where $\overline{y}$ is a log of aggregate output. So, on the aggregate level the role of price elasticity of demand disappears. Then, to simplify calculations, for aggregate output we proceed with the rearrangement of (14), as we know dm and also $d\overline{p}$.

Results of the model depend on the size of she policy and idiosyncratic shock compared to firms' inaction band. Therefore, we fix the size of the inaction band and calibrate monetary policy and idiosyncratic shocks in corresponding units. We normalize the size of the inaction band $a + b = 100$. In this case an idiosyncratic shock of size w, can be interpreted as the shock of w% of the inaction band. The same is true for monetary policy—its size will be measured as a corresponding percent of an inaction band. Then, the asymmetry of the pricing policy can be described by parameter a. If $a = 50$, there is no asymmetry in firms' pricing strategy. If $a < 50$ firms tolerate larger price deviations above the optimal price compared to the deviations below it. If $a > 50$ situation is reversed.

We assume the idiosyncratic shocks are normally distributed with zero mean and variance that is measured in units comparable to the size of the inaction band. Variance being equal to w, means shock are drawn from $N(0,\sigma)$, which corresponds to the shock variance being equal to the σ^2 % of the inaction band. If w is small enough, we have demonstrated that the time invariant price deviation distribution density is given by (9). However, when shock variance increases the long run distribution departs from the one derived analytically. Larger mass of firm's will be adjusting each period to optimal price and as a consequence larger mass will be concentrate at $x = 0$. In order to permit the system to converge to the time invariant distribution before starting a policy experiment we initialize the system with a uniform distribution of x over the interval $[-a;b]$ and let the system run without any aggregate shock for 3000 periods[12]. Once the system has settled to the time invariant distribution we conduct a policy experiment—we introduce a monetary shock of certain size and analyze the system's response to it.

We study the economy populated by 1000 firms. For reporting each result we conduct 150 Monte-Carlo simulation and report the average values across all 150 runs. In all cases standard deviations are extremely small, therefore they are not reported on graphs below.

3.2. Results

A major contribution of the paper to the literature is that we can discuss the implications of the extent of the asymmetry of the adjustment bands. Recall that we have normalized $a + b = 100$. Then parameter a completely characterizes the adjustment band asymmetry. Asymmetry of inaction interval (a) is one of the major parameters in our investigation. This is because the results of a recent study by Álvarez *et al.* (2007) that has assembled the evidence from european countries suggests variation in levels of asymmetry across countries.

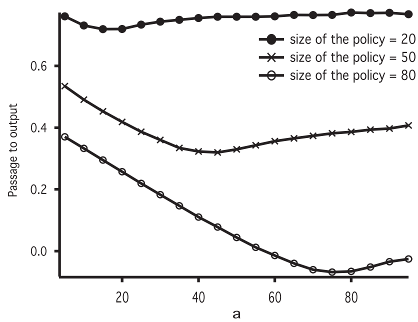
Figure 2 presents the results of agent-based model that demonstrates the effect of the asymmetry on the effectiveness of the monetary policy. On the bottom axis the parameter a is plotted, while on the vertical axes we have plotted the share of the monetary shock passed to output. The value 0.3 on the vertical axis should be interpreted as 30% of the shock being passed to output while 70% being absorbed by the prices. The graph is reproduced

12. Numerous simulations show that in case of sufficiently low variance long-run equilibrium is indeed the one given by (9). As a consequence, the results reported in this paper are not dependent on initial conditions unlike, for example, Caplin and Spulber (1987) where initial distribution is crucial for basic results of the model.

by setting the variance of the idiosyncratic shock to unity, which is a low enough level for the time invariant price deviation distribution to be well described by the analytical one presented by equation (9). We are usually interested by the left half of the graph because that half implies $a < b$ which is a realistic case based on the empirical findings reviewed in this paper.

Figure 2 presents three sequences for three different sizes of the expansionary monetary policy: for 20, 50 and 80% of the inaction band. As one can clearly see the asymmetry plays virtually no role if the magnitude of the monetary policy is small. With increasing size of the monetary intervention role of asymmetry becomes prominent. For instance with policy size of 80 greater asymmetry (going to the left on the graph) implies higher efficiently of the policy. This is intuitive as larger asymmetry leaves fewer firms at the right edge of the x distribution which will adjust prices when policy is implemented. Fewer firms adjusting prices induces larger share of monetary shock being passed to output.

Figure 2. The effect of the adjustment band asymmetry

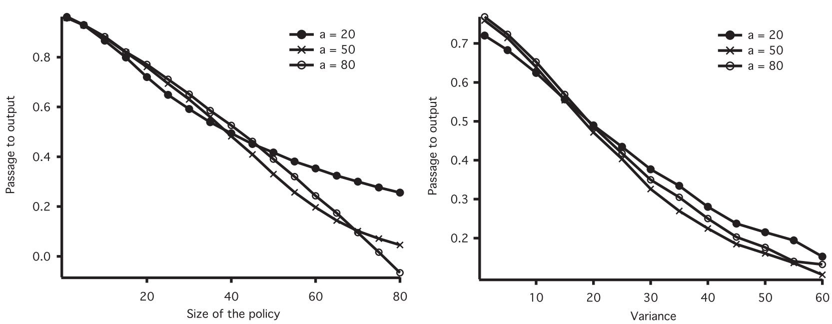


Another important effect that has been demonstrated by figure is the impact of the policy size on its passage to output. This effect is better demonstrated by the left panel of figure 3. Similar to the figure , in this figure $\sigma^2 = 1$. The model predicts that the size of the monetary intervention negatively affects its efficiency. The logic behind this result is that larger monetary shock knocks more firms out of the inaction bands, induces them to adjust to the optimal price and as a result drives up the inflation instead of affecting real economy.

Although in all the figures we present in this paper we are discussing expansionary monetary policy, we can in fact draw conclusions also about the effects of the contractionary policies. This is due to the symmetry of the results. If $a = 50$ we do not have asymmetry in inaction bands and the size of the results of positive and negative monetary interventions are equal. However, in case of asymmetry for any given a we can construct a scenario to derive the corresponding results for the contractionary policy. Consider an arbitrary a. We know that $b = 100 - a$. Therefore, the contractionary monetary policy for inaction band asymmetry being described by a, is exactly equal to the result generated by $100 - a$. This means that in figure 2 the effect of the positive and negative monetary policy are given by mirroring at $a = 50$. For example, when $a = 20$, 25% of positive monetary shock of size 80 is passes to output, as documented by the graph. However, contractionary monetary policy is sterile, which is seen by observing the passage to output being equal to zero at $a = 80$ (which is a mirror to $a = 20$).

The implications of the model in this regard are easily seen on the left panel of figure 3. In this figure we plot three series, each corresponding to different values of asymmetry. Two of them correspond to $a = 20$ and $a = 80$ which are the mirror cases comparable to each other. The discrepancy between these two series implies differential response to positive and negative monetary shocks. As we know that reality calls for $a < 50$, and figure presents results for the expansionary monetary policy, it is intuitive to view results of $a = 20$ as response to expansionary monetary policy and that of $a = 80$ as contractionary monetary policy.

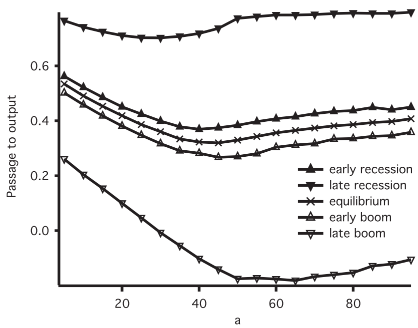
Figure 3. The effects of the policy size and idiosyncratic shock variance



As one can clearly see the two series depart from each other as policy size grows. The results predict that if there exists an asymmetry in adjustment bands (of sort that $a < b$), for large enough monetary intervention, positive expansionary monetary policies are more effective than contractionary monetary policies. This is in line with the empirical findings surveyed above.

Another result concerns the analytical long-run price deviation distribution derived above and its implications. Recall that for the derivation of function (9) we had to assume infinitesimally small idiosyncratic shocks, which in case of numerical simulations means $\sigma^2 \to 0$. As σ is measured as the constant share of the inaction band size, this effectively means infinitely large inaction interval. This is, clearly not realistic. Agent-based simulations present us a chance to explore the effect of relaxing this assumption and exploring the effects on the monetary policy.

Figure 4. Output response during boom and recession



The right panel of the figure 3 presents results where we vary the value of σ^2. As we have anticipated in the text above, larger variance would imply larger mass of firms leaving adjustment bands and reseting themselves to the mode of the distribution at $x = 0$. This would effectively mean that at any point in time greater number of firms holding optimal prices and monetary policy being less effective. As results presented in figure 3 show, this is indeed the case: for any size of asymmetry effectiveness of monetary policy is strictly decreasing in idiosyncratic shock variance. This result stresses the importance of the size of the inaction bands when taking the decision on the size of the monetary policy.

3.2.1. Monetary policy during booms and recessions

Here we present results of our model regarding the asymmetric response to aggregate monetary shocks during different phases of business cycle. The current model is a kind of hybrid of sticky and flexible price models. Everything depends on the distribution of price deviations and the direction of the monetary shock. For example, if economy is in a boom, that is, it has been hit with several positive shocks, the distribution of price deviations shifts to the left border of inaction interval. And any further positive monetary shock induces a large number of firms to raise their prices. The model gets closer to flexible price models and the output response is dampened. But this is only for positive monetary shocks. If, in this situation, the economy is hit by a negative monetary shock the distribution will shift to the right and basically no firm will adjust prices. Then, the model gets closer to sticky price models and the whole shock is passed to the output. So, the regime of output responses crucially depends on the direction of the aggregate shock.

Figure 4 presents five series. One of them, termed "equilibrium" is the series generated the same way as all the series up to now. The other four series represent responses to expansionary monetary policy during the different phases of a business cycle. Cycles in our computational environment are generated artificially by shocking the economy in several consecutive times. More precisely, early boom and early recession is generated by introducing policy of size +1 and -1 respectively, while late boom and recession are generated by introducing shocks of the same size for 20 consecutive time periods. After we bring the system to the state of boom or recession we exercise expansionary monetary policy and calculate the response that is presented on the figure.

The results are close to linear and conform to our conjectures. Expansionary monetary policy is becoming increasingly ineffective as we progress further into the boom and it becomes increasingly effective as we go deeper into the recession. It is worth mentioning that this statement is valid only in the case of positive monetary shocks. For negative ones, the situation is the mirror image. In case of contractionary policy, it is absorbed by prices in recessions but passed to the output in booms. But, the point is that this particular kind of heterogeneity of agents is able to produce some type of asymmetry. Stemming from the theoretical considerations above, these results can be derived from any (S,s) pricing model. The asymmetry of the bands is not required for this result. It is purely due to the shifts of the price deviation density to one of the edges of the distribution. So, asymmetry on the micro level is not the cause of the aggregate output having two regime property, but rather this is due to (S,s) pricing behavior itself. Thus, this kind of aggregate asymmetry is the direct consequence off heterogeneity of agents, no matter whether their micro policies are symmetric or asymmetric.

4. Conclusion

Individual prices change rarely, and there is a staggering in the adjustment since the price changes across the firms differ in time. This behavior is due to some costs involved in the price adjustment process: costs of gathering information about the market conjuncture, costs of loosing the market share, etc. So, the adjustment cost is a wider notion than "menu cost"; the latter is one of the components of the former. Due to the fact that some ingredients of price adjustment costs are asymmetric for price changes in different directions, the adjustment costs, as a whole, are also different for price upgrades and downgrades.

In the current paper we presented the model where individual firms follow asymmetric (S,s) pricing behavior. This is due to the asymmetry in the adjustment costs mentioned in the previous paragraph. We investigate few important questions such as asymmetry in responses to expansionary and contrationaty monetary policies and variance of the effectiveness of the policy during different phases of the business cycle. We also investigate the role of the asymmetry in adjustment bands in these processes.

The basic results were derived by numerically simulating the model. However, for the small idiosyncratic shocks the time invariant price deviation distribution had been analytically derived. This distribution does not depend on the initial conditions of the model. One more specific character of the current paper is that, unlike the most similar papers, we did not use simple binomial random walk for the description of shock process. Rather we used more elaborate shock process that allows for the variance in the size of the idiosyncratic shocks. This is important as it highlights the importance of firms located in the interior of the adjustment bands. This contrasts to the models with binomial shocks (e.g. Caballero, 1992), that but emphasis on firms located on margins of the inaction interval. We have also explored the effects of the changing variance in shock process.

We have explored at the implications of the asymmetric (S,s) pricing behavior of firms for two kinds of stylized facts about the asymmetry in the aggregate output dynamics. The first is the asymmetric response of output to positive and negative monetary shocks. Here the finding is that in the case of sufficiently high shocks, the model is able to produce significant asymmetry on the aggregate level between responses. The second type of asymmetry is that the aggregate output has low and high response regimes with respect to monetary shocks, depending on whether the economy is in boom or in recession. Although the model is able to produce this kind of effect for positive shocks, the main conclusion is that this is not due to the asymmetry on the micro level. Instead, firm heterogeneity itself creates the asymmetry on aggregate level.

References

Akerlof G, Yellen J, 1985. "A Near-Rational Model of the Business Cycle with Wage and Price Inertia". *Quarterly Journal of Economics* 100: 823-838.

Álvarez L, Hernando I, 2004. "Price Setting Behavior in Spain: Stylised Facts Using Consumer Price Micro Data". *European Central Bank Working Paper* 416.

Álvarez L, Dhyne E, Hoeberichts M, Kwapil C, Le Bihan H, Lunnemann P, Martins F, Sabbatini R, Stahl H, Vermuelen P, Vilmunen J, 2007. "Sticky Prices in the Euro Area: A Summary of New Micro-Evidence". *Journal of European Economic Association* 4: 575-584.

Arrow K, Harris T, Marschak J, 1951. "Optimal Inventory Policy". *Econometrica* 19: 205-272.

Aucremanne L, Dhyne E., 2005. "Time-Dependent Versus State-Dependent Pricing: A Panel Data Approach to the Determinants of a Belgian Consumer Price Changes". *European Central Bank Working Paper* 462.

Babutsidze Z, 2006. (S,s) "Pricing: Does the Heterogeneity Wipe Out the Asymmetry on Micro Level?", *UNU-MERIT Working Paper* 033.

Ball L, Mankiw G, 1994. "Asymmetric Price Adjustment and Economic Fluctuations". *The Economic Journal* 104: 247-261.

Barro R, 1977. "Unanticipated Money Growth and Unemployment in the United States". *American Economic Review* 67: 101-115.

Barro R, Rush M., 1980. "Unanticipated money and Economic Activity". In: Fischer S(Ed), *Rational Expectations and Economic Policy*. University of Chicago Press: Chicago.

Baudry L, Le Bihan H, Sevestre P, Tarrieu S, 2004. "Price Rigidity: Evidence from the French CPI Micro-Data". *European Central Bank Working Paper* 384.

Bernanke B, Gertler M, Gilchrist S, 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework". In *Handbook of Macroeconomics*, J. Taylor and M. Woodford (eds). Elsevier Science: Amsterdam.

Blinder A, Canetti E, Lebow D, Rudd J, 1998. *Asking About Prices: A New Approach to Understanding the Price Stickiness*. Russel Sage Foundation: New York.

Borenstein S, Cameron A, Gilbert R, 1997. "Do Gasoline Prices respond Asymmetrically to Crude Oil Price Changes". *Quarterly Journal of Economics* 112: 305-339.

Bowman D., 2002. "Sticky Prices, No Menu Costs". *International Finance Discussion Paper at Board of Governors of Federal Reserve System* 743.

Bunn P, Ellis C, 2012. "How Do Individual UK Producers Behave?". *The Economic Journal* 122: f16-f34.

Caballero R, 1992. "A Fallacy of Composition". *American Economic Review* 82: 1279-1292.

Calvo G, 1983. "Staggered Pricing in a Utility Maximizing Framework". *Journal of Monetary Economics* 12: 383-398.

Caplin A, Leahy J, 1991. "State Dependent Pricing and the Dynamics of Money and Output". *Quarterly Journal of Economics* 106: 683-708.

Caplin A, Leahy J, 1997. "Aggregation and Optimization with State-Dependent Pricing". *Econometrica* 65: 601-625.

Caplin A, Spulber D, 1987. "Menu Costs and Neutrality of Money". *Quarterly Journal of Economics* 102: 703-725.

Carlton D, 1986. "The Rigidity of Prices". *American Economic Review* 76: 637-658.

Cecchetti S, 1986. "The Frequency of Price Adjustment: A Study of Newsstand Prices of Magazines". *Journal of Econometrics* 31: 255-274.

Chu J, Ratti R, 1997. "Effects of Unanticipated Monetary Policy on Aggregate Japanese Output: The Role of Positive and Negative Shocks". *Canadian Journal of Economics* 30: 723-741.

Costain J, Nakov A, 2011. "Price Adjustments in a General Model of State-Dependent Pricing". *Journal of Money, Credit and Banking* 43: 385-406.

Cover J, 1992. "Asymmetric Effects of Positive and Negative Monetary-Supply Shocks". *Quarterly Journal of Economics* 107: 1261-1282.

Damjanovic V, Nolan C, 2007. "Aggregation and Optimization with State-Dependent Pricing: A Comment". *Econometrica* 74: 565-573.

Danziger L, 1988. "Costs of Price Adjustment and the Welfare Economics of Inflation and Disinflation". *American Economic Review* 78: 633-646.

DeLong B, Summers L, 1988. "How Does Macroeconomic Policy Affect Output". *Brookings Papers on Economic Activity* 2*:* 433-494.

Devereux M, Siu H, 2007. "State Dependent Pricing and Business Cycle Asymmetries". *International Economic Review* 48: 281-310.

Dias M, Dias D, Neves P, 2004. « Stylised Features of Price Setting Behavior in Portugal: 1992-2001". *European Central Bank working paper* 332.

Dixit A, Stiglitz J, 1977. "Monopolistic Competition and Optimal Product Diversity". *American Economic Review* 67: 297-308.

Fabiani S, Druant M, Hernando I, Kwapil C, Landau B, Loupias C, Martins F, Matha T, Sabbatini R, Stahl H, Stokman A, 2006. "What Firms' Surveys Tell Us about Price-Setting Behavior in the Euro Area". *International Journal of Central Banking* 2: 3-47.

Fagiolo G, Roventini A, 2012. "Macroeconomic Policy in DSGE and Agent-Based Models". *Revue de l'OFCE-Débats et politiques,* this issue.

Fuhrer C, Moore G, 1995. "Forward-Looking Behavior and the Stability of a Conventional Monetary Policy Rule". *Journal of Money, Credit and Banking* 27*:* 1060-70.

Gali J, Gertler M, 1999. "Inflation Dynamics: A Structural Econometric Analysis". *Journal of Macroeconomics* 44: 195-222.

Garcia R, Schaller H, 2002. "Are the Effects of Interest Rate Changes Asymmetric?". *Economic Inquiry* 40: 102-119.

Greenwald B, Stiglitz J, 1993. "Financial Market Imperfections and Business Cycles". *Quarterly Journal of Economics* 108: 77-114.

Greenslade J, Parker M, 2012. "New Insights into Price-Setting Behaviour in the UK: Introduction and Survey Results". *The Economic Journal* 122:1-15.

Hansen P, 1999. "Frequent Price Changes Under Menu Costs". *Journal of Economic Dynamics and Control* 23: 1065-1076.

Jackson W, 1997. "Market Structure and the Speed of Price Adjustment: Evidence of Non-Monotonicity". *Review of Industrial Organization 12:* 37-57.

Karras G, 1996. "Are the Output Effects of Monetary Policy Asymmetric? Evidence from a Sample of European Countries". *Oxford Bulletin of Economics and Statistics* 58: 267-278.

Kaufmann S, 2002. "Is there an Asymmetric Effect of Monetary Policy over Time? A Bayesian Analysis using Austrian Data". *Empirical Economics* 27: 277-297.

Klenow P, Kryvtsov O, 2008. "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?". *Quarterly Journal of Economics* 123: 863-904.

Klenow P, Malin B, 2011. "Microeconomic Evidence on Price-Setting". In *Handbook of Monetary Economics 3A*, B. Friedman and M. Woodford (eds). Elsevier, 231-284.

Kwapil C, Baumgartner J, Scharler J, 2005. "The Price-Setting Behavior of Austrian Firms: Some Survey Evidence". *European Central Bank Working Paper* 464.

Lach S, Tsiddon D, 1992. "The Behavior of Prices and Inflation: An Empirical Analysis of Disaggregated Price Data". *Journal of Political Economy* 100: 349-388.

Levy D, Bergen M, Dutta S, Venable R., 1997. "The Magnitude of Menu Costs: Direct Evidence from Large US Supermarket Chains". *Quarterly Journal of Economics* 112: 791-825.

Lo M, Piger J., 2005. "Is the Response of Output to Monetary Policy Asymmetric? Evidence from a Regime-Switching Coefficients Model". *Journal of Money Credit and Banking* 37: 865-886.

Loupias C, Ricart R, 2004. "Price Setting in France: New Evidence from Survey Data". *European Central Bank Working Paper* 423.

Lunnemann P, Matha T, 2004. "How Persistent is Disaggregate Inflation? An Analysis Across EU15 Countries and HICP Sub-Indices". *European Central Bank Working Paper* 415.

Mankiw G, 1985. "Small Menu Costs and Large Business Cycles: A Macroeconomic Model of Monopoly". *Quarterly Journal of Economics* 100: 529-537.

Mankiw G, 2001. "The Inexorable and Mysterious Tradeoff between Inflation and Unemployment". *Economic Journal* 111:45-61.

Mankiw G, Reis R, 2002. "Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve". *Quarterly Journal of Economics* 117:1295-1328.

Mishkin F, 1982. "Does Anticipated Policy Matter? An Econometric Investigation". *Journal of Political Economy* 40: 22-51.

Napoletano M, Gaffard J-L, 2012. "Introduction to Special Issue on New Advances in Agent-Based Modeling: Economic Analysis and Policy". *Revue de l'OFCE-Débats et politiques,* this issue.

Peersman G, Smets F, 2001. "Are the Effects of Monetary Policy in the Euro Area Greater in Recessions than in Booms?". *European Central Bank Working Paper* 52.

Peltzman S, 2000. "Prices Rise Faster than They Fall". *Journal of Political Economy* 108: 466-502.

Ravn M, Sola M, 2004. "Asymmetric Effects of Monetary Policy in the United States". *Federal Reserve Bank of St. Louis Review* 86: 41-60.

Sheshinski E, Weiss Y, 1977. "Inflation and Cost of Price Adjustment". *Review of Economic Studies* 44: 287-303.

Sheshinski E, Weiss Y, 1983. "Optimum Pricing Policy Under Stochastic Inflation". *Review of Economic Studies* 50: 513-529.

Stigler G, Kindahl J, 1970. "The Behavior of Industrial Prices". *Columbia University Press: New York*.

Taylor J, 1980. "Aggregate Dynamics and Staggered Contracts". *Journal of Political Economy* 88: 1-24.

Tobin J, 1972. "Inflation and Unemployment". *American Economic Review* 62: 1-18.

Tsiddon D, 1991. "On the Stubbornness of Sticky Prices". *International Economic Review* 32: 69-75.

Wolman A, 2000. "The Frequency and Costs of Individual Price Adjustment". *Federal Reserve Bank of Richmond Economic Quarterly* 86 n°4.

Woodford M, 2009. "Information-Constrained State-Dependent Pricing". *Journal of Monetary Economics* 56: 100-124.

MACROPRUDENTIAL POLICIES IN AN AGENT-BASED ARTIFICIAL ECONOMY

Silvano Cincotti and Marco Raberto
DIME-DOGE.I, Università di Genova
Andrea Teglio
Universitat Jaume I, Campus del Riu Sec

Basel III is a recently-agreed regulatory standard for bank capital adequacy with focus on the macroprudential dimension of banking regulation, *i.e.*, the system-wide implications of banks' lending and risk. An important Basel III provision is to reduce procyclicality of present banking regulation and promote countercyclical capital buffers for banks. The Eurace agent-based macroeconomic model and simulator has been recently showed to be able to reproduce a credit-fueled boom-bust dynamics where excessive bank leverages, while benefitting in the short term, have destabilizing effects in the medium-long term. In this paper we employ the Eurace model to test regulatory policies providing time varying capital requirements for banks, based on mechanisms that enforce banks to build up or release capital buffers, according to the overall conditions of the economy. As conditioning variables for these dynamic policies, both the unemployment rate and the aggregate credit growth have been considered. Results show that the dynamic regulation of capital requirements is generally more successful than fixed tight capital requirements in stabilizing the economy and improving the macroeconomic performance.

Keywords: Basel III, Macroprudential Regulation, Agent-Based Models and Simulation.

The recent economic and financial crisis has cast serious doubt on the idea of efficient self-regulating financial and credit markets, and consequently the need for a more effective regulation of these markets unquestionably has arisen. As a response to the crisis, a new global regulatory standard has been proposed under the name of Basel III[1], with the objective to improve the resilience of the banking system.

The rational behind Basel III regulation, which is founded on the same three pillars[2] characterizing its previous version Basel II, is that one of the main reasons why the economic crisis became so severe was that the banking sectors of many countries had built up excessive on—and off-balance sheet leverage. The erosion of the level and quality of the capital base determined that the banking system was not able to absorb systemic trading and credit losses nor could it cope with the large off-balance sheet exposures. The crisis was further amplified by a procyclical deleveraging process. The weaknesses in the banking sector were rapidly transmitted to the rest of the financial system and the real economy, resulting in a massive contraction of liquidity and credit availability (for more details see BIS, 2011).

Some previous works by the authors reproduced these economic mechanisms by means of the agent-based model and simulator Eurace (see Raberto *et al.* 2012; Teglio *et al.* 2012). In particular, Raberto *et al.* (2012) shows that excessive bank leverages can drive economies into severe recession in the medium-long run. The pressure on wages and labor costs during credit-fueled economic booms, in conjunction with the speed of growth of credit-money, causes a rise of inflation, that in turn can determine higher interest rates. Excessively indebted firms may be unable to fulfill their financial commitments with the cash proceedings of their revenues, and may be obliged to take new loans to pay interests on their debt, therefore entering in a Ponzi scheme. However, the deterioration of firms creditworthiness causes a further rise of interest rates due to the widening of the risk spread on policy rates. This, in turn, affects the balance sheet of highly indebted firms, which may become soon insolvent. Debt write-offs reduce banks's equity and their lending capacity, thus causing a widespread credit rationing and a forced deleveraging of the corporate sector that may trigger a possible wave of bankruptcies of even good but illiquid firms. A credit-fueled economic boom may thus turn out in a depression.

1. For details and documents on Basel III, please visit the Bank for International Settlements (BIS) website at: http://www.bis.org/bcbs/basel3.htm
2. The three pillars are: minimum capital requirements, supervisory review process and market discipline. http://www.bis.org/publ/bcbs107.htm

Given this economic dynamics, emerged in Raberto *et al.* (2012), the aim of this paper is to understand if some of the post-crisis measures proposed by the Basel Committee can have a positive impact in our model.

According to Borio (2011), the institutional response after the crisis has taken two forms. Policymakers have been strengthening the systemic (or macro-prudential) orientation of regulatory and supervisory frameworks, and they have begun to question the premise that financial stability can be secured without a more active support of macroeconomic policies. The established pre-crisis policy framework was focused on the stability of individual institutions (micro-prudential orientation) while the reforms introduced in Basel III provide a macro-prudential approach to regulation and supervision that has a system-wide focus, with the goal to limit the risk of episodes of financial distress with serious consequences for the real economy ("systemic risk").

Claiming that one of the most destabilizing elements of the crisis has been the procyclical amplification of financial shocks throughout the banking system, the new regulatory framework introduces some measures in order to make banks more resilient to such procyclical dynamics, like encouraging banks to create countercyclical capital buffers in order to accumulate capital when the economy is strong and use it when the economic conditions are bad. The problem, pointed out again by the Basel Committee in BIS (2011) and also emerged from the computational experiments of Raberto *et al.* (2012), is that losses incurred in the banking sector can be extremely large when a downturn is preceded by a period of excess credit growth. These losses can destabilize the banking sector, creating a credit crunch that contributes to a downturn in the real economy that then feeds back to the banking sector again.

According to Drehmann *et al.* (2010), the main target of countercyclical capital standards is to encourage banks to build up buffers in good times that can be drawn down in bad ones. In this paper, we design two endogenous adaptive policy rules for the Eurace agent-based model; the two rules set capital requirements in the spirit of encouraging banks to build up capital buffers in good times. Details about the implementation of these policy rules, as identifying bad and good times and choosing the conditioning variables which could guide the build-up and release of capital, are discussed in section 3.1.

The issue of pro-cyclicality in regulatory policy has been widely discussed in the literature of the last 20 years. Blum and Hellwig (1995) already observed that a "rigid link between bank equity and bank lending may act as an automatic amplifier for economic fluctuations, inducing banks to lend more when times are good and to lend less when times are bad, thus reinforcing underlying shocks". They propose a simple stylized macroeconomic model where banks must satisfy a minimum-reserve requirement and a capital adequacy requirement, in order to study the effects of demand disturbance for different levels of capital requirements. Their conclusion, later extended by Cecchetti and Li (2008) in a more complete economic framework, is that capital requirements have a significant macroeconomic impact. Heid (2007) presents a model with a representative bank which invests in riskless bonds and loans, subject to regulatory constraint, explaining the cyclical effects of capital requirements also in the case of banks which always hold a positive capital buffer. In Raberto *et al.* (2012) and Teglio *et al.* (2012), we addressed similar issues using an agent-based methodology, confirming and extending the relevance of the macroeconomic implications of capital requirements. We are both able to reproduce the endogenous amplification of economic fluctuations and to observe how these fluctuation are affected by different levels of capital requirements.

In the last ten years, and markedly after the 2007 crisis, the discussion on bank regulation and pro-cyclicality significantly increased, incorporating new concepts as "systemic risk". Acharya (2009) shows that capital adequacy requirements fail to mitigate systemic risk, using a multi-period general equilibrium model with many agents and markets, inspired to the Allen and Gale (2000) one-period model of bubbles and crisis. In order to assess the cyclaclity of capital requirements, several macroeconometric models have been proposed and estimated on data of different countries. Andersen (2011) and Antão and Lacerda (2011) simulate the IRB (internal rating based) approach of Basel II using Norwegian and Portuguese data respectively, both confirming the cyclicality of capital requirements and comparing the new regulatory framework with the previous one of Basel I.[3] More recently, in particular after the appearance of Basel III, a discussion about the utility and the correct implementation of macroprudential regulation of the banking sector has arisen. Repullo and Saurina (2011) present a critical assessment of the countercyclical capital buffer in Basel III, evaluating the "conditioning variables" suggested for taking buffer decisions (see Drehmann *et al.*, 2010). Their conclusion is that the choice of the credit-to-GDP gap as the "common reference point" for taking buffer decisions can be misleading because its correlation with the GDP growth is generally negative, and this contradicts the necessity of building buffers of resources in good times that can draw down when conditions deteriorate. They also claim that credit growth "appears to be a much better common reference point for the countercyclical capital buffer". As it will be shown in section 4, our results confirm the efficacy of credit growth as conditioning variable.

In this paper we study the macroeconomic implications of macroprudential policy regulations using an agent-based approach. With respect to the previous literature, mainly consisting in general equilibrium models or macroeconometric models, the Eurace agent-based model and simulator is a much more complex environment where all the economic adjustments are endogenous and produced by the interaction of many heterogeneous agents acting in different markets. The completeness of the Eurace framework is particularly important in this study, where it is necessary to consider the interplay and spillover between the production, the financial as well as the credit sector of the economy. In the last decade, several agent-based economic models have been developed in order to focus on the relation between the credit and financial factors and the real economy, see e.g. Delli Gatti *et al.* (2005, 2009); Raberto *et al.* (2008a); Dosi *et al.* (2010); Chiarella and Di Guilmi (2011). However, the novelty and the advantage of the Eurace framework is the simultaneous presence of the most important economic agents interacting in many different markets. This feature allows for an endogenous and realistic representation of the whole economic system in an evolving dynamic

3. In Basel II, with respect to Basel I, the capital charges depend on asset quality, based on public or internal ratings, rather than on asset type.

setting, which has no antecedents in the history of economic modelling.

The paper is divided as follows. Section 1 presents an overview of the general structural features of the Eurace model with the related references. Section 2 reports a detailed description of a new model for capital goods' demand within Eurace, while a throughout description of the implementation of the Basel II capital requirements rule and of the new countercyclical policies is reported in section 3. Computational results are presented and discussed in Section 4, while Section 5 draws our concluding remarks.

1. An overview of the Eurace model

Eurace is a model and simulator of an artificial economy which belongs to the class of agent-based computational models (see Tesfatsion and Judd (2006) for a review). The agent-based approach to economics addresses the modelling of economic systems as complex adaptive systems, *i.e.*, systems made by many self-interested interacting units (economic agents here) that may change their behavior in order to adapt to the changing (economic) environment and to the change of other units' behavior. The main distinguishing features of an agent-based artificial economy with respect to the mainstream dynamic stochastic general equilibrium (DSGE) modelling can be summarized as follows: out-of-equilibrium dynamics versus market equilibrium, decentralized markets with pairwise bargaining and price dispersion versus centralized markets and the law of one price, adaptive expectations with myopic behavior versus rational expectations and infinite foresight, endogenous shocks and business cycles versus exogenous shocks.

The Eurace artificial economy has been constantly evolving since the start in 2006 of the Eurace project within a EU-funded research grant under the sixth framework programme. Eurace is a fully-specified agent-based model of a complete economy that includes different types of agents and integrates different types of markets. Agents include households which act as consumers, workers and financial investors, consumption goods producers as well as capital goods producers, banks, a government and a central bank. Agents interact in different types of markets, namely markets for consumption goods and capital goods, a labor market, a credit market and a financial market for stocks and government bonds. Except for the financial market, all markets are characterized by decentralized exchange with price setting behavior on the supply side. Agents' decision processes are characterized by bounded rationality and limited information gathering and computational capabilities; thus, agents' behavior follows adaptive rules derived from the management literature about firms and banks, and from experimental and behavioral economics of consumers and financial investors.

In the following, we outline the structural features of the Eurace economy with respect to the agents considered, the types of real and financial assets owned and exchanged by agents as well as the related payment commitments over time. Finally, table 2 presents the balance sheets entries of the Eurace agents. The balance-sheet variables can be regarded as the state variables of any agent and, along with wages, interests and prices, are endogenously determined within the system. In particular, wages and consumption goods prices are heterogeneous and fixed by any CGP according to labor market conditions and costs, interest rates are fixed by banks and are heterogeneous as well, because they depend on the creditworthiness of the borrower as well as on the central bank rate;[4] in both cases see Raberto *et al.* (2012); Cincotti *et al.* (2012), for further details. The only exogenous variables are the price of energy (or raw materials) which is considered constant and, accordingly, the price of capital goods which is constant as well, being a fixed mark-up on the energy price. The number of the different types of agents is also fixed.

1.1. Types of agents

- Households (Hous) (indexed by *h*)
- Consumption goods producers (CGP) (also named *firms* and indexed by *f*)
- Investment goods producer (KGP)
- Banks (B) (indexed by *b*)

4. The central bank policy rate is endogenously determined *via* a Taylor rule

- Foreign Sector (FgnS)
- Government (Gov)
- Central Bank (CB)

1.2. Types of Assets

1.2.1. Real assets

- homogeneous consumption goods (C Goods): (q);
- homogeneous capital goods (K Goods): (K);
- homogeneous energy or raw materials;
- homogeneous labor units: (L).

1.2.2. Financial assets

- deposits (of households, firms and the KGP) at banks;
- deposits (of banks and the government) at the central bank;
- loans from banks to firms;
- loans from the central bank to banks;
- equity shares (for firms, banks and the capital good producer);
- government bonds.

Table 1. Interaction matrix of the Eurace model

	Hous	CGP	KGP	B	FgnS	Gov	CB
Hous	**equity shares** **gov. bonds**	C GOODS **equity shares** *dividends*		**deposits**		*transfers* **gov. bonds** *coupons*	
CGP	LABOR		K GOODS	**loans** **deposits**			
KGP				**deposits**	RAW MAT.		
B		*interests* *principal*					**loans** **deposits**
FgnS							
Gov	*taxes*	*taxes*	*taxes*	*taxes*			*seignorage*
CB				*interests*		**bonds**	
				principal		*coupons*	

The matrix should be read as follows: row agents are the ones demanding or receiving real assets (denoted in small caps), financial assets (denoted in bolded style) and their related monetary payment commitments over time (denoted in italics); column agents are the ones supplying the corresponding real assets, financial assets and their related monetary flows.

Table 2. Balance sheets of the agents in the Eurace economy

Agent	Assets	Liabilities
Household	Liquidity Equity shares Gov bonds	Equity
CGP	Capital goods (K^f) Inventories Liquidity	Loans $(D^f = \sum_b \ell_b^f)$ Equity (E^f)
KGP	Liquidity	Equity
Bank	Loans $(\sum_f \ell_b^f)$ Liquidity	Deposits (Liquidity of Hous, CGP and the KGP) Standing facility with the CB Equity
Government	Liquidity	Bonds Equity
Central Bank	Standing facility with Banks Gov bonds Liquidity	Outstanding fiat money Deposits (Liquidity of Banks and the Gov) Equity

Agents' behaviors are thoroughly described in our previous works, see e.g. Raberto *et al.* (2012); Teglio *et al.* (2010b); Cincotti *et al.* (2012) about decision making hypotheses in real (consumption goods and labor) markets as well as in credit markets. In particular, consumption goods producers as well as banks are short-term profit maximizers that fix prices (the price of consumption goods and the lending rate) based on a fixed mark-up on their costs (wages and cost of capital for CGPs and the central bank policy rate for banks). CGPs make their production decisions according to standard results from inventory theory (Hillier and Lieberman, 1986). Households' saving-consumption decision is modelled according to the theory of buffer-stock saving behaviour (Carroll, 2001; Deaton, 1992), which states that households consumption depends on a precautionary saving motive, determined by a target level of wealth to income ratio. Households can invest their savings in the asset market, by buying and selling equity shares or government bonds. Households' portfolio allocation is modeled according to a preference structure designed to take into account the psychological findings emerged in the framework of behavioral finance and in particular of prospect theory (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). Households' behavior in the financial market has been thoroughly described in Raberto *et al.* (2008b); Teglio *et al.* (2009). It is worth noting, however, that only the equity shares of CGPs are exchanged among households in the stock market and only CGPs are allowed to issue new equity shares to be sold to households (see the related cells in the Interaction Matrix in Table 1). Equity shares of KGP and banks are equally distributed among households and can not be traded; put in another way, profits of the KGP and of banks are equaly shared and distributed to households. Finally, the balance sheet approach employed in agent-based modeling, as outlined in Table 2, is described in details in Teglio *et al.* (2010a) and in Cincotti *et al.* (2010), where computational experiments about the use of quantitative easing policies have been reported.

The computational experiments reported in this study have been realized within an enriched version of the Eurace model. In particular, beside the new regulatory policies designed and investigated in this study, the Eurace model and simulator employed here is characterized by a new model for capital goods demand and a better founded model for the estimation of bankruptcy probability for firms, as described in the following sections.

2. Capital goods investments

2.1. Demand of capital goods

Physical capital K is employed with labor L by consumption goods producers (CGPs) to produce an amount of consumption goods q according to a Cobb-Douglas technology, as follows

$$q = \gamma L^{\alpha} K^{\beta}, \tag{1}$$

where γ is positive and constant returns to scale are considered, *i.e.* α and β are positive constants with the constraint $\alpha + \beta = 1$. CGPs need new capital goods both to replace capital depreciation and to expand the production capacity. In the previous version of the Eurace model, see e.g. Raberto *et al.* (2012), CGPs computed the desired amount of physical capital, given the planned production quantity, by means of a static optimization method based on the Cobb-Douglas technology and isoquants. In the new model presented here, the demand for capital goods is based on the net present value (NPV) approach, which, according to a recent empirical survey (Graham and Harvey, 2001, 2002), is, along with the internal rate of return (IRR),[5] the most popular method used by managers to evaluate investments. The new investments model should be therefore considered better founded on the realistic behavior of economic actors.

CGPs compute their demand for new capital goods, henceforth investments or I, at the beginning of any production month.[6] The amount I of new capital goods will then be delivered and will be ready to use during the next production month. The first step in computing the demand of new capital goods is the estimation[7] by every CGP of the expected demand $\hat{q}^d$ of consumption goods it should face during the month. Based on this estimate and on the inventories amount ν, the CGP computes the production needs $\overline{q}$ as $\overline{q} = \hat{q}^d - \nu$, if $\nu < \hat{q}^d$, or $\overline{q} = 0$ otherwise. Given the production goal $\overline{q}$, its present capital endowment K and the Coob-Douglas technology, any CGP computes the necessary workforce L^d in order to meet the production goal as:

$$L^d = \left(\frac{\overline{q}}{\gamma K^{\beta}}\right)^{\frac{1}{\alpha}}. \tag{2}$$

Considering the present workforce L, the CGP opens new vacancies, if $L^d > L$, or decides layoffs if $L^d < L$.

It is worth noting that, if the present workforce L and the present endowment of physical capital K are not sufficient to meet the production goal, *i.e.*, $\overline{q} > \gamma L^{\alpha} K^{\beta}$, increasing the workforce to L^d is the only feasible way to meet the production goal during the present month. The reason is that the new demanded physical capital would be available to the CGP only from the next month. Nevertheless, the expected demand $\hat{q}^d$ is still used to determine the demand of new capital goods. The rationale is that the CGP has to check if the increase of the stock of physical capital to meet the

5. It can be shown, however, that the IRR is generally equivalent to the NPV, except for some special cases.
6. It is worth remembering that 1 month, *i.e.*, 20 business days, is the basic time span for the production process in Eurace and that the starting production day of the month is fixed for each firm but can be different (asynchronous) among different firms.
7. The estimate is based on a linear interpolation of the sales made during a given number of previous months.

production goal, given the present workforce, is profitable in the reasonable hypothesis that q^d would be a lower bound for demand in the future.

If $\overline{q} > \gamma L^{\alpha} K^{\beta}$, the CGP computes the desired endowment of physical capital K^d as follows:

$$K^d = \left(\frac{\overline{q}}{\gamma L^{\alpha}}\right)^{\frac{1}{\beta}}. \tag{3}$$

The difference $K^d - K$ is a reference point for investments. It is worth noting, however, that according to the realistic hypothesis of imperfect capital markets in Eurace and according to the difference between internal and external financing for producers, we stipulate, in line with Fazzari *et al.* (2008), that cash flows are a bound for nominal investments and we consider last month CGP revenues as a proxy for cash flows.

The CGP computes the NPV using a grid of investment values I in the range between 0 and I_{max}, where I_{max} is given by the minimum between $K^d - K$ and last months revenues divided by the price of capital p_K, set by the capital good producer.[8]

The NPV is computed considering the present cost of investments, *i.e.*, $p_K I$ and the discounted values of the future cash flow given by the augmented productive capacity.

In particular, we consider

$$NPV(I) = -p_K I + \sum_{m=1}^{m^*} \frac{\hat{p}_C^m \Delta q_C^m}{(1+\frac{r}{12})^m}, \tag{4}$$

where $\hat{p}_c^m$ is the expected consumption goods price level[9] at month m, r is the weighted average yearly cost of capital for the CGP, and Δq_c^m is the additional amount of monthly production given by the capital investment I, after properly taking into account the depreciation given by ξ, *i.e.*,

$$\Delta q_C^m = \gamma L^{\alpha} (K + (1-\xi)^{(m-1)} I)^{\beta} - \gamma L^{\alpha} K^{\beta}. \tag{5}$$

8. See e.g. Raberto *et al.* (2012) for some details about the stylized modeling of capital goods producers in Eurace.

9. As common in agent-based computational models, the approach is the one of adaptive expectations. Expected inflation is then computed based on past inflation, measured in a given moving time window.

The end value m^* of the sum is given by the number of months when investment is still higher than a given threshold after considering the monthly depreciation rate ξ. The CGP demands the amount I of capital goods that maximizes the $NPV(I)$. If $NPV(I) < 0$. for any I in the grid, no investment should be done. Finally, it is worth remembering that CGPs are never rationed in the demand for capital goods; however, they may be obliged to reduce the investment schedule because rationed in the credit market.

2.2. Supply of capital goods

Capital goods are offered with infinite supply by a single capital goods producer (KGP), which follows a build to order production approach. No inventories and financing needs are then considered for the KGP. Energy and raw materials are the only factor of production and are assumed to be imported from abroad. The price of energy and raw materials is exogenously given and set to a constant value. The price of capital goods p_K is a fixed mark-up on energy prices and therefore is also constant in this setting. Profits of the capital good producer are distributed in equal shares among all households. Put differently, it is assumed that all households own equal shares of the capital goods producer and that shares are not traded in any market. Therefore, the amount payed by consumption goods producers for capital goods is partially (the part related to mark-up) channeled back into the economy, while the part related to energy costs leaves the Eurace economy.

3. Credit financing of investments

Consumption goods producers face the liquidity needs necessary to finance the production and investment plans as well as the scheduled financial payments, *i.e.* interests, debt installments, taxes and dividends. They decide between internal and external capital according to the pecking-order theory (Myers and Majluf, 1984), which states that, because of information asymmetries present in both credit and equity markets, firms prefer to meet their financial payments first by using their internal liquidity, then by means of new debt, if liquidity is not sufficient, then by issuing new equity if rationed in the credit market.

It is worth noting that, if the producer is unable to meet its financial payments, it goes into bankruptcy. Two types of bankruptcies are considered, *i.e.*, insolvency and illiquidity bankruptcy. The first type is when CGP's equity goes negative. The second type is when the CGP is unable to pay its financial commitments but still owns a positive equity. The significative difference between the two types of bankruptcies is that, in case of insolvency bankruptcy, CGP's debt is restructured, *i.e.* the debt is reduced to a target fraction of CGP's total assets and the corresponding loans in the portfolios of lending banks are written-off (as consequently banks' equity is written-off as well). When a CGP goes into bankruptcy, either of the insolvency or of the illiquidity type, it fires all its employees, stopping production for a period necessary to raise new equity capital in the stock market in order to further strengthen its capital base and increase its liquidity.

In the following, we will discuss in detail how the supply of credit is determined within the banking sector.

3.1. Credit supply

Credit to consumption goods producers is provided by banks, who are supposed to be short-term profit maximizers. Households' deposits are not rewarded and interests due to loans from the central bank are the only costs for banks. Given the positive difference between the lending rate to the corporate sector and the central bank rate, banks can always increase their profits by increasing their lending. Banks are then supposed to always fulfil loan requests if regulatory constraints are satisfied, *i.e.* if capital requirements are met. In particular, following the Basel II capital adequacy rules, we stipulate that a minimum percentage of the risk-weighted loan portfolio should be held by the bank in the form of equity capital as a buffer for possible loan write-offs and equity losses. We denote this minimum percentage as κ and we call it capital requirement. The risk weight ω^{λ} of any loan λ depends on the borrower credit worthiness after the loan and is measured by using its balance sheet entries. Let us suppose that a bank b, with equity E^b and risk weighted loan portfolio W^b, receives a loan request amount λ^f from CGP f, characterized by debt D^f and equity E^f, then the bank is allowed to lend to the CGP an amount $\ell^{bf} \leq \lambda^f$ given by:

$$\ell^{bf} = \begin{cases} \lambda^f & if \quad E^b \geq \kappa(W^b + \omega^\lambda \lambda^f), \\ \dfrac{E^b - \kappa W^b}{\kappa \omega^\lambda} & if \quad \kappa W^b < E^b < \kappa(W^b + \omega^\lambda \lambda^f), \\ 0 & if \quad E^b \leq \kappa W^b. \end{cases} \tag{6}$$

The rationale of the rule is that the bank is allowed to lend up to the amount ℓ^{bf} that, weighted by the loan riskiness ω^λ and summed to its present risky weighted portfolio W^b, can be sustained by its equity base E^b, which should be at least κ % of $W^b + \omega^\lambda \ell^{bf}$. The loan riskiness is computed by considering first the borrower's default probability, along the lines of the Moody's KMV model , then by adopting an ad-hoc approximation of the so-called Basel II internal ratings approach that, given the credit rating of the borrower (here given by the default probability), provides the loan riskiness ω^λ. In particular, inspired by the Moody's KMV model (Saunders and Allen, 2010), we consider the balance sheet entries of the borrower as an indicator of its distance to default or, alternatively, probability of default and credit rating. In this context, the probability of default π^f of borrower f is defined as:

$$\pi^f = \frac{D^f + \lambda^f}{D^f + \lambda^f + E^f}, \tag{7}$$

where the rationale is that the lower is the capital base of the borrower with respect to its debt, the higher is the likelihood of default, because of possible equity losses due to negative earnings. We then stipulate that the risk weight of the loan depends on π^f through an ad-hoc function as follows:

$$\omega^\lambda = 2.5(\pi^f)^3. \tag{8}$$

This particular cubic function has to be considered as an approximation of the so-called Basel II internal ratings approach, after considering its graphical representation as in Yeh *et al.* (2005) (Figure 2).

3.2. Adaptive minimum capital requirements

Following the recommendations of the Basel committee, we implement into the model a mechanism to encourage[10] banks to build up and release capital buffers, according to the overall economic conditions of the economy, that are generally identified with "bad times" and "good times". The main idea is that in "good times" banks should build up a precautionary capital buffer and release it in bad times when credit is scarcely available. The concept of good and bad times is clearly related with expansion and contraction of the economic cycle, but the tricky part is to choose the conditioning variable that permits to identify if times are good or bad. Several possibilities have been already suggested. Drehmann *et al.* (2010) suggests three typologies of conditioning variables. The first one includes measures of the aggregate macroeconomic conditions (GDP growth, credit growth, or their ratio), the second one focuses on the banking sector activity (banks aggregate profits and losses, banks credit growth), while the last one is related to the cost of funding for banks (spreads and cost of liquidity). We choose to use a conditioning variable for each of the first two typologies, the unemployment rate as a measure of aggregate macroeconomic conditions and aggregate banks' loans as a measure of the banking sector activity.

The first rule for setting the minimum capital requirement κ is a piecewise linear rule that maps the unemployment rate into κ. The logic is that the capital requirement should be tighter when we are in "good times" (low unemployment rate) while they should be more relaxed when we are in "bad times" (high unemployment rate). The capital requirement[11] κ_t^u is then given by:

$$\kappa_t^u = \begin{cases} \kappa_{max} - (\kappa_{max} - \kappa_{min}) \dfrac{u_t}{\overline{u}} & if\ u_t < \overline{u}, \\ \kappa_{min} & if\ u_t \geq \overline{u}. \end{cases} \tag{9}$$

We use the notation κ_t^u here to highlight the dependence on time and on the unemployment rate. The values of κ_t^u then lies in an interval between a minimum level κ_{min}, reached when the unemployment is higher than a given threshold $\overline{u}$, and a maximum value κ_{max} that is assumed at full employment. In the

10. Let us note that we are "compelling" more than encouraging banks, because our mechanism cyclically moves minimum capital requirements, as shown in Equation (9) and (10). This may be questionable (see Repullo *et al.*, 2010; Repullo and Saurina, 2011; Gordy, 2009) but it appears to us an acceptable simplification, considering that in the model banks only have loans as risky assets in their portfolio

11. We omit the adjective "minimum" here to avoid confusion between κ_t^u that is the minimum level of capital required to banks at time t, and κ_{min} that is the minimum level under which the time varying κ can not drop, and that is independent from time.

computational experiments presented here, we have set $\kappa_{min} = 8\%$, which is the actual reference in the Basel II accords, $\kappa_{max} = 12\%$ and $\overline{u} = 25\%$. The parameter $\overline{u}$ determines the slope of the straight line in the (u_t, κ_t^u) plane, and represents the threshold unemployment rate above which the maximum leverage is set to κ_{min}.

The mechanism of the second rule is essentially the same as the previous one, using the aggregate loan portfolio of all banks as the conditioning variable. If L_t is the sum of all outstanding bank loans at month t, the decision making about capital requirement is made according to:

$$\kappa_t^c = \begin{cases} \kappa_{min} + (\kappa_{max} - \kappa_{min}) \dfrac{\Delta L_t}{\eta L_t} & if \ \dfrac{\Delta L_t}{L_t} < \eta, \\ \kappa_{max} \quad if \ \dfrac{\Delta L_t}{L_t} \geq \eta. \end{cases} \tag{10}$$

where $\Delta L_t/L_t$ is the percentage increase (or decrease) of aggregate credit L from month t–1 to month t. The parameter η represents in this case the threshold monthly credit growth above which κ_t^c is set to κ_{max}. The notation κ_t^c highlights the dependence on time and on the credit growth. In the set of simulations presented in this paper, $\eta = 5\%$, *i.e.*, if monthly credit growth is higher that 5%, capital requirement is the maximum one (12%). The rationale behind the rule is again to force banks to build up buffers in good times, which are usually characterized by a rapid growth of credit, that can be drawn down in bad ones. Buffers should be understood as capital in excess that is available to absorb losses in bad times. Real world experience as well as Eurace simulations have shown that a too rapid credit growth, while benefitting in the short term, can result in a bubble burst in the long run, when high levels of leverage become unsustainable. In this perspective, this new rule for κ_t^c can also be interpreted on one hand as a way to moderate a too rapid growth of credit by increasing the minimum requirement of capital for banks and, on the other hand, as an attempt to prevent the effects of banks's loans write-off on the credit supply. Finally, it is worth noting that whatever the negative value $\Delta L_t / L_t$ can be, we stipulate that capital requirement can never drop under the minimum level of κ_{min}, set to 8%.

4. Computational experiments

This section shows the results of computational experiments, performed in a setting with 2000 households, 20 consumption goods producers, 3 banks, 1 investment goods producer, 1 government, and 1 central bank.

We present several tables resuming the average values (with relative standard errors) of some of the main economic indicators in order to compare the overall performance of the artificial economy for three different values of minimum capital requirements. In the same tables, we also present the values of the same economic indicators in the two implemented cases of macro-prudential policy, *i.e.*, a countercyclical buffer capital mechanism that uses the unemployment rate as conditioning variable, as shown in Equation (9), and a second one based on banks credit growth, as in Equation (10). Moreover, we separate the simulations into two time periods. The first period shows the first 5 years of simulation while the second period represents the remaining 25 years.

Even in a different setting, results are in agreement with the explications discussed in Raberto *et al.* (2012). When capital requirement is low, the economic performance is good in the short term, due to the major amount of credit granted by banks, but turns bad in the long term because of the augmented financial fragility of firms that leads to a higher risk of insolvency. This mechanism closely mimics what is reported in BIS (2011), *i.e.*, that banks capital losses "can destabilize the banking sector, which can bring about or exacerbate a downturn in the real economy. This in turn can further destabilize the banking sector...". This is exactly what happens in our model. In the short run, real consumption, employment and investments (and therefore real GDP) are higher for loose capital requirements, due to the easier access to bank loans. Indeed, a worrying element is already present in the short run: firms leverage increases when relaxing capital requirement, raising economy's financial fragility (or systemic risk). Table 3 corroborates this narrative with the proper economic indicators. To observe the vicious circle between banking sector and real economy, emphasized by BIS (2011), we have to move to the medium-long run. Table 4 shows that the hierarchy emerged in the short run is completely reversed. Tighter capital requirements allow for better economic performances in the long run, in terms of consumption goods production, unemployment rate and real GDP level. Basically, loose capital requirements tend to push firms towards an unbalanced debt-equity ratio, with excess of debt. This can be observed in Figure 5, where firms leverage is plotted for a sample case. The higher financial fragility of firms determines a higher probability for firms to get into insolvency bankruptcy, as shown in Table 6, with obvious damage for commercial banks equity capital (see Figure 4) that undermines the resilience of the banking system. Banks capital losses trigger in turn a credit contraction that prevents firms to roll over debt, therefore exacerbating the downturn in the real economy thought chains of illiquidity bankruptcies. In Table 6 it can be observed a clear trend for banks equity and firms insolvency bankruptcies, showing that these economic indicators deteriorate when relaxing capital requirements.

Table 3. Values of the main economic indicators in the first 5 years of simulation for different capital requirements (κ)*

κ (%)	cons. goods production	inv. goods production	unempl. rate (%)	banks' loans	firms' leverage
8	9629 (25)	1620 (14)	2.84 (0.18)	151393 (1042)	3.21 (0.02)
10	9530 (29)	1530 (18)	2.86 (0.18)	143659 (1090)	2.85 (0.02)
12	9486 (27)	1486 (14)	2.88 (0.18)	138126 (840)	2.60 (0.01)
κ_t^u	9411 (40)	1442 (24)	2.95 (0.20)	136749 (1246)	2.58 (0.02)
κ_t^c	8518 (57)	1069 (12)	7.54 (0.57)	121650 (512)	2.23 (0.01)

* An unemployment rule κ_t^u (average value = **11.54**) and a credit rule κ_t^c (average value = **9.65**) has been used to set dynamic requirements. Values are averaged over 20 different random seeds (standard errors are in brackets).

Table 4. Values of real variables in the last 25 years of simulation for different capital requirements (κ)*

(%)	cons. goods production	inv. goods production	real GDP level	unempl. rate (%)
8	14296 (160)	3650 (188)	17946 (333)	8.1 (0.5)
10	14637 (126)	3460 (154)	18097 (243)	7.6 (0.4)
12	15081 (154)	3729 (137)	18810 (270)	6.2 (0.4)
κ_t^u	15040 (157)	3686 (131)	18727 (270)	5.5 (0.5)
κ_t^c	15419 (151)	3901 (103)	19320 (224)	3.2 (0.4)

* An unemployment rule κ_t^u (average value = **11.34**) and a credit rule κ_t^c (average value = **9.61**) has been used to set dynamic requirements. Values are averaged over 20 different random seeds (standard errors are in brackets)

Table 5. Values of price variables in the last 25 years of simulation for different capital requirements (κ)*

(%)	price index	inflation rate (%)	wage index	interest rate (%)
8	1.53 (0.03)	5.92 (0.20)	6.76 (0.13)	8.41 (0.21)
10	1.53 (0.02)	6.12 (0.23)	6.84 (0.09)	8.61 (0.23)
12	1.55 (0.02)	6.33 (0.17)	7.03 (0.12)	8.82 (0.15)
κ_t^u	1.51 (0.02)	6.22 (0.21)	6.97 (0.12)	8.58 (0.14)
κ_t^c	1.43 (0.02)	6.13 (0.12)	7.27 (0.15)	6.90 (0.25)

* An unemployment rule κ_t^u (average value = **11.34**) and a credit rule κ_t^c (average value = **9.61**) has been used to set dynamic requirements. Values are averaged over 20 different random seeds (standard errors are in brackets).

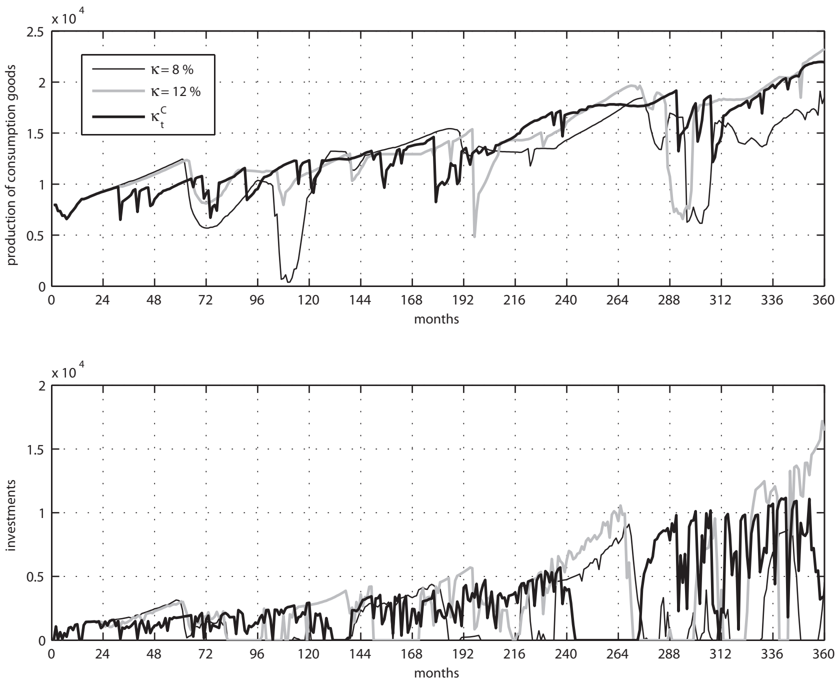
Table 6. Values of risk signaling variables in the last 25 years of simulation for different capital requirements (κ)

(%)	total loans	banks' equity	firms' leverage	illiquidity bankrupt.	insolvency bankrupt.
8	720375 (25922)	78003 (4331)	7.19 (0.85)	11.0 (0.5)	4.5 (0.2)
10	681196 (23748)	93022 (4280)	6.74 (0.58)	8.3 (0.4)	4.2 (0.1)
12	722717 (25495)	117812 (5448)	7.12 (1.26)	7.4 (0.4)	4.1 (0.2)
κ_t^u	698161 (17714)	105051 (3474)	6.76 (1.55)	9.0 (0.4)	3.8 (0.1)
κ_t^c	590332 (10978)	53261 (1841)	5.69 (3.31)	15.1 (0.8)	0.6 (0.1)

* An unemployment rule κ_t^u (average value = **11.34**) and a credit rule κ_t^c (average value = **9.61**) has been used to set dynamic requirements. Columns 4 and 5 show the annual bankruptcy probability for a firm. Values are averaged over 20 different random seeds (standard errors are in brackets).

From Figure 2 it clearly emerges the difference between $\kappa = 8\%$ (that is the minimum capital currently required by Basel II) and $\kappa = 12\%$. In both cases a turbulent period starts in the 5^{th} year (look at Figure 2), first with a severe crisis and then with a period of stagnation that ends in the 7^{th} year of simulation. Observing the dynamics of the aggregate banks equity capital in Figure 4, it can be noticed that a tighter capital requirement ($\kappa = 8\%$) forces banks to raise their equity while, when the requirement is lower ($\kappa = 12\%$), banks can match the credit demand with a lower level of equity capital. The capital surplus (or buffer) owned by banks in the case of $\kappa = 12\%$ makes the difference when facing the following crisis starting in the year, allowing the economy to recover quickly from the recession by means of an injection of new credit money by the banking sector. On the other hand, when equity capital is too low, the banking system is not able to fuel the economy with new credit, even if capital requirement is looser, and the depression continues much deeper and for a much longer time. It is interesting to observe the clear bifurcation of the two time trajectories representing the total outstanding credit, emerging as a consequence of the 9^{th} year crisis, and revealing the apparent paradox that banks will be much more "generous" lenders for the following 20 years in the case of higher capital requirements.

Figure 1. Time evolution of GDP components*



* Two lines correspond to different fixed values for minimum capital requirement κ, while the third one represents the outcome of the capital buffer rule based on credit growth as conditioning variable (κ_t^c).

Figure 2. Time evolution of GDP and unemployment rate*

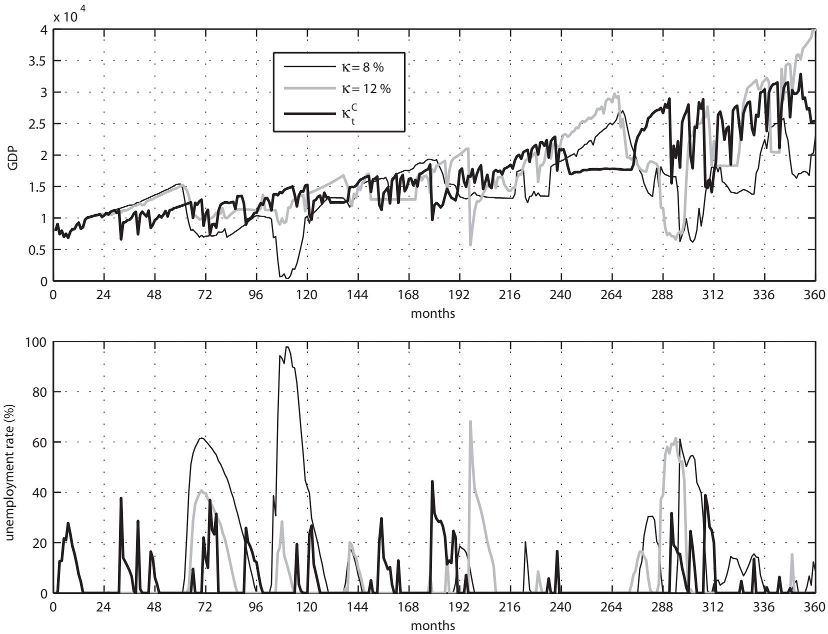
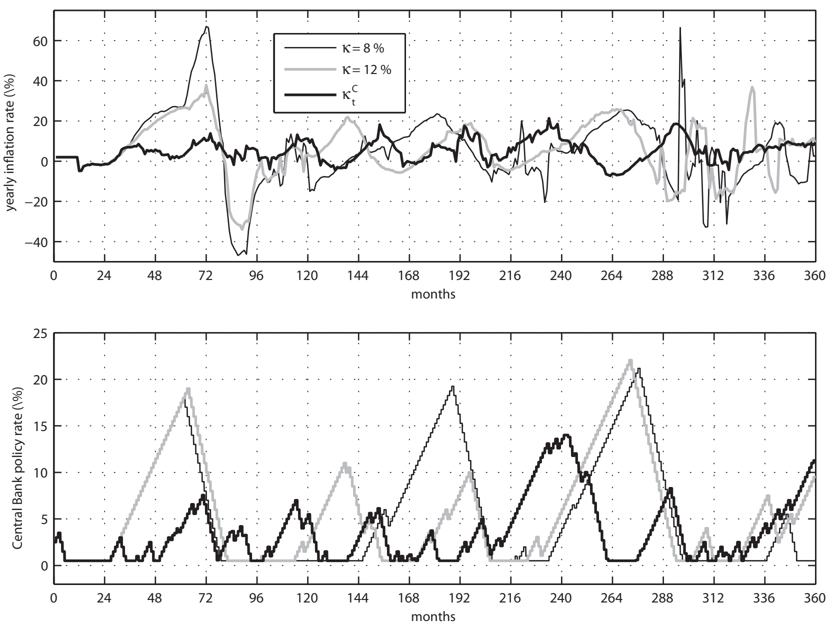


Figure 3. Time evolution of yearly inflation and base interest rate*



* Two lines correspond to different fixed values for minimum capital requirement κ, while the third one represents the outcome of the capital buffer rule based on credit growth as conditioning variable (κ_t^c).

Figure 4. Time evolution of total credit and aggregate banks equity capital*

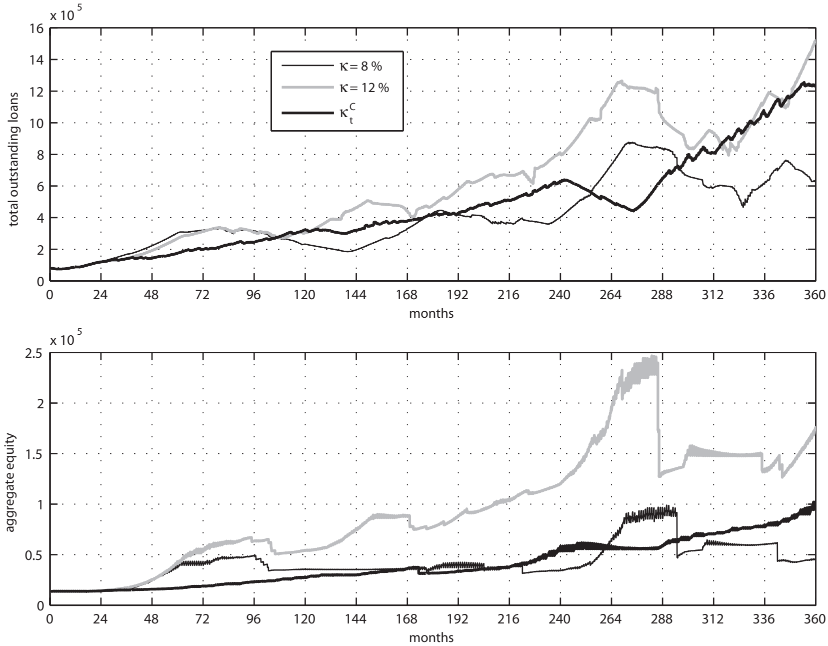
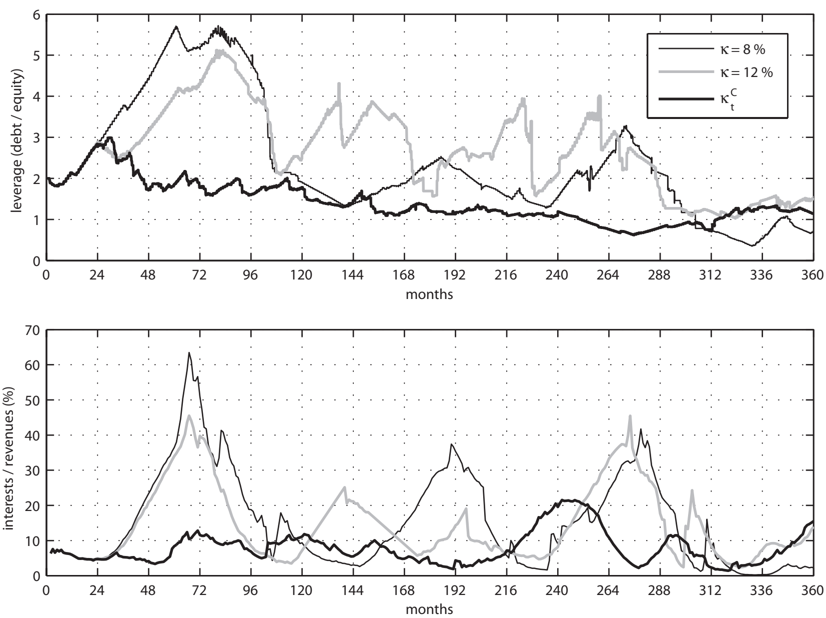


Figure 5. Time evolution of firms financial fragility indicators*



* Two lines correspond to different fixed values for minimum capital requirement κ, while the third one represents the outcome of the capital buffer rule based on credit growth as conditioning variable (κ_t^c).

The lesson learned by this example is that having a capital buffer in bad times is useful. As extensively commented in the introduction, the same lesson seems to be shared by the main banking policy regulators, thus we present here the results of two computational experiments where two different rules to build-up capital buffers have been implemented. The first one uses the unemployment rate as a conditioning variable, as reported in Equation (9), while the second uses banks loans, as in Equation (10).

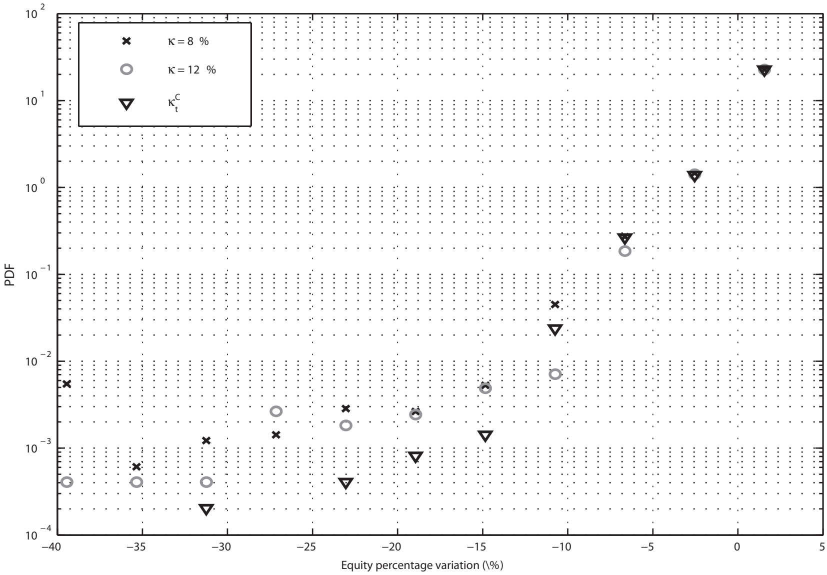
The two rules seem to have both a positive impact on the economic performance. In particular, the capital buffer rule based on the unemployment rate (*i.e.*, κ^u_t) improves all the economic indicators in the long run. These results confirm and extend preliminary explication discussed in Teglio *et al.* (2012). Table 4 shows that in the case of κ^u_t the average unemployment rate is lower than for fixed capital requirements, and the average output is higher. The interesting point about this result is that dynamic requirements κ^u_t, varying in a range from 8% to 12% with mean 11.34, perform better that the two range extremes, including the highest one (12%). So it is not true that tighter capital requirements are a better option overall, because it depends from the state of the economy. In some cases, according to our outcomes, relaxing the capital requirement is beneficial.

In order to analyze in more detail the effects of these policies, we plotted the time trajectories of the simulation with bank credit growth as a conditioning variable (case κ^c_t) along with the two extreme cases $\kappa = 8\%$ and 12% (for the seek of plots readability we omitted to superimpose the case of κ^u_t). Before looking at them, let us highlight the excellent performance of the "credit rule" κ^c_t. Whilst suffering in the first five years (see Table 3) the economy of the "credit rule" is characterized by the best economic indicators: consumption, investments, unemployment rate, real wages, firms financial fragility, and others. Tables from 4 to 6 attest these results. The only apparently jarring note in this idyllic picture is banks equity, significantly lower than the rest of the cases. Moreover, from an overall glance at the presented plots it clearly appears that the economy is considerable more stable in the case of the "credit rule". In the following, we try to interpret these outcomes.

First, let us note that the initial conditions are the same for all cases, and that banks are characterized at the beginning by a strong capitalization with equity equal to 20% of weighted assets. We remind also that banks raise their equity capital when the capital requirement constraint is binding, *i.e.*, when they are not able to satisfy the credit demand because of the low equity. The "credit rule" states that when credit grows, capital requirement also grows, but it remains lower when credit stagnates. So, at the beginning of the κ_t^c case simulation, there is no need for banks to raise the equity because of the low credit growth (see Figure 4). In the other two cases (κ= 8% and 12%), on the other hand, banks have to raise equity to match the growing credit demand. However, looking to Figure 5, we can see how firms high borrowing raises the interest bill, increasing consequently firms bankruptcy risk. On the contrary, in the κ_t^c case the situation is much more stable. In fact, the high level of firms debt ends in a big economic crash around month 100, accompanied by a quick deleveraging process. This crisis, strong for κ= 8% and milder for κ= 12%, does not affect the κ_t^c case at all. When the "credit rule" rule is active, the total outstanding credit is growing much more smoothly. The same holds for banks equity, that is rarely affected by falls caused by insolvency bankruptcies (Figure 6). Illiquidity bankruptcies are more frequent but the banking system is not directly affected, remaining robust and propagating its robustness to the real economy.

Figure 6 compares banks equity variations considering the whole simulation set. The higher robustness of the banking sector in the case of κ_t^c is therefore extended from a single simulation seed to the general case. In particular, it can be observed from the plot that the falls of banks equity are much more infrequent in case banks build up capital buffers according to the rule based on credit growth. These results tell us that lower aggregated values of outstanding bank loans and of banks equity capital do not have negative implications on the economy if credit is granted at the right time. In this sense, it seems to be desirable to have active (public) institutions in charge of ruling banks credit supply according to the macroeconomic conjuncture.

Figure 6. Probability density function (PDF) estimation of the equity capital percentage variations for three different regulatory policies



5. Concluding remarks

After the recent financial and economic crisis, the Basel III global regulatory standard has been proposed in order to improve the resilience of the banking system. This new framework is oriented towards a more active support of macroeconomic policies, and presents a set of macro-prudential regulations with the objective to limit systemic risk. In particular, it has been identified the procyclical amplification of financial shocks through the banking system as a critical issue, and in order to cope with such procyclical dynamics new countercyclical capital buffers regulations have been proposed. The rational is to encourage banks to accumulate capital during good times and use it when the economic conditions are bad.

A coherent economic analysis has emerged from some previous works by the authors, where computational experiments were performed with the Eurace model. In particular, it has been shown that excessive bank leverages can drive economies into severe recession in the medium-long run. In this paper, we implement into the model a mechanism to encourage banks to build up and release capital buffers, according to the overall economic conditions of the economy. As conditioning variables that set banks capital requirements, thus ruling the capital buffers build up and release phases, we use both the unemployment rate and aggregate banks loans growth.

Moreover, the setting of the model has been improved incorporating a new demand for firms investments that is based on the net present value (NPV) approach, which is one of the most popular methods used by managers to evaluate investments. Therefore, the new model for investments can be considered well grounded on a more realistic behavior of economic actors.

Results confirm and extend our previous studies, showing that loose capital requirements can affect the economic performance in the medium-long run, raising the financial fragility (or systemic risk) in the economic system and potentially triggering chains of firms insolvency bankruptcies. The situation is generally better when setting tighter capital requirements. Furthermore, results have shown that the dynamic regulation of capital requirements successfully stabilizes the economy and improves the main economic indicators. In particular, when the "credit rule" is adopted, the economic scenario seems to change in a significant way, showing a much more solid banking sector with a resulting positive effect on the real economy.

References

Acharya VV., 2009. "A theory of systemic risk and design of prudential bank regulation." *Journal of Financial Stability* 5(3): 224–255.

Allen F., and D. Gale, 2000. "Bubbles and crises." *Economic Journal* 110(460): 236–55.

Andersen H., 2011. "Procyclical implications of basel ii: Can the cyclicality of capital requirements be contained?." *Journal of Financial Stability* 7(3): 138–154.

Antão P., and A. Lacerda, 2011. "Capital requirements under the credit risk-based framework." *Journal of Banking & Finance* 35(6): 1380–1390.

BIS, 2011. *Basel III: A global regulatory framework for more resilient banks and banking systems*. Tech. rep., Basel Committee on Banking Supervision, URL http://www.bis.org/publ/bcbs189.htm

Blum J., and M. Hellwig, 1995. "The macroeconomic implications of capital adequacy requirements for banks." *European Economic Review* 39(3-4): 739–749.

Borio C., 2011. "Rediscovering the macroeconomic roots of financial stability policy: journey, challenges and a way forward." Tech. rep., *Bank for International Settlements Working Paper* 354.

Carroll CD., 2001. "A theory of the consumption function, with and without liquidity constraints." *Journal of Economic Perspectives* 15(3): 23–45.

Cecchetti SG., and L. Li, 2008. "Do capital adequacy requirements matter for monetary policy?." *Economic Inquiry* 46(4): 643–659.

Chiarella C., and C. Di Guilmi, 2011. "The financial instability hypothesis: A stochastic microfoundation framework." *Journal of Economic Dynamics & Control* 35(8): 1151–1171.

Cincotti S., M., Raberto, and A.Teglio, 2010. "Credit money and macroeconomic instability in the agent-based model and simulator eurace." The Open-Access, *Open-Assessment E-Journal* 4 (2010-26), URL http://dx.doi.org/10.5018/economics-ejournal.ja.2010-26

Cincotti S., M. Raberto and A. Teglio, 2012. "The Eurace macroeconomic model and simulator. In: Agent-based Dynamics, Norms, and Corporate Governance." The proceedings of the 16-th World Congress of the International Economic Association, *Palgrave*, vol II, forthcoming.

Deaton A., 1992. "Household saving in ldcs: credit markets, insurance and welfare." *The Scandinavian Journal of Economics* 94 (2): 253–273.

Delli Gatti D., C. Di Guilmi, E. Gaffeo, G. Giulioni, M. Gallegati and A. Palestrini, 2005. "A new approach to business fluctuations: heterogeneous interacting agents, scaling laws and financial fragility." *Journal of Economic behavior & organization* 56 (4): 489–512.

Delli Gatti D., M. Gallegati, BC. Greenwald, A. Russo, JE. Stiglitz, 2009. "Business fluctuations and bankruptcy avalanches in an evolving network economy." *Journal of Economic behavior & organization* 4 (2): 195–212.

Dosi G, G. Fagiolo and A. Roventini, 2010. "Schumpeter meeting keynes: A policy-friendly model of endogenous growth and business cycles." *Journal of Economic Dynamics and Control* 34 (9): 1748–1767.

Drehmann M., C. Borio, L. Gambacorta, G. Jiminez, and C. Trucharte, 2010. *Countercyclical capital buffers: exploring options*. Bis working papers, Bank for International Settlements.

Fazzari SM., P. Ferri, and E. Greenberg, 2008. "Cash flow, investment, and Keynes-Minsky cycles." *Journal of Economic Behavior & Organization* 65 (3-4): 555–572.

Gordy M., 2009. *First, do no harm – a hippocratic approach to procyclicality in Basel II*. Paper presented at the conference Procyclicality in the financial system, jointly organised by the Netherlands Bank and the Bretton Woods Committee.

Graham J., and C. Harvey, 2001. "The theory and practice of corporate finance: evidence from the field." *Journal of Financial Economics* 60 (2-3): 187–243.

Graham J., and C. Harvey, 2002. "How do CFOs make capital budgeting and capital structure decisions?." *Journal of Applied Corporate Finance* 15 (1):8–23.

Heid F., 2007. "The cyclical effects of the Basel II capital requirements." *Journal of Banking & Finance* 31(12): 3885–3900.

Hillier F., and G. Lieberman, 1986. *Introduction to Operations Research*. McGraw-Hill.

Kahneman D, and A. Tversky, 1979. "Prospect theory: an analysis of decision under risk." *Econometrica* 47(2):263–292.

Myers S., and N. Majluf, 1984. "Corporate financing and investment decisions when firms have information that investors do not have." *Journal of Financial Economics* 13(2):187–221.

Raberto M., A. Teglio, and S. Cincotti, 2008 a. "Integrating real and financial markets in an agent-based economic model: an application to monetary policy design." *Computational Economics*." Vol 32 (1-2), pp. 147-162.

Raberto M, A. Teglio and S. Cincotti, 2008 b. "Prospect Theory Behavioral Assumptions in an Artificial Financial Economy, Lecture Notes in *Economics and Mathematical Systems*, vol 614, Springer, pp 55–66.

Raberto M., A. Teglio and S. Cincotti, 2012. Debt deleveraging and business cycles. an agent-based perspective. The Open-Access, Open-Assessment E-Journal, Vol. 6, 2012-27. URL http://dx.doi.org/10.5018/economics-ejournal.ja.2012-27

Repullo R., and J. Saurina, 2011. *The countercyclical capital buffer of basel iii: A critical assessment*. CEPR Discussion Papers.

Repullo R., J. Saurina and C. Trucharte, 2010. "Mitigating the pro-cyclicality of basel ii." *Economic Policy* 25: 659–702.

Saunders A., and L. Allen, 2010. *Credit Risk. Measurement In and Out of the Financial Crisis*. Wiley.

Teglio A., M. Raberto, and S. Cincotti, 2009. "Explaining equity excess return by means of an agent-based financial market." Lecture Notes in *Economics and Mathematical Systems*. Vol 631, Springer Verlag, chap 12, pp 145–156.

Teglio A., M. Raberto and S. Cincotti, 2010a. "Balance sheet approach to agent-based computational economics: The Eurace project." In: *Combining Soft Computing and Statistical Methods in Data Analysis, Advances in*

Intelligent and Soft Computing. Vol 77, Springer Berlin / Heidelberg, pp. 603–610.

Teglio A., M. Raberto and S. Cincotti, 2010 b. "Endogenous credit dynamics as source of business cycles in the Eurace model." In: Li Calzi M, Milone L, Pellizzari P (eds) Progress in Artificial Economics, Lecture Notes in *Economics and Mathematical Systems*. Vol 645, pp 203–214.

Teglio A., M. Raberto and S. Cincotti, 2012. *The impact of banks' capital adequacy regulation on the economic system: an agent-based approach*. Advances in Complex Systems. Vol. 15, issue supp2.

Tesfatsion L., K. Judd, 2006. "Agent-Based Computational Economics." *Handbook of Computational Economics*. Vol 2. North Holland.

Tversky A., and D. Kahneman, 1992. "Advances in prospect theory: cumulative representation of uncertainty." *Journal of Risk and Uncertainty* 5 (4): 297–323.

Yeh A., J. Twaddle and M. Frith, 2005. "Basel II: A new capital framework." Tech. rep., *Reserve Bank of New Zealand: Bulletin*. Vol. 68, no. 3.

WAGE FORMATION, INVESTMENT BEHAVIOR AND GROWTH REGIMES: AN AGENT-BASED ANALYSIS[1]

Mauro Napoletano
OFCE and SKEMA Business School, Sant'Anna School of Advanced Studies
Giovanni Dosi and Giorgio Fagiolo
University of Verona, Sant'Anna School of Advanced Studies
Andrea Roventini
University of Verona, Sant'Anna School of Advanced Studies, OFCE

Using the "Keynes+Schumpeter" (K+S) agent-based model developed we study how the interplay between firms' investment behavior and income distribution shapes the short—and long-run dynamics of the economy at the aggregate level. We study the dynamics of investment under two different scenarios. One in which investment is fully determined by past profits, and one in which investment is tied to expectations about future consumption demand. We show that, independently from the investment scenario analyzed, the emergence of steady growth with low unemployment requires a balance in the income distribution between profits and wages. If this is not the case, the economy gets locked either in stagnation equilibria, or into growth trajectories displaying high volatility and unemployment rates. Moreover, in the demand-led scenario we show the emergence of a non-linear relation between real wages and unemployment. Finally, we study whether increasing degrees of wage-flexibility are able to restore growth and unemployment and reduce the volatility in the economy. We show that this is indeed the case only when investment is profit-led. In contrast, in the scenario where investment is driven by demand expectations wage-flexibility has no effect on either growth and unemployment. In turn, this result casts doubts on the ability of wage-flexibility policies to stabilize the economy.

Keywords: Agent-Based Models, Growth Regimes, Income Distribution, Wage-Flexibility.

1. The authors gratefully acknowledge the financial support of the Institute for New Economic Thinking (INET) grant #220, "The Evolutionary Paths Toward the Financial Abyss and the Endogenous Spread of Financial Shocks into the Real Economy".

This work studies how the interplay between firms' investment behavior, wage formation and income distribution affect the short-run and long-run aggregate dynamics of an economy. We study the aggregate behavior of the economy under two different scenarios: one wherein investment in new productive capacities is determined by past profits of firms ("profit-led scenario") and one in which firm investment depends on expectations about future consumption demand ("demand-led" scenario). We show that in both scenarios the distribution of income between profits and wages crucially affects the characteristics of the growth path of the economy. Independently from the investment rule adopted by firms, the emergence of long-run growth associated with low rates of unemployment and short-run volatility always requires a balance in the distribution between profits and wages. Lacking such a requirement, the economy can get stuck either into stagnation equilibria with low growth and high unemployment, or into trajectories characterized by high and volatile growth. Furthermore, we study the relation between the level of real wages and unemployment in the economy. We show that the economy displays the Neo-Classical (positive) relation between the two variables only in the profit-led scenario. In contrast, in the demand-led scenario such a relation is non-linear. In particular, it exists a threshold below which unemployment increases (rather than decreasing) with a reduction in real wages. Finally, we explore whether increasing degrees of nominal wage sensitivity to unemployment variations are effective in curbing unemployment and stabilizing the economy. Our results show that this is the case only in the profit-led scenario, provided that the characteristics of the income distribution allow the economy to grow in the long-run. In contrast, in the demand-led scenario, wage-flexibility is never able to reduce unemployment, and can sometimes increase the incidence of economic crises.

Our work is motivated by two different streams of literature. First, we refer to a central debate in macroeconomics, *i.e.* the one about the role of wages in determining unemployment in the economy. On one side, the "Neo-Classical View" identifies into real wage rigidity the main source of unemployment in labor markets. Following an adverse shock to the economy, production and labor demand will fall, and real wages must also decrease in order to equilibrate demand and supply of labor. A rigid real wage impedes such an adjustment and thus leads to unemployment in the labor market. In contrast, downward flexibility of nominal wages allows the reduction of the real wage down to the new level compatible with full employment and thus the reduction of unemployment.[2] On the opposite camp, we find all works that starting from Keynes (1936) pointed to deficiencies in effective demand as the main source of unemployment. Moreover, in line with the intuition of Keynes, these works (see Howitt, 1986; Amendola *et al.*, 2004) warned against the destabilizing effect of the downward nominal wage flexibility . This is because a reduction of nominal wages is likely to adversely affect consumption demand and to induce deflationary pressures in the economy, with the consequence of increasing (rather than decreasing) the level of unemployment in the economy.[3]

Second, our work is related to both evolutionary (Freeman and Perez, 1988; Coriat and Dosi, 2000; Chiaromonte *et al.*, 2000) and French "Régulation" research programs (see Aglietta, 1979, Boyer, 1988, see Lordon, 1991, for a survey), that study how "growth regimes", as well as crises, are generated by the matching or mismatching between, on the one hand, processes of technical change and, on the other hand, the characteristics of the processes governing firms' behavior and the division of income in the economy.

We contribute to the above strands of literature along several dimensions. First, we show how different growth regimes emerge out of micro-interactions between heterogeneous agents. In that fashion, our work provides a *micro-foundation* of early evolutionary and "regulationist" theories of the role of institutions, demand-formation patterns, and technical change in determining both

2. This approach is also largely followed by New Keynesian DSGE models (e.g. Smets and Wouters, 2007). These models claim that price and wage rigidities constitute the main impediments to full employment. Without such rigidities and in line with the Neo-Classical View the economy would be able to adjust to whatever shock and keep the labour market at the full employment equilibrium.

3. Another strand of research in the New Keynesian literature (Greenwald and Stiglitz, 1993) warns against the perils of downward flexibility of nominal wages. Still, in these models unexpected increases in real wages increase firms costs, thus leading to a reduction of firms' net worth and to lower levels of investment and output.

growth and business cycles. Second, we highlight the role of income distribution as a crucial factor determining the characteristics of growth paths followed by the economy. Third, by studying the relation between wages and unemployment into different investment scenarios, we study under which conditions one observes the validity of either the Neo-Classical or Keynesian view about the effectiveness of policies increasing the degree of wage flexibility in the economy. Finally, we extend the analysis of wage-flexibility to encompass also its long-term effects on the economy.

We perform the above investigations by extending the "Keynes+Schumpeter" (K+S) model developed in Dosi *et al.* (2010) and Dosi *et al.* (2012), that bridges Keynesian theories of demand-generation and Schumpeterian theories of technology-fueled economic growth. In Dosi *et al.* (2010), we studied the consequences of different "innovation regimes"—and related policies—and their interaction with (Keynesian) demand management. In Dosi *et al.* (2012) we focused instead on the interactions between income distribution on one hand, and monetary and fiscal policies on the other hand. In this paper we exclude fiscal and monetary policies from the picture, and we rather study how the interactions between income distribution and firm investment behavior affect the short—and long-run dynamics at the aggregate level. In addition, we try to assess whether different levels of nominal wage sensitivity to variations in unemployment may promote long-run growth and reduce output volatility and unemployment.

The work is structured as follows. Section 1 briefly presents the K+S model. Section 2 presents the simulation results, starting with the analysis of growth regimes under different income distributions and investment behavior (Section 2.1), and then moving to the analysis of the effects of wage-flexibility to unemployment variations (Section 2.2). Section 3 concludes.

1. The K+S model

Let us now briefly discuss the K+S model developed in Dosi *et al.* (2010) and extended in Dosi *et al.* (2012), to which we refer for more details. The model portrays an economy composed of a machine-producing sector made of F_1 firms (denoted by the subscript i), a consumption-good sector made of F_2 firms (denoted by the subscript j), L^S consumers/workers, and a public sector. Capital-good firms invest in R&D and produce heterogeneous machines. Consumption-good firms combine machine tools bought by capital-good firms and labor in order to produce a final product for consumers. Capital-good firms are paid in advance for the machines they have to produce. Consumption-good firms finance their production and investment expenditures by using internal funds and external financing provided by an un-modeled banking sector. The latter provides credit to firms up to a credit ceiling that depends of firms' past sales.[4] Finally, the public sector levies taxes on firms' profits and pays unemployment benefits. In what follows, we present the timeline of events in the K+S model. Next we will briefly describe each part of it.

1.1. The timeline of events

In any given time period (t), the following microeconomic decisions take place in sequential order:

1. Policy variables are fixed (e.g. the "Government" setting tax rates and unemployment benefits, etc.).
2. Machine-tool firms perform R&D trying to discover new products and more efficient production techniques and to imitate the production technology and the products of their competitors. Capital-good firms advertise their machines to consumption-good firms.
3. Consumption-good firms decide how much to produce and invest. If investment is positive, consumption-good firms choose their supplier, send their orders and pay for the machines. When internal funds are not enough to finance production and investment plans, firms borrow up to a ceiling.
4. In both industries firms hire workers according to their production plans if below their credit ceiling or at the ceiling otherwise and start producing.
5. Imperfectly competitive consumption-good market opens. The market shares of firms evolve according to their price competitiveness.

4. In Dosi *et al.* (2012) we model a banking sector that gathers deposits from firms and provides credit to them on a pecking order that depends on the firms' past net-worth-to-sales ratio.

6. Firms in both sectors compute their net cash flow, pay back their due loans to the bank to the extent that they have cash flow to do that.
7. Entry and exit take place. In both sectors firms with near-zero market shares and/or a negative stock of net liquid assets are eschewed from their industry and replaced by new firms (for simplicity, we keep the number of firms fixed; any dead firm is replaced by a new one; and entrant firms are random copies of incumbent ones).
8. Machines ordered at the beginning of the period are delivered and become part of the capital stock at time $t + 1$.

At the end of each time step, aggregate variables (e.g. GDP, investment, employment) are computed, summing over the corresponding microeconomic quantities.

1.2. The capital-good industry

The technology of capital-good firms evolves along the vintages of produced machine-tools. Each firm-specific generation of machine-tools has indeed a distinct production cost and distinct labour productivity for the user. The price of machines is set with a mark-up rule[5] over production costs. The quality of each vintage is measured by the productivity of machines in the consumption-good sector.

Innovation and imitation are costly processes: firms invest in R&D a fraction of their revenues and hire researchers at the current market wage.

Both innovation and imitation follow a two-steps stochastic process. In the first step, the resources allocated to search determine in probability whether the events "innovation" and "imitation" are drawn. Note that the newly discovered capital goods might be a "failed innovation", because production costs might be higher and/or user-efficiency might be lower than the currently manufactured machines. Indeed, at the second stochastic stage, each firm draws the characteristics of the would-be machine and decide whether to keep on producing the current generation of machines or to switch to the new vintage, by evaluating the

5. This in line with survey data evidence on firm pricing behavior (see Fabiani *et al.*, 2006).

possible trade-off between production costs and productive efficiencies. Once the machine tool is chosen, capital-good firms try to reach their customers under conditions of imperfect information: hence, we assume that they send a "brochure" with the price and the productivity of their machines to both their historical clients and a random sample of potential new customers.

1.3. The consumption-good industry

Consumption-good firms produce an homogenous good using capital (*i.e.* their stock of machines) and labor under constant returns to scale. We assume alternative scenarios for firms' production and investment decisions.[6] In the first scenario, we assume as in Greenwald and Stiglitz (1993) and Delli Gatti *et al.* (2005) that desired production (Q_j^d) is determined by the level of firm stock of liquid assets (NW_j), according to[7]:

$$Q_j^d(t) = \sigma NW_j(t-1), \quad \sigma > 0 \tag{1}$$

In turn, the level of firms' net worth is determined—*via* cash flows—by the past level of profits. Following Boyer (1988) we label the above *profit-led investment scenario*. Notice that we here we attempt to describe the economic dynamics in a highly hypothetical scenario wherein both desired production and desired investment are not limited by demand.[8]

In contrast, in the second scenario, that we label *demand-led investment scenario*, firms plan their production according to adaptive demand expectations (D_j^e):

$$D_j^e(t) = f(D_j(t-1), D_j(t-2), \ldots, D_j(t-h)), \tag{2}$$

where $D_j(t-1)$ is the demand actually faced by firm j at time $t-1$ (h positive integer)[9]. The desired level of production depends on

6. To simplify notation and unless it needed for the sake of clarity, in what follows we suppress the time index to indicate variables in the text.
7. This kind of firm behavior may emerge in models where firms are equity rationed and face positive bankruptcy costs. See Greenwald and Stiglitz (1993) for more details.
8. As it is discussed at more length in Boyer (1988), this kind of scenario however captures some key features of the dynamics of investment in capitalistic economies of the 19th century.
9. For maximum simplicity, here we use the rule $D_j^e(t) = D_j(t-1)$. In Dosi *et al.* (2006) we check the robustness of the simulation results employing more sophisticated expectation-formation rules. We found that increasing the computational capabilities of firms does not significantly change either the average growth rates or the stability of the economy. These properties still hold in the model presented here.

the expected demand as well as on the desired inventories (N_j^d) and the actual stock of inventories (N_j):

$$Q_j^d(t) = D_j^e(t) + N_j^d(t) - N_j(t-1), \tag{3}$$

with $N_j^d(t) = \iota D_j^e(t), \iota \in [0,1]$.

Finally, in both scenarios, the output of consumption-good firms is constrained by their capital stock (K_j). If the desired capital stock (K_j^d)—computed as a function of the desired level of production—is higher than the current capital stock, firms invest (EI_j^d) in order to expand their production capacity[10]:

$$EI_j^d(t) = K_j^d(t) - K_j(t). \tag{4}$$

Consumption-good firms have a capital stock composed of heterogenous machines having different productivity levels. Firms decide whether to scrap their machines following a payback period rule, that is they assess whether the substitution cost of any current machine, *i.e.* the price of a new one, can be recovered in a given number of years through the savings obtained in production costs (new machines have lower unit production cost than incumbent ones). In this way, technical change and capital-good prices affect the replacement decisions of consumption-good firms.[11] The latter choose their capital-good supplier comparing the price and productivity of those machine tools which they know via the brochures they received. Machine production is a time-consuming process: consumption-good firms receive the ordered machines at the end of the period.[12] Gross investment of each firm is the sum of expansion and replacement investment. Aggregate investment is just the sum of the investments of all consumption good firms.

Given their current stock of machines, consumption-good firms compute their average productivity and unit costs of production. Firms fix prices applying a variable mark-up (μ_j) over the latter. More precisely, we set an initial value for the mark-up $\bar{\mu}(0)$, which

10. We assume that in any give period firm capital growth rates cannot exceed a fixed maximum threshold consistent with the maximum capital growth rates found in the empirical literature on firm investment patterns (Doms and Dunne, 1998).

11. This in line with a large body of empirical papers (Feldstein and Foot, 1971; Eisner, 1972; Goolsbee, 1998) showing that replacement investment is typically not proportional to the capital stock.

12. The presence of gestation-lag effects in firm investments expenditures is supported by a large body of empirical literature (see Del Boca *et al.*, 2008).

is equal across firms. The variation of mark-ups over time are regulated by the evolution of firms' market shares (f_j): firms raise (cut) mark-up whenever the growth rate of their market shares is positive (negative):

$$\mu_j(t) = \mu_j(t-1)\left(1 + \upsilon \frac{f_j(t-1) - f_j(t-2)}{f_j(t-2)}\right), \tag{5}$$

with $0 \leqslant \upsilon \leqslant 1$. This process in turn implies that the average mark-up rate $\overline{\mu}(t)$ (as well as firms' ones) fluctuate around a sort of peg represented by the initial mark-up rate $\overline{\mu}(0)$. The level of real wages (*w/cpi*) is determined by the average mark-up rate in the consumption-good sector, it follows that by by tuning up and down the level $\overline{\mu}(0)$ one can vary the long-term income distribution between wages and profits in the economy.

Prices are one the key determinants of firms' competitiveness. The other ones are the levels of unfilled demand. If firms cannot fully satisfy their customers, their competitiveness is accordingly reduced.

Market shares evolve according to a replicator-type dynamics operating under conditions of imperfect information,[13] so that even if the product is homogeneous, firms may charge different prices. In such dynamics, firms with above-average competitiveness expand their market shares, while those below shrink (or even die).

1.4. The labor market

We do not impose any assumption of labor-market clearing: as a consequence involuntary unemployment as well as labor rationing are the rule rather than the exception. The aggregate labor demand is computed summing up the labor demand of capital—and consumption—good firms. The aggregate supply of labor is exogenous and inelastic. Aggregate employment is then the minimum between labor demand and supply. The wage is set according to:

$$w(t) = w(t-1)\left(1 + \psi_1 \frac{\Delta \overline{AB}(t)}{\overline{AB}(t-1)} + \psi_2 \frac{\Delta cpi(t)}{cpi(t-1)} + \psi_3 \frac{\Delta U(t)}{U(t-1)}\right), \tag{6}$$

13. See Rotemberg (2008) for a survey of the empirical literature on consumers' imperfect price knowledge.

where $\overline{AB}(t)$ is the average labor productivity, $cpi(t)$ is the consumer price index, and $U(t)$ is the unemployment rate. The wage rate is determined by institutional and market factors, with both indexation mechanisms upon consumption prices and average productivity, on the one hand, and, adjustments to unemployment rates, on the others. Notice that, by varying the magnitude of the parameters ψ_1,ψ_2 and ψ_3 in Equation (6), and the initial mark-up rate $\bar{\mu}(0)$ we are able to tune the distribution of productivity gains between workers and firms as well as the sensitivity of wages no unemployment variations (see Section 2 fore more discussion). In this way, and inn line with works in the "Régulation" literature, we capture different institutional regimes governing labor market dynamics and demand formation (see Boyer, 1988, for a taxonomy of those regimes). More precisely, in Section 2.1 we begin our analysis of the role of wages and income distribution in determining aggregate dynamics by restricting ourselves to a regime wherein wage just grows with average productivity, *i.e.* where $\psi_2 = \psi_3$ and = 0 and $\psi_1 > 0$. In Section 2.2, we remove this hypothesis and we perform experiments where nominal wages are also a function of variations in unemployment.[14]

1.5. Consumption, taxes, and public expenditures

As in Dosi *et al.* (2010) and Dosi *et al.* (2012) the model has a public sector that levies taxes on firm profits and worker wages (or on profits only) and pays to unemployed workers a subsidy, that is a fraction of the current market wage. In those models, redistributive fiscal policies significantly affect the aggregate dynamics both in the short- and in the long-run. Therefore, in what follows we set both the tax and unemployment subsidy rate to zero. This allows us to better analyze the role of income distribution and wage formation on aggregate dynamics, and the one of nominal wage flexibility on unemployment in particular.

All wages are consumed in the model. The aggregate consumption (C) is the sum of income of both employed and unemployed

14. In the experiments we present below average inflation is always very close to zero (see Tables 2 and 3). However, we also experimented with regimes wherein wage move also as a function of inflation rates (*i.e.* where $\psi_2 > 0$). All the properties discussed below robustly hold.

workers, as the model satisfies the standard national account identities: the sum of value added of capital—and consumption-goods firms (Y) equals their aggregate production since in our simplified economy there are no intermediate goods, and that in turn coincides with the sum of aggregate consumption, investment and change in inventories (ΔN):

$$\sum_{i=1}^{F_1} Q_i(t) + \sum_{j=1}^{F_2} Q_j(t) = Y(t) \equiv C(t) + I(t) + \Delta N(t). \tag{7}$$

The micro decisions of a multiplicity of heterogenous, adaptive agents and their interaction mechanisms is the explicit microfoundation of the dynamics for all aggregate variables of interest (e.g. output, investment, employment, etc.).

2. Simulation results

In line with Dosi *et al.* (2010) and Dosi *et al.* (2012), we investigated the micro and macro properties of the model through extensive Monte-Carlo simulations.[15] We perform our simulation analysis in two complementary steps (see also the paper by Fagiolo and Roventini in this issue for a discussion of this methodology). First, we identify a "benchmark" setup for which the model is empirically validated (see Table 1), *i.e.* it is able to replicate a wide spectrum of microeconomic and macroeconomic stylized facts. Next, we turn to a battery of "experiments", by identifying sets of parameters (e.g. the level of the initial mark-up rate, the degree of wage-indexation to unemployment) whose values capture different structural conditions and/or policies.

The macro and micro stylized facts robustly replicated by the model are the same statistical regularities produced by and discussed at much greater length in Dosi *et al.* (2010) and Dosi *et al.* (2012). There we show that the model is able to generate macroeconomic time-series of output, consumption and aggregate investment characterized by self-sustained growth patterns and by persistent fluctuations. Moreover, aggregate investment is more volatile than GDP whereas consumption is less volatile.

15. All results discussed below refer to averages over $MC = 50$ Monte-Carlo iterations. Each iteration has $T = 600$ time-steps.

Table 1. Benchmark Parameters

Description	Symbol	Value
Number of firms in capital-good industry	F_1	50
Number of firms in consumption-good industry	F_2	200
R&D investment propensity	υ	0.04
R&D allocation to innovative search	ξ	0.50
Firm search capabilities parameters	$\zeta_{1,2}$	0.30
Beta distribution parameters (innovation process)	(α_1, β_1)	(3,3)
Beta distribution support (innovation process)	$[\underline{x}_1, \overline{x}_1]$	[−0.15,0.15]
New-customer sample parameter	γ	0.50
Desired inventories	ι	0.10
Payback period	b	3
"Physical" scrapping age	η	20
Capital-good firm mark-up rate	μ_1	0.04
Consumption-good firm initial mark-up	$\bar{\mu}(0)$	0.10
Coefficient in the consumption-good desired production rule *(profit-led scenario)*	σ	2
Coefficient in the consumption-good firm mark-up rule	υ	0.10
Competitiveness weights	$\omega_{1,2}$	1
Replicator dynamics coefficient	χ	1
Wage setting $\Delta \overline{AB}$ weight	ψ_1	1
Wage setting Δcpi weight	ψ_2	0
Wage setting ΔU weight	ψ_3	0
Tax rate	tr	0
Unemployment subsidy rate	φ	0
Loan-to-value ratio	λ	2
Baseline Interest Rate	r	0.025

In addition, the model replicates the empirically observable co-movements between a large set of macro time series and GDP (net investment and consumption pro-cyclical coincident, inflation pro-cyclical and lagging, counter-cyclical mark-up rates, etc.). In particular, the K+S model is also able to replicate two among the most relevant statistical properties characterizing labor markets, namely the fact that unemployment is strongly counter-cyclical whereas real wages display little variation at business cycles frequencies. At the same time, at the microeconomic level, the model matches a wide set of stylized facts concerning firm dynamics (including right-skewed distribution of firm sizes, fat-tailed distributions of firm growth rates, wide and persistent productivity differences across firms, lumpy investment dynamics).

Figure 1. Profit-led investment scenario. Average GDP growth rate as a function of the mark-up rate (confidence bands in gray)

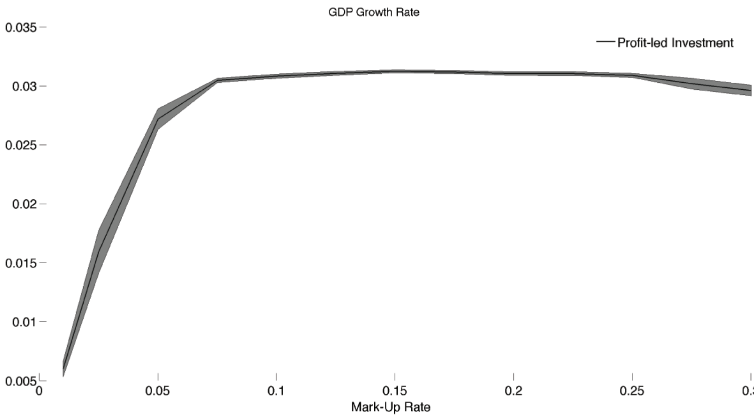


Figure 2. Profit-led investment scenario. Average unemployment rate as a function of the mark-up rate (confidence bands in gray)

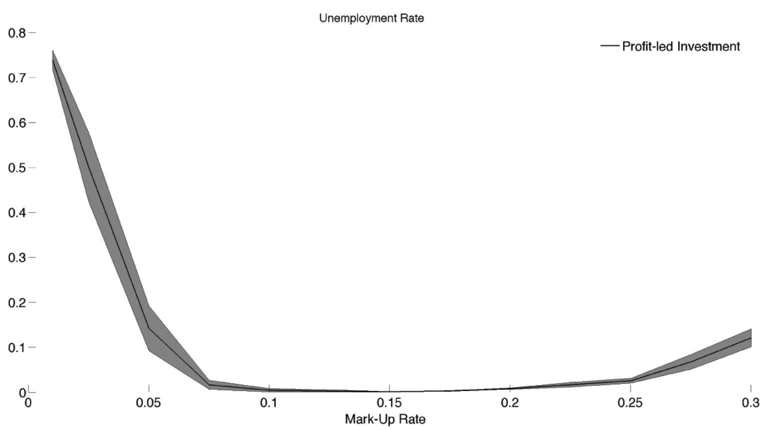


Figure 3. Profit-led investment scenario. Standard deviation of GDP growth rate as a function of the mark-up rate (confidence bands in gray)

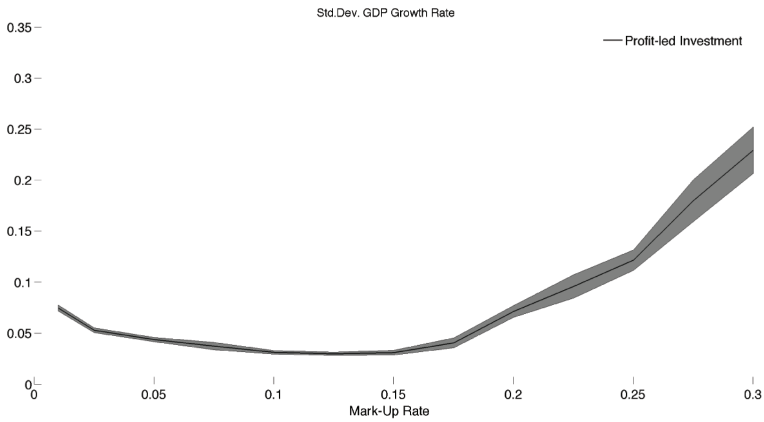
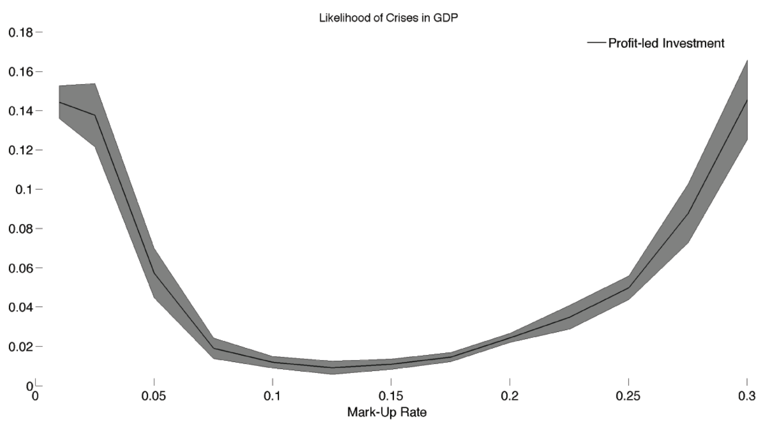


Figure 4. Profit-led investment scenario. Likelihood of crises in GDP as a function of the mark-up rate (95% confidence bands in gray)



Note: Crises are defined as time periods with growth rates lower than –3%.

Encouraged by the empirical performance of the model, in the next sections we turn to explore the behavior of the economy under different income distribution hypotheses, and under different regimes of adjustment of nominal wages to variations in unemployment.

2.1. Income distribution and growth regimes

Real wages in the model have a dual role. On the one hand, they affect the level of firms' costs and therefore profits, thus determining the ability of firms to survive and to finance production and investment expenditures. On the other hand, real wages determine aggregate consumption and via the latter they impact on firms' investment decisions. The first effect is the one already emphasized by Neo-Classical economists as well as by the works in the New Keynesian literature with financial market imperfections (see Greenwald and Stiglitz, 1993), and it should have a detrimental effect on growth and employment. In contrast, the second effect, as emphasized by Keynes in his General Theory (see e.g. also Howitt, 1986, for a discussion), should affect aggregate demand and thus output and unemployment. In what follows, we study the behavior of macro-variables in the model under different levels of the wage-share in the economy. We change the dynamics of real wages (and of profits) by tuning up and down the level of the initial mark-up rate of consumption good firms. The latter indeed determines consumption-good prices in the model, and therefore it determines how real income is divided between profits and wages (see also Section 1.3). Moreover, we perform the aforementioned experiment on income distributions under different regimes of firms' investment behavior. In the first scenario, desired investment is "profit-led". More precisely, desired investment in new productive capacity is determined only by the degree of financial robustness of the firm, proxied by its stock of net liquid assets. In the second scenario, instead, desired expansionary investment is "demand-led", *i.e.* driven by expectations about future consumption demand (see also Section 1.3). Finally, notice that in both scenarios desired investment plans can be constrained by the availability of internal and external finance, respectively determined by the level of past profits, and by the external credit ceilings.

Let us start by discussing the result of the experiment under the "profit-led" investment scenario. Figures 1 and 2, show average growth and unemployment in relation to the (initial) mark-up rate. Both figures indicate that very low levels of the mark-up rate (and, thus, very high levels of real wages) have a detrimental effect on average growth and average unemployment. These outcomes correspond to the "Neo-Classical" result, according to which high real wages have a negative effect on long-run growth and employment. More specifically, a high level of real wages implies lower firms' profits, and therefore a lower incentive of firms to invest in new productive capacities, and thus to absorb the existing labor supply. In addition, also productivity growth is lower, as the introduction of more productive machines is limited to the substitution of the existing capital stock.

Furthermore, Figures 3 and 4 show that volatility and (even more so) the likelihood of crises are high when the mark-up rate is very low. In this profit-led investment scenario, short-run business fluctuations are generated by a Goodwin-type dynamics (Goodwin, 1967). An increase in profits promotes *via* investment an increase in both production and productivity (due to introduction of new and more productive machines). Productivity gains however imply also an increase in real wages and a reduction in the profit rate thereby creating the conditions for a fall in economic activity. The lower is the mark-up rate the stronger is the above described predator-prey dynamics. Accordingly, both GDP volatility and the likelihood of crises turns out to be high at low levels of the mark-up rate.

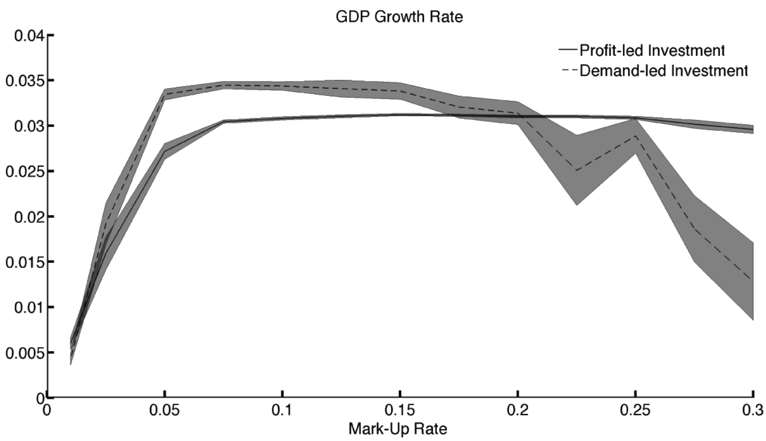
What happens if one increases the mark-up and therefore decreases the level of real wages in this scenario? As Figures 1-4 show quite starkly, increasing the mark-up has the effect of increasing growth and reducing unemployment. The economy enters into a "Classical Growth Regime" (see Boyer, 1988) where productivity increases are able to promote profits, hence investment and effective demand, which enhance employment. In addition, in this growth scenario, the above-described dynamics between profits and wages is dampened by a lower mark-up rate because the process of workers' appropriation of productivity gains through real wages is weakened. As a result, both GDP volatility and the likelihood of crises are lowered. In particular, notice that the economy is characterized by a wide region of high and stable growth (corresponding to levels of mark-up between 0.075 and 0.15) wherein, on one side, average growth is maximized and, on the other side, unemployment, volatility and the probability of crises are zero.

As real wages are reduced further ($\bar{\mu}(0) > 0.15$), however, the economy enters a region where high growth are associated to wild fluctuations, higher incidence of crises and positive unemployment. This occurs because the aforementioned "Classical" engine of growth is hampered by the credit constraints imposed on firms investment. More precisely, in this region the strong incentives of firms to invest in large productive capacities imply also high debt accumulation by firms. As time goes by, increasingly high debt burdens erode profits, thereby creating the conditions for the incoming recession. Accordingly, both volatility and the incidence of crises become higher.

How do the above results change when firm investment is "demand-led"? Figures 5, 6 and 7 compare the dynamics of macro-variables in the profit-led vs. the demand-led scenarios, always as a function of the mark-up rate. As in the profit-led scenario, low levels of the mark-up rate are associated with low growth and high unemployment and crisis incidence also in the demand-led regime. The mechanisms generating this result are however completely different between the two scenarios. In the profit-led regime, low growth and high unemployment rates are determined by the low incentive of firms to invest in new productive capacities. In contrast, in the demand-led scenario, the incentive to invest is high when the mark-up rate is low, because wages and thus expectations of consumption demand are very high. However, firms desired investment plans are constrained by the availability of internal and external financing, which are on average lower in presence of a low profits. Moreover, a higher incentive to invest induce a dynamics of debt accumulation similar to the one analyzed in the case of the high-growth profit-led regime, thereby causing higher volatility (not shown) and crises incidence (*cf.* Figure 7). In light of the above, it comes as no surprise that a small reduction in the (already high) level of real wages is able to restore growth and reduce volatility and unemployment (see Figure 5).

As the distribution of income between wages and profits gets more balanced ($0.05 \leq \bar{\mu}(0) \leq 0.15$), the economy enters a regime characterized high average growth rates and low unemployment and volatility rates. Following the taxonomy suggested in Boyer (1988), we shall label this a "Fordist Growth Regime", roughly matching the institutional conditions characterizing advanced economies in the post-WWII period. Such a regime associates investment led by demand expectations with a balanced sharing of productivity gains between workers and firms. In that, improvements in productivity lead to significant increases in wages, thereby rising aggregate demand *both* via consumption *and* investment by firms. Notice that this regime is characterized by significantly higher average growth rates than the Classical one discussed before. In turn, this occurs because high levels of consumption demand now reinforce *via* expectations aggregate investment, rather than crowding it out as in the Classical regime.

Figure 5. Average GDP growth rate as a function of the mark-up rate (confidence bands in gray)



The high dependence of growth on expectations about consumption demand also implies the emergence of a wide region wherein real wages and unemployment are inversely (rather than directly) related. More precisely, in the demand-led scenario, further increasing the profits rate in the economy result into a significant increase of unemployment and of the incidence of crises (see Figures 6 and 7). Moreover, very high levels of the mark-up rate ($\bar{\mu}(0) > 0.20$, *cf.* Figure 5) locks the economy into a low-growth trajectory (similar to what very high real wages do). This is explained by the fact (see also Dosi *et al.*, 2012 for a more detailed discussion) that in presence of very low levels of expected demand, firms have low incentive to undertake investments into new and productive machines. This hampers the growth of productivity and therefore the long-term growth prospects of the economy.

Figure 6. Average unemployment rate as a function of the mark-up rate (confidence bands in gray)

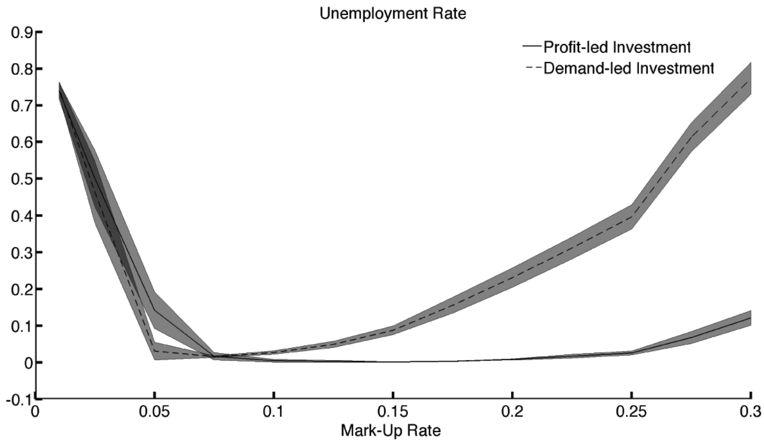
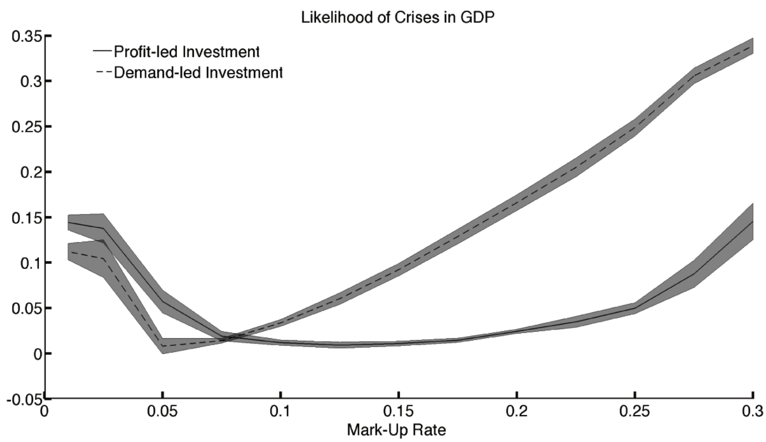


Figure 7. Likelihood of crises in GDP as a function of the mark-up rate (95% confidence bands in gray)



Note: Crises are defined as time periods with growth rates lower than -3%.

Let us now summarize the different results of the section as well as provide some implications for theory. First, our results indicate that the economy is not always characterized by an inverse relation between real wages and employment. The presence of such a relation is indeed closely associated to a profit-led investment behavior by firms. In contrast, under regimes where investment is driven by firms expectations such a relation is non-linear.

In particular, there exists a threshold (corresponding to a low level of mark-ups) above which unemployment decreases (rather than increasing) with the level of the real wage in the economy. Second, in line with Boyer (1988) and Freeman and Perez (1988), our experiments indicate that endogenous innovation is not able by itself to guarantee high and stable growth without the presence of specific institutional conditions determining investment behavior and the distribution of income between wages and profits. However, differently from that literature our analysis shows that independently from firm investment behavior the presence of high growth together with low levels of volatility and unemployment, always requires a balance in the distribution of income between profits and wages. When such a condition is not fulfilled the economy gets either into a low and highly volatile growth trap (low mark-up rate in either investment scenario, high mark-up rate in the demand-led scenario), or into a situation where high growth rates are associated with significant volatility and high unemployment rates (high mark-up rate in the profit-led scenario). In the next section we investigate the robustness of the above results in presence of wage flexibility to unemployment variations.

2.2. Nominal wages flexibility, growth and unemployment

In the previous section we have shown that the distribution of income between profits and wages crucially affects the properties of the macroeconomic dynamics and the rate of unemployment in the economy. However, these results were obtained assuming that in the labor market wages were only a function of labor productivity. In this section, we remove this hypothesis, and we allow for nominal wages that depend also on changes of the unemployment rate. More precisely, we select different levels of the mark-up rate and we run different Montecarlo experiments by varying the level of the parameter ψ_3 in Equation (6), *i.e.* the parameter tuning the response of nominal wages to relative changes in the unemployment rate U. We perform the experiment for the scenario wherein investment is profit-led and for the one in which investment is demand-led. In this way, we study how all the results discussed in the previous section change under higher levels of nominal wages flexibility to unemployment variations.

Table 2 and 3 present the results of the above described experiments. Notice that the levels of mark-ups presented in the tables correspond to different growth regimes (see previous section) . Let us start by discussing the results for the scenario with profit-led investment and low mark-up rate ($\bar{\mu}(0) = 0.01$). In this region of the parameters' space, the high-level of real wages depresses profits and therefore the incentive to invest. One could therefore expect a positive effect of downward wage flexibility on the economy. This is not the case, however. Introducing wage flexibility does not affect growth and it does not reduce unemployment (see Table 2, first and second column). Furthermore, increasing nominal wage-flexibility exacerbates the volatility of the economy. More precisely, increasing the sensitivity of wages to unemployment implies a significant increase both in the standard deviation of GDP growth and in the likelihood of crises (Table 2, columns 3 and 4). The same occurs for the standard deviation of the inflation rate (column 6). The significant increase in volatility is explained by the fact that wage flexibility strengthen the competition between profits and wages, which is one of the main mechanisms generating fluctuations in the profit-led investment scenario.

The situation dramatically changes when we consider higher levels of the mark-up rate. Take for instance the scenario with mark-up rate $\bar{\mu}(0) = 0.05$. There, the distribution of income between profits and wages ensures that firms have enough incentive to invest and therefore growth unfolds according to the "Classical" mechanisms described in the previous section. In that growth regime, wage-flexibility is able to significantly reduce unemployment. In addition, the increases in growth and inflation volatility are small and in most cases not significant. Notice, however, that wage flexibility also induces a small reduction in the average growth rate of the economy. In turn, this outcome is explained by the fact that wage flexibility reduce firm incentives to replace their machines, thus curbing productivity growth.

All the patterns described above emerge even more starkly in the high mark-up scenario ($\bar{\mu}(0) = 0.03$ see Table 2). More precisely, wage-flexibility significantly reduces unemployment and brings a small reduction in average growth. However, and differently from the case with $\bar{\mu}(0) = 0.05$ discussed before, now an increase in wage-flexibility results in a reduction both of the standard deviation of GDP growth and of the likelihood of crises.

Table 2. Profit-led investment scenario. Experiments with different degrees of nominal wages flexibility to unemployment variations (Monte-Carlo simulations standard errors in brackets)

	Avg.GDP growth rate	Avg.unempl. rate	St.dev. GDP growth rate	Avg.likel. GDP crises	Avg. infl. rate	St.dev. infl. rate
Mark-Up Rate 0.01						
$\psi_3 = 0.0$	0.0060 (0.0003)	0.7386 (0.0108)	0.0746 (0.0015)	0.1442 (0.0042)	0.0001 (0.0000)	0.0591 (0.0017)
$\psi_3 = 0.2$	0.0068 (0.0003)	0.7113 (0.0110)	0.0746 (0.0014)	0.1963 (0.0057)	0.0005 (0.0000)	0.0673 (0.0019)
$\psi_3 = 0.4$	0.0086 (0.0004)	0.6300 (0.0206)	0.0819 (0.0043)	0.2023 (0.0055)	0.0010 (0.0001)	0.0732 (0.0016)
$\psi_3 = 0.6$	0.0074 (0.0004)	0.6703 (0.0193)	0.1294 (0.0350)	0.1816 (0.0089)	0.0009 (0.0002)	0.0723 (0.0022)
$\psi_3 = 0.8$	0.0040 (0.0010)	0.7124 (0.0174)	0.4883 (0.1149)	0.2106 (0.0134)	0.0006 (0.0002)	0.0708 (0.0022)
Mark-Up Rate 0.05						
$\psi_3 = 0.0$	0.0270 (0.0005)	0.1463 (0.0287)	0.0456 (0.0014)	0.0582 (0.0066)	0.0001 (0.0000)	0.0353 (0.0018)
$\psi_3 = 0.2$	0.0266 (0.0005)	0.1098 (0.0248)	0.0491 (0.0014)	0.0733 (0.0067)	0.0019 (0.0001)	0.0413 (0.0017)
$\psi_3 = 0.4$	0.0247 (0.0007)	0.0694 (0.0187)	0.0562 (0.0021)	0.0770 (0.0054)	0.0035 (0.0002)	0.0534 (0.0028)
$\psi_3 = 0.6$	0.0198 (0.0011)	0.0576 (0.0188)	0.0600 (0.0059)	0.0752 (0.0067)	0.0053 (0.0004)	0.0531 (0.0032)
$\psi_3 = 0.8$	0.0177 (0.0013)	0.0391 (0.0134)	0.0548 (0.0023)	0.0714 (0.0048)	0.0069 (0.0005)	0.0565 (0.0040)
Mark-Up Rate 0.20						
$\psi_3 = 0.0$	0.0309 (0.0001)	0.0032 (0.0010)	0.0332 (0.0015)	0.0118 (0.0014)	0.0000 (0.0000)	0.0147 (0.0007)
$\psi_3 = 0.2$	0.0311 (0.0001)	0.0019 (0.0003)	0.0470 (0.0035)	0.0261 (0.0024)	0.0017 (0.0002)	0.0288 (0.0019)
$\psi_3 = 0.4$	0.0278 (0.0007)	0.0028 (0.0006)	0.0574 (0.0047)	0.0414 (0.0050)	0.0027 (0.0003)	0.0389 (0.0035)
$\psi_3 = 0.6$	0.0250 (0.0012)	0.0024 (0.0003)	0.0526 (0.0041)	0.0362 (0.0046)	0.0033 (0.0004)	0.0347 (0.0028)
$\psi_3 = 0.8$	0.0207 (0.0015)	0.0030 (0.0008)	0.0476 (0.0035)	0.0334 (0.0041)	0.0054 (0.0007)	0.0373 (0.0031)
Mark-Up Rate 0.30						
$\psi_3 = 0.0$	0.0295 (0.0002)	0.1318 (0.0097)	0.2360 (0.0091)	0.1596 (0.0089)	-0.0001 (0.0000)	0.0243 (0.0008)
$\psi_3 = 0.2$	0.0307 (0.0002)	0.0712 (0.0113)	0.1849 (0.0108)	0.1311 (0.0080)	0.0014 (0.0002)	0.0398 (0.0025)
$\psi_3 = 0.4$	0.0263 (0.0008)	0.0532 (0.0111)	0.1707 (0.0100)	0.1431 (0.0072)	0.0033 (0.0004)	0.0593 (0.0042)
$\psi_3 = 0.6$	0.0222 (0.0012)	0.0521 (0.0112)	0.1605 (0.0111)	0.1352 (0.0076)	0.0049 (0.0005)	0.0627 (0.0047)
$\psi_3 = 0.8$	0.0173 (0.0014)	0.0519 (0.0112)	0.1435 (0.0124)	0.1168 (0.0088)	0.0061 (0.0007)	0.0572 (0.0043)

Table 3. Demand-led investment scenario. Experiments with different degrees of nominal wages flexibility to unemployment variations (Monte-Carlo simulations standard errors in brackets)

	Avg.GDP growth rate	Avg.unempl. rate	St.dev. GDP growth rate	Avg.likel. GDP crises	Avg. infl. rate	St.dev. infl. rate
Mark-Up Rate 0.01						
$\psi_3 = 0.0$	0.0042 (0.0003)	0.7447 (0.0092)	0.1952 (0.0237)	0.1121 (0.0045)	0.0001 (0.0000)	0.0687 (0.0026)
$\psi_3 = 0.2$	0.0042 (0.0006)	0.7356 (0.0135)	0.2080 (0.0193)	0.1206 (0.0048)	0.0005 (0.0001)	0.0767 (0.0034)
$\psi_3 = 0.4$	0.0045 (0.0004)	0.7364 (0.0115)	0.1881 (0.0226)	0.1222 (0.0050)	0.0007 (0.0001)	0.0772 (0.0037)
$\psi_3 = 0.6$	0.0040 (0.0006)	0.7432 (0.0130)	0.1997 (0.0208)	0.1336 (0.0051)	0.0011 (0.0001)	0.0766 (0.0022)
$\psi_3 = 0.8$	0.0038 (0.0007)	0.7629 (0.0136)	0.2884 (0.0697)	0.1465 (0.0059)	0.0012 (0.0002)	0.0824 (0.0034)
Mark-Up Rate 0.05						
$\psi_3 = 0.0$	0.0334 (0.0003)	0.0307 (0.0119)	0.0293 (0.0008)	0.0080 (0.0043)	0.0002 (0.0000)	0.0137 (0.0010)
$\psi_3 = 0.2$	0.0333 (0.0002)	0.0318 (0.0080)	0.0291 (0.0008)	0.0092 (0.0033)	0.0025 (0.0001)	0.0133 (0.0007)
$\psi_3 = 0.4$	0.0330 (0.0003)	0.0509 (0.0150)	0.0304 (0.0012)	0.0169 (0.0058)	0.0048 (0.0001)	0.0170 (0.0013)
$\psi_3 = 0.6$	0.0335 (0.0002)	0.0285 (0.0086)	0.0285 (0.0008)	0.0080 (0.0028)	0.0074 (0.0001)	0.0175 (0.0009)
$\psi_3 = 0.8$	0.0331 (0.0003)	0.0540 (0.0178)	0.0301 (0.0012)	0.0151 (0.0053)	0.0093 (0.0003)	0.0218 (0.0014)
Mark-Up Rate 0.20						
$\psi_3 = 0.0$	0.0314 (0.0006)	0.2307 (0.0132)	0.7430 (0.0680)	0.1660 (0.0043)	0.0000 (0.0001)	0.0858 (0.0033)
$\psi_3 = 0.2$	0.0308 (0.0006)	0.2356 (0.0124)	0.7018 (0.0751)	0.1683 (0.0038)	0.0016 (0.0001)	0.0826 (0.0031)
$\psi_3 = 0.4$	0.0296 (0.0012)	0.2289 (0.0123)	0.8236 (0.0914)	0.1586 (0.0042)	0.0026 (0.0001)	0.1004 (0.0047)
$\psi_3 = 0.6$	0.0280 (0.0015)	0.2374 (0.0173)	0.8185 (0.1028)	0.1612 (0.0055)	0.0037 (0.0002)	0.1007 (0.0036)
$\psi_3 = 0.8$	0.0289 (0.0015)	0.2356 (0.0105)	0.6773 (0.0632)	0.1624 (0.0037)	0.0046 (0.0003)	0.1132 (0.0046)
Mark-Up Rate 0.30						
$\psi_3 = 0.0$	0.0128 (0.0021)	0.7733 (0.0216)	1.7748 (0.0998)	0.3388 (0.0042)	-0.0003 (0.0001)	0.1810 (0.0088)
$\psi_3 = 0.2$	0.0128 (0.0020)	0.8144 (0.0226)	1.8682 (0.0831)	0.3416 (0.0042)	-0.0003 (0.0001)	0.1835 (0.0091)
$\psi_3 = 0.4$	0.0128 (0.0017)	0.7836 (0.0215)	2.0117 (0.0765)	0.3390 (0.0037)	-0.0006 (0.0002)	0.1884 (0.0102)
$\psi_3 = 0.6$	0.0125 (0.0019)	0.8259 (0.0237)	1.9303 (0.0869)	0.3401 (0.0044)	-0.0004 (0.0001)	0.1705 (0.0100)
$\psi_3 = 0.8$	0.0136 (0.0020)	0.8018 (0.0249)	1.8278 (0.0794)	0.3384 (0.0045)	-0.0004 (0.0001)	0.1816 (0.0103)

Overall, the results of the above experiments indicate that the flexibility of nominal wages to unemployment variations can stabilize the economy and reduce unemployment, when growth is driven by a Classical mechanism, linking productivity growth to investment determined by firms' profits. Moreover, the role of wage-flexibility as a stabilizing device is stronger the more the distribution of income is biased towards profits.

How does the above picture change when we turn to the demand-led scenario? Notice that in that scenario, investment is tied to the level of real wages via expectations of consumption demand. In turn, such a Keynesian expectation effect can be stronger than the cost-saving effect when wages are reduced following an increase in unemployment. Accordingly, higher wage-flexibility to unemployment can de-stabilize (rather than stabilize) investment and aggregate demand (see Keynes, 1936; Howitt, 1986; Amendola *et al.*, 2004). A rise in the degree of wage-flexibility does not provoke a reduction in average unemployment in any of the scenarios considered in the table. Consider for instance the experiments with high mark-up rates (respectively $\bar{\mu}(0) = 0.20$ and $\bar{\mu}(0) = 0.30$). Except for inflation volatility, none of the macroeconomic statistics reported in Table 3 display a statistically significant variation when we increase the degree of wage-flexibility.

One could expect the above results to be different in the scenarios with low mark-up rates ($\bar{\mu}(0) = 0.05$ and $\bar{\mu}(0) = 0.01$). There, real wages and the level of consumption demand are expected to be high, and therefore the cost-saving effect associated with wage reductions in presence of unemployment could dominate the Keynes' expectation effect. Once again, the results in Table 3 show that this does not happen. Higher degrees of wage-flexibility do not lead to unemployment reductions. On the contrary, unemployment increases (even though the variation is not statistically significant). In addition, high levels of wage-flexibility lead to an increase in the likelihood of crises when the level of the mark-up is very low ($\bar{\mu}(0) = 0.01$).

3. Concluding remarks

Using the Keynes+Schumpeter (K+S) agent-based model developed in Dosi *et al.* (2010, 2012) we investigated the characteristics of growth regimes emerging different rules for firms' investment decisions and different distributions of income between profits and wages. We studied the aggregate dynamics when investment is determined by the stock of liquid assets of the firm ("profit-led investment scenario") and when it is determined according to expected demand ("demand-led investment scenario"). We showed that, independently from the investment scenario, the occurrence of stable growth associated with low unemployment requires a balanced distribution of income between profits and wages. Moreover, we showed that in the demand-led scenario the economy is characterized by a non-linear relation between the level of real wages and the unemployment rate. In particular, we find a threshold below which further reductions in the average real wage are associated with an increase (rather than an decrease) in unemployment. Finally, we investigated whether the introduction of increasing degrees of nominal wage-flexibility to unemployment variations are able to restore growth and curb unemployment. We showed that this is indeed the case when investment is profit-led and the mark-up rate in the economy is able to ensure positive growth. In contrast, in the demand-led scenario wage-flexibility never brings reductions in unemployment and in some cases can also result into a higher incidence of crises in the economy. Accordingly, our results cast doubts on the effectiveness of policies promoting wage-flexibility when investment is related to expectations about future consumption demand.

The present work could be extended in at least two directions. First, one could extend the above analysis by considering several open economies, whose ability to export may depend on the level of wages, and there study how the results of this work may change in that context. Second, we considered only a very stylized representation of labor market interactions. However, one could easily extend the above framework to introduce a full-fledged analysis of the labor market, e.g. like in Fagiolo *et al.* (2004); Dawid *et al.* (2011) and study the inter-play between this market and the processes of technical change and of the determination of income-distribution in the economy.

References

Aglietta M., 1979. A Theory of Capitalist Regulation: *The US Experience*. New Left Books.

Amendola M., J. Gaffard and F. Saraceno, 2004. "Wage exibility and unemployment: the Keynesian perspective revisited". *Scottish Journal of Political Economy*. 515, 654–674.

Boyer R., 1988. "Formalizing Growth Regimes". In *Technical Change and Economic Theory*, G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete (eds). Francis Pinter. pp. 609–629.

Chiaromonte F., G. Dosi and L. Orsenigo, 2000. "Innovative learning and institutions in the process of development: on the microfoundations of growth regimes". *In Innovation, organization and economic dynamics: selected essays*, G. Dosi (ed). Edward Elgar Publishing.

Coriat B. and G. Dosi, 2000. "The Institutional Embeddedness of Economic Change. An Appraisal of the'Evolutionary'and'Regulationist'Research Programmes". *In Innovation, organization and economic dynamics: selected essays*, G. Dosi (ed). Edward Elgar Publishing.

Dawid H., S. Gemkow, S. van der Hoog and M. Neugart, 2011. "The Eurace@Unibi Model: An Agent-Based Macroeconomic Model for Economic Policy Analysis". *Working paper*. Universität Bielefeld.

Del Boca A., M. Galeotti, C. P. Himmelberg and P. Rota, 2008. "Investment and Time to Plan and Build: A Comparison of Structures vs. Equipment in A Panel of Italian Firms". *Journal of the European Economic Association*. 6, 864-889.

Delli Gatti D., C. Di Guilmi, E. Gaffeo, G. Giulioni, M. Gallegati and A. Palestrini, 2005. "A New Approach to Business Fluctuations: Heterogeneous Interacting Agents, Scaling Laws and Financial Fragility". *Journal of Economic Behavior and Organization*. 56, 489–512.

Doms M. and T. Dunne, 1998. "Capital Adjustment Patterns in Manufacturing Plants". *Review Economic Dynamics*. 1, 409-29.

Dosi G., G. Fagiolo, M. Napoletano and A. Roventini, 2012. "Income Distribution, Credit and Fiscal Policies in an Agent-Based Keynesian Model". *LEM Papers Series* 2012/03. Laboratory of Economics and Management (LEM), Sant'Anna School of Advanced Studies, Pisa, Italy.

Dosi G., G. Fagiolo and A. Roventini, 2006. "An Evolutionary Model of Endogenous Business Cycles". *Computational Economics*. 27, 3–34.

Dosi G., G. Fagiolo, and A Roventini., 2010. Schumpeter Meeting Keynes: A Policy-Friendly Model of Endogenous Growth and Business Cycles. Journal of Economic Dynamics & Control. 34, 1748–1767.

Eisner R., 1972. "Components of Capital Expenditures: Replacement and Modernization versus Expansion". *The Review of Economics and Statistics*. 54, 297-305.

Fabiani S., M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Mathä, R. Sabbatini, H. Stahl and A. Stokman, 2006. “What Firms' Surveys Tell Us about Price-Setting Behavior in the Euro Area”. *International Journal of Central Banking*. 2, 3–47.

Fagiolo G., G. Dosi and R. Gabriele, 2004. “Matching, Bargaining, and Wage Setting in an Evolutionary Model of Labor Market and Output Dynamics”. *Advances in Complex Systems*. 14, 237–73.

Feldstein M. and D. Foot, 1971. “The Other Half of Gross Investment: Replacement and Modernization Expenditures”. *The Review of Economics and Statistics*. 53, 49-58.

Freeman C. and C. Perez, 1988. “Structural Crises of Adjustment, Business Cycles and Investment Behaviour”. In *Technical Change and Economic Theory*, G. Dosi, C. Freeman, R. Nelson, G. Silverberg and L. Soete (eds). Francis Pinter. pp. 38-66.

Goodwin R., 1967. *A growth cycle*. *In Socialism, capitalism and economic growth*, C. Feinstein (ed). Cambridge University Press.

Goolsbee A., 1998. “The Business Cycle, Financial Performance, and the Retirement of Capital Goods”. *Review of Economic Dynamics*, 1(2), pp. 474-96.

Greenwald B. and J. Stiglitz, 1993. “Financial Market Imperfections and Business Cycles”. *Quarterly Journal of Economics*. 108, 77–114.

Howitt P., 1986. “Wage exibility and employment”. *Eastern Economic Journal*. 123, 237–242.

Keynes J. M., 1936. *The General Theory of Employment, Interest, and Money*. New York, Prometheus Books.

Lordon F., 1991. “Théorie de la croissance : quelques développements récents”. *Revue de l'OFCE*. 37, 191–243.

Rotemberg J., 2008. “Behavioral Aspects of Price Setting, and Their Policy Implications”. *Working Paper* 13754. NBER.

Smets F. and R. Wouters, 2007. “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach”. *American Economic Review*. 97, 586–606.

PRODUCTION PROCESS HETEROGENEITY, TIME TO BUILD, AND MACROECONOMIC PERFORMANCE

Mario Amendola
University of Rome "La Sapienza"
Jean-Luc Gaffard
OFCE and SKEMA Business School
Francesco Saraceno
OFCE

This paper describes the out-of-equilibrium approach to the analysis of economic processes. We argue that such an approach is adapted to study qualitative (or structural) changes, like technical progress or changes in preferences. Truly sequential analyses manage to capture the essential features of qualitative change. In particular, we show how this approach shifts the focus from the issue of optimality to the one of *viability* of the processes of change. The objective of the paper is, first, to highlight the analytical elements of an out-of-equilibrium approach, so as to serve as a guide for the construction of this type of models; second to show, how this analysis allows to see controversial phenomena, like for example the debate on wage rigidity or the productivity paradox in a new and different light ; third to identify the real causes of the on-going crisis.

Keywords: Out-of-equilibrium, Structural change, Credit constraints, Agent-based models, Adaptive expectations, Time-to-build, Sequential analysis, Macroeconomic policy, Productivity paradox, Wage rigidity

The purpose of this article is twofold: first, to highlight the interest of out-of-equilibrium analysis, that can as a first approximation be defined as the construction of models that allow dealing with phenomena that are in the nature of *qualitative change*. Changes, in other words, that entail modifications in the structure of the economy, and that are in the nature of *processes* that take place over time. An equilibrium analysis that, by its very nature, is limited to the comparison of equilibrium states (or paths), is not apt to the analysis of what happens during the process triggered by qualitative change. Out-of-equilibrium analysis allows studying the process of change as a sequence of constrained choices: which, as we shall see, shifts the focus from the question of *optimality* of the path followed by an economy to its *viability*.

The second purpose of the article is to detail the logical structure of out-of-equilibrium models, and to emphasize the necessary analytical departures from the simplifying hypotheses used in standard equilibrium analysis. This article should hence serve also as a guide for the construction of this type of models.

The paper is structured as follows. Section 1 discusses the out-of-equilibrium approach, emphasizing its departure from standard equilibrium growth theory and its methodological pillars. Section 2 expounds the logical structure of a typical out-of-equilibrium model, based on a sequential structure and on irreversibility in production and in decision making. Section 3 shows how this approach allows shedding a different light on some long standing controversies such as, e.g., the desirability of wage rigidity or the so-called "productivity paradox". It will be shown that the specific results obtained hint at more general conclusions that are a guide line for policy interventions quite different and sometimes opposite to those inspired by the prevailing equilibrium analytical approach. We conclude the paper by stressing the relevance of our approach for a better understanding of the current crisis and possible ways out of it (section 4). The analytical elements are finally discussed in an appendix.

1. The out-of-equilibrium approach

Most processes of economic change are not 'quantitative'—that is, a simple modification of the intensity of a given functioning of the economy—but 'qualitative' changes, that is, changes in the very way of functioning (like changes in technology, a speeding up of the growth rate, the introduction of new products, the entering of new markets, the irruption of new countries and new firms in the international trade, changes in the distribution of income, and so on). This means by definition the breaking of equilibrium. It implies the disruption of the existing productive structure, on which the behaviour of the economy depends, and the construction of a new and different one. This is a process that takes time, and the market economy is not necessarily self-adjusting. Co-ordination failures are unavoidable, because problems of co-ordination arise in the first place in the production process itself. These problems, on the other hand, extend to the whole economic system. New goods or techniques imply new types of production processes and new activities that, in turn, call for new forms of interaction among the existing agents and institutions, or even the appearance of new actors and institutions. The heterogeneity of the agents involved is an essential feature of processes of structural change, as we shall stress in what follows.

The *viability* of the process of change, and even its outcome, are not predetermined, but depend on the way the co-ordination problems, both at the micro and macro level, are dealt with. *Different outcomes* may in fact be associated, e.g., with a given technological advance, depending on the effective development of the process, *i.e.* in the way in which co-ordination is (or is not) restored. As a matter of fact, technical changes that potentially allow for substantial increases in productivity may actually result, at the end of the process, in a waste of productive resources.

The focus on processes of change has momentous analytical implications. In the first place, the usual distinction between a long-term, where equilibrium obtains, and a disequilibrium short term disappears[1]. A process is neither a short nor a long term: it is a sequence of linked disequilibria that shape the evolution of the process itself. This is no longer seen as a transition path between equilibrium positions. In an out-of-equilibrium perspective the point of arrival becomes blurred. It is lost to sight, not necessarily in the sense that it ceases to exist, but because everything that matters is inside the process.

The analysis of this process does not call for a traditional type of model, that is, a model capable of generating a 'solution' in the sense of a behavior of the economy characterized by certain

1. The standard view, which considers trend and cycle as unrelated phenomena, is misleading. "When we turn our attention to long-run questions, we aren't turning away from co-ordination and adjustment problems, we are simply looking at them from a different perspective" (Howitt, 1994, p. 765).

specific features (efficiency, optimality, and so forth). What becomes important, instead, is to follow the evolution of the economy, traced out step by step by the sequence of disequilibria. The essence of a thorough process is in its going on, that is, in its being viable. This calls for a monitoring of the process itself to bring to light its salient moments: which can only be achieved by means of numerical experiments, that is, by simulations that, under certain conditions (chosen so as to stress aspects relevant to the analysis) allow to unveil what happens "along the way".

In this light also the usual distinction between the terms 'exogenous' and 'endogenous' must be interpreted in a different way. "In a model there are variables and there are parameters, which reflect the existing constraints. In the standard analysis the constraints, which exist outside and above the economy and which determine its behaviour, are taken to be exogenous. But once we recognize that the time over which change takes place is an irreversible process that shapes the change itself, 'it is impossible to assume the constancy of anything *over time*}...The only truly exogenous factor is *whatever exists at a given moment of time,* as a heritage of the past' (Kaldor, 1985, p.61). While the standard approach focuses on the right place to draw the line between what should be taken as exogenous and what should be considered instead as endogenous in economic modelling—a line that moves according to what we want to be explained by the model—out of equilibrium the question is no longer that of drawing a line here or there but rather one of the time perspective adopted. Everything can be considered as given at a certain moment of time, while everything becomes endogenous over time" (Amendola and Gaffard, 1998, pp.32-3).

Out-of-equilibrium, oscillations no longer appear as short-term deviations due to demand shocks from a long-term trend determined beforehand by fundamentals alone. Focusing on coordination mechanisms implies to recognize that the short term actually determines what the long term will be, and that supply conditions and demand conditions interact with each other.

Finally, it must be stressed that different evolution paths can be associated in fact to given fundamentals, according to how the out-of-equilibrium process actually evolves, and the fundamentals themselves undergo a change during this process, given the very definition of qualitative change. The 'fundamentals', in other words, are no longer fundamental.

A framework that helps dealing with the co-ordination problems concerning the heterogeneity of the entities involved in out-of-equilibrium paths is Agent Based Modelling (ABM), that "seeks to model the process by which one among many possible futures is selected, rather than imposing constraints on the model that ensure only a single equilibrium outcome" (Mehrling, 2006, p. 77). With this kind of model, 'agent' refers broadly to an encapsulated collection of data and methods representing an entity residing in a computationally constructed world. Individual biological life forms, social groupings, institutions, and physical entities can all be represented as agents" (LeBaron and Tesfatsion, 2008, p. 246).

Our sequence analysis shares most of the properties of ABM modelling, namely, heterogeneity of agents, bounded rationality, and non-market clearing. Here, those are production processes of different ages, and incorporating different technologies, which are represented as agents. Economic dynamics is mainly driven by the evolution of the composition of these processes, that is, by the time structure of productive capacity.

2. The logical structure of the model

The main objective of out-of equilibrium models is to allow the study of processes of change. Standard equilibrium analysis is carried out by comparison of equilibria, be them points or steady state paths. Transitional dynamics are most of the time neglected because inherently short term phenomena, and because they are pre-determined from the beginning, and as such not particularly informative (for example, the saddle path adjustment). It was argued above that a meaningful analysis of the transition may add substantial information to our understanding of the economy, notably as regards the *viability* of the out-of-equilibrium path undertaken following a structural change of the economy. This section aims at discussing the building blocks of out-of-equilibrium models, without reference to any specific set of equations. Some examples may be found in the appendix.

Out-of-equilibrium analysis shows that once we release some simplifying assumptions of standard equilibrium theory, ongoing processes of change become interesting to study because, far from being pre-determined, they are shaped by the interaction of agent's behavior, institutional factors, and environmental characteristics.

The starting point is a standard general equilibrium model, with households and price-setting firms using labor and capital. Typical examples of this type of models may come from the New Keynesian literature (see e.g., Woodford, 2003).

It is important to notice that the use of capital is not strictly necessary, as it can be substituted by dated labor as was for the example the case in classical analysis (see the discussion in Garegnani, 1984).

A sequential analysis focuses both on the supply and on the demand side of the economy. As for the supply side, the standard hypothesis to be dropped is the perfect substitutability of factors (usually embedded in a Cobb-Douglas production function) and the instantaneous utilization of labor and capital. In out-of-equilibrium models, production takes time, and is characterized by complementarity rather than substitutability of the production factors. This complementarity can be modeled through the definition of a productive process as a scheme by which a flow of labor inputs is converted into a flow of output and the consideration of a construction period, with inputs but no final output, which is followed by a running-it period (Hicks, 1970; Hicks, 1973). It can also modeled by using a CES function with a sufficient degree of complementarity between capital/dated labor) and labor (the limit case would be a standard Leontief function).

Analytically, time-to-build and complementarity are both necessary, as they create sunk costs, and make choices at a certain moment in time depend on the stock of capital/dated labor available for the firm. Suppose for example that you had no complementarity. Then no matter what their past choices and the stock of capital/dated labor were, firms would always be able to choose the desired level of output through an appropriate choice of factor quantities.

A second simplifying assumption, rational expectations, also prevents fully-fledged transitional dynamics. In out-of-equilibrium analysis agents have bounded rationality, especially when facing complex environments. “Innovativeness raises uncertainties. The *future outcome* of an *innovative action* poses ambiguity: the law of 'unanticipated consequences' applies (Merton, 1936) entrepreneurs have to act on their 'animal spirits', as John Maynard Keynes (1936) put it; in the view of Friedrich Hayek (2002), innovations are launched first, the benefit and the cost are 'discovered' afterward" (Phelps, 2007, p. 544). A backward-looking component of expectation formation is necessary for reasons analogous to time-to-build, *i.e.* to create a link between past and current actions, and hence to link periods into a sequence. While there is some ground for rejecting backward looking behavior in equilibrium models, it is much more difficult to do so in out-of-equilibrium contexts, in which "knowledge of the model" is of little use, and in which at least short term fluctuations, cannot be properly predicted. In these situations, agents usually resort to "rules of thumb" that resemble the adaptive behavior embedded in out-of-equilibrium models (see for example Hommes, 1998).

A third important simplifying assumption of standard theory is market clearing. Instantaneous price adjustments rule out, by definition, the possibility of disequilibrium. In sequential analysis, this assumption is released in order to allow the emergence of quantity constraints (the short side rule), and undesired stocks (both real and financial). These stocks contribute to link the periods in a sequence. This does not mean, of course, that prices do not change: "The fix-price method is a disequilibrium method [...] If flow demand is less than flow supply, a stock will have to be carried over; we say here that it has to be carried over, for the alternative policy of cutting price so as to dispose of them within the current period is not seriously considered. (And is not that, very often, realistic?) [...] In describing this model as a fix-price model, it is not assumed that prices are unchanging over time, or from one single period to its successor; only that they do not necessarily change whenever there is demand-supply disequilibrium." (Hicks, 1956, p. 232).

Finally, out-of-equilibrium agents may be constrained, in their transactions, by financial resources availability, strictly relevant in a context where costs are dissociated in time from receipts. This can be obtained by introducing missing markets, or more simply, through cash-in-advance constraints. The first road is necessary if the focus is on the working of credit and financial markets. The out-of-equilibrium literature, so far, has been more concerned with a macroeconomic approach, and hence has made use of cash-in-advance constraint that emerges because markets open sequentially.

All the modifications to the standard approach listed above have been extensively used in the literature. As for price rigidity, the early reappraisal of Keynes's economics (Clower, 1965; Leijonhufvud, 1968) or temporary equilibrium models (Hicks, 1939; Malinvaud, 1977; Bénassy, 1982), assume that prices only adjust between periods. Nevertheless, by releasing the hypothesis of full rationality, out-of-equilibrium analysis has to deal with the appearance of unsold stocks. The New Keynesian literature (Clarida, Gali and Gertler, 1999; Woodford, 2003) also makes reference to a monopolistic competition sticky prices environment, emphasizing short run quantity adjustments in response to shocks, even if in a context in which fluctuations are exclusively technology driven. Time-to-build has also been extensively studied (Kydland and Prescott, 1982), usually (but not only) in RBC type models, although not with reference to fully vertically integrated production processes. Finally cash-in-advance and credit constraints are rather commonly studied in the mainstream literature.

Nevertheless, to the best of our knowledge, these hypotheses were never considered jointly, so that their potential to analyze sequential economies has not been fully exploited (Saraceno, 2004). It is easy to see why their interaction is relevant when we are interested in analyzing out-of-equilibrium dynamics: each period begins with state variables determined in the previous one, and with imbalances that constrain agents in their subsequent decisions. Expectations and the structure of productive capacity further link the periods in a sequence. As a consequence, it is impossible to consider each period in isolation, as for example in the temporary equilibrium literature. A shock (no matter of what type) disrupts the coordination between agents and between phases of production (construction and utilization) that characterizes the equilibrium. *Ex ante* disequilibria (*i.e.* inconsistency of agent's plans) within the period are eliminated by rationing and stock accumulation rather than by price adjustments. The "success" of the subsequent transition lies in the ability of the system in recovering coordination.

In order to embed the assumptions above into a coherent framework, out-of-equilibrium models are usually built into a sequence. The sequential opening of markets (for financial resources, for labor, and for goods) creates binding constraints. The sequence broadly speaking consists of three different moments (the equation numbers refer to the equations in the appendix):

(A) At the end of every period we have a number of state variables:

- The productive capacity, represented by a population of processes, or by a stock of dated labour (see eqs A.3 and A.4) he stocks (of goods or financial means) in the hands of agents (that appeared as rationing in the previous period (eq. A.8). A set of prices and wages
- Some stock variables as labor supply or total money/credit supply.

(B) At the junction between periods

- If imbalances in the labor and goods markets appeared, wages and prices change (eqs. A.6 or A.7).
- The productive capacity 'ages': every productive process becomes one year older.

(C) In the new period things happen in the following order:

- On the production side firms determine, based on expectations (eq. A.5), and the stocks left from the previous period, both the desired quantity to be produced and investment.
- In the next step the desired production is compared with the productive capacity. This may either result in a constraint, or in a decision to keep part of productive capacity idle (equivalently, scrapping of processes can occur).
- Once desired/feasible production is determined, firms can compute the wage bill necessary to carry on production, and investment. The difference between the desired wage fund and available internal resources gives the amount of external financing required by firms.
- The short side rule (eq. A.8). then determines (given the supply from helicopter money or the financial sector) the equilibrium quantity of financial resources. If the demand side is rationed, investment and production are affected, and this alters the structure of productive capacity.

— After taking into account all the constraints affecting desired production, we can compute labor demand. The labor market opens, and demand and supply are matched. If rationing occurs (eq. A.8)., either unemployment or non desired money balances by firms appear
— Production takes place, and wages are paid. This allows computing aggregate demand for the goods.
— The matching of demand and supply in the goods markets is the final step, and the short side rule (eq. A.8). determines unsold stocks or non-desired money holdings by households.

While the details (on utility, technology, rationing, financial markets) may change, most out-of-equilibrium models share a sequential structure of this type that allows analyzing at each step the emergence of constraints affecting the subsequent choices (a *constraint-decision-constraint* sequence). Laying down the sequence allows to realize that problems in the matching of demand and supply may arise because of a number of constraints. A firm may fail to meet the demand it expects because it has not enough productive capacity, or because it faces a human resources constraint, or again because it fails to raise the funding needed to pay for wages and investment. In other words, problem of coordination may arise for a number of reasons, and the smooth functioning of the economy along a regular path appears to be the exception rather than the norm.

As a consequence of a change in the environment (the appearance of a new and superior technology, the degree of extent of the market due to a change in the distribution of income or to the globalisation of the economy), firms have to adapt their productive capacity, in fact they have to adopt a new one adapted to the new environment. In any case, the new productive capacity must be built before being used. And most of the time, there is a divorce between the investment in terms of capacity and the investment at cost. Whatever the reason for changing, the new capacity requires a higher construction cost more than compensated by a lower utilisation cost. This inevitably creates distortions in the structure of productive capacity, which engender fluctuations in output and employment. As a matter of fact, in absence of a change in the external resources available for carrying production processes, the investment in capacity decreases and, mechanically, after a while, the final output decreases with the consequence of diminishing the level of employment as well as the level of productivity. This is the result of coordination failures due to both irreversibility of investment decisions and imperfection of market information. What really happens along the way is the consequence of the distortion in the structure of production capacity, which the main event associated with any structural change (and the main property of our model). It will depend on the reaction of agents, that is, the way in which firms reacts to successive market disequilibria and government reacts to global imbalances (unemployment, inflation pressures).

3. Macroeconomic controversies revisited

An out-of-equilibrium analytical perspective allows shedding a different light on issues that have been at the center of important debates in macroeconomics.

The prevailing policy consensus, reflecting the equilibrium view of the existence of unique attractors defined with respect to the properties of technological changes and/or other fundamentals, maintains that the long term must prevail over the short term, that the supply conditions are more important than the demand conditions.

In an out-of-equilibrium perspective—where short term oscillations appear no longer as deviations from a fixed trend but rather as the way in which a process of change takes shape and gets realised—the short term actually determines what the long term will be, supply conditions and demand conditions interact with each other and there is a strict relation between the distribution and the creation of wealth, that is, between equity and efficiency.

These are the general methodological conclusions that result from the analysis of some relevant theoretical issues and controversies, presented in the following sections: conclusions that provide a guide line for policy interventions quite different and sometimes opposite to those inspired by the prevailing equilibrium analytical approach. For all the controversies listed below, here we only want to give a sense of how the out-of-equilibrium approach allows gaining a different perspective from equilibrium analysis. While the structure of the models used broadly corresponds to what we outlined in the previous section, we refer the reader to the cited papers for details on the specific choices made in each of the models.

3.1. Wage rigidity and keynesian unemployment

In an equilibrium construct in which markets are complete, the price vector conveys all the information necessary to fully coordinate agent's decisions. In this framework, unemployment can only stem from nominal rigidities in the relevant market, namely the one for labor.

The focus on flexibility comes from Neo-classical and New Keynesian models which both explain involuntary unemployment as the result of real wage rigidity. The Neo-classical analysis also postulates a positive correlation between nominal and real wages (generally confirmed by empirical observations), so that any cut in money wages should result in a cut in real wages. As a consequence, money wage rigidities associated with some specific institutional rules would be responsible for involuntary unemployment and should be reduced. The New-Keynesian analysis focuses on the bargaining arrangements on the labor market. In particular, entrepreneurs would fix a real wage rate above the equilibrium one—the efficiency wage—in order to induce the workers to reveal their actual level of productivity Stiglitz (1987). Once again, better information would result in a lower real wage and a higher employment level.

According to Keynes, the persistence of unemployment is due to a fall in the marginal efficiency of capital, which is not compensated by an equivalent fall in the real interest rate. In other words, it is due to capital market imperfections. "It was Keynes' position that it is the failure of the incomplete market mechanism to reconcile the implied values of forward demand and supplies [...] that is the source of the trouble. Unemployment of labor and other resources is a derivative phenomenon" (Leijonhufvud, 1968, p. 276). Co-ordination failures at the system level rather than failures in the labor market are responsible for unemployment, which will therefore be involuntary in the strict sense. In this context, disequilibria on the labor market call for wage adjustments, but a fall in the money wage and in the price level, far from leading to a decrease in the real wage and a re-absorption of unemployment, results in a debt deflation process and a cumulative depression. The reason invoked by Keynes is that wage bargains between entrepreneurs and workers do not determine the real wage, and 'there may exist no expedient by which labour as a whole can reduce its *real* wage to a given figure by making revised *money* bargains with the entrepreneurs' (Keynes, 1936, p. 13).

If the "source of the trouble" does not lie in the labor market, whose disequilibrium is only a "derivative phenomenon", the hypothesis of fixed wages that stirred so much controversy acquires a very precise meaning. Keynes writes that "if money-wages were to fall without limit whenever there was a tendency for less than full employment, [...] there would be no resting place below full employment until either the rate of interest was incapable of falling further, or wages were zero. In fact, we must have some factor, the value of which in terms of money is, if not fixed, at least sticky, to give us any stability of values in a monetary system" (Keynes, 1936, p. 303). Keynes reverses the common wisdom on wage rigidity that, in his framework, becomes a necessary institutional feature to avoid the implosion of the system rather than a source of disequilibrium.

The crucial role of co-ordination, however, is better tackled by abandoning the equilibrium approach that also characterises Keynes' *General Theory*, and by seeing the working of the economy as a sequential out-of-equilibrium process (Amendola, Gaffard and Saraceno, 2004b; Saraceno, 2004). This is a complex process that originates on the production side of the economy, but involves all the relevant economic variables, as discussed in the previous section. Involuntary unemployment appears then as the result of a lack of co-ordination that emerges in the economy as a whole along the way, at each step, and cannot disappear simply by allowing price and wage flexibility. As a matter of fact, the standard treatment for taking care of unemployment, a reduction in wages, may result in a sequential process and in further distortions of productive capacity rather than in re-establishing co-ordination and hence re-absorbing unemployment.

Whatever the nature of the original shock experienced by the economy, it implies a distortion of its productive capacity, the dissociation in time of costs and proceeds, a reduction in the resources allocated to investments and hence in the demand for labor. Consequently income earned by the workers and their demand is reduced. An excess supply appears in the market for final output, and undesired real stocks cumulate. The production demanded no longer matches the existing productive capacity, and the firms scrap some processes in the utilisation phase. As production drops, the excess supply on the labor market persists. An excess of demand may then appear on the market for final output so that we can have an alternation of excesses of supply and demand that, by amplifying the distortion of productive capacity, result in ever increasing fluctuations of the economy. These can be either reduced or amplified according to whether co-ordination is successfully re-established or less. The prevailing wage regime has an essential role in this. Flexibility interpreted as quick adjustment feeds over-reactions in one or the other direction that result in a stronger alternation of excesses of supply and demand, and amplify the distortion of productive capacity. The relation between employment and flexibility then appears under a completely different light. Employment is in fact the result of a complex adjustment process and depends on how this process actually evolves. Price variability implies trading at false prices that create constraints and incentives, which in turn induce firms to take wrong production and investment decisions. Thus the problem lies not so much in the persistence of a wrong price than in the effects of an excessive variability of prices on the structure of productive capacity. In this case a certain wage rigidity prevents the fluctuations from becoming too strong and representing a threat to the viability of the economy. However, as the source of the problem is not in the labour market but in the conditions under which the investment creates, amplifies or eliminates distortions in the productive capacity of the economy, the issue of flexibility versus rigidity should be viewed in the light of how the investment issue is taken care of. If the latter is properly dealt with, it does not matter how flexible wages are: the wage–employment dilemma does not exist.

Prices and wages volatility induces quantity and hence investment volatility, that is, distortions in the age structure of productive capacity, which are responsible for stronger and stronger fluctuations of final output. By the way, in case of technological change, reducing the real wage would be at the opposite of what should happen as the consequence of a higher labour productivity level.

3.2. Financial constraints and monetary policy

The dissociation in time of inputs from output and costs from proceeds that characterizes every structural change calls for a central role of liquidity in ensuring the viability of the processes of change. Additional liquidity is required to build the bridge through time at the heart of the production process, destroyed by the distortion of productive capacity: and this can only be the outcome of an external intervention. Credit, or money creation, have the crucial role of re-establishing consistency over time of construction and utilization, investment and consumption, supply and demand. This is a general feature of out-of-equilibrium models, but was studied in detail in (Amendola and Gaffard, 1998; Amendola *et al.*, 2004b) where it is shown that the provision of liquidity must be articulated over time so as to properly interact with the modification in the structure of productive capacity which is taking place sequentially; which means, in particular, being harmonized with the time profile of internally generated financial resources during the process of change. Following a positive technology shock, although the natural rate of interest should finally increase, during the transition the lack of financial resources makes it necessary to conduct a loose monetary policy. It will be carried out through a reduction in the monetary interest rate, which will respond to the temporary reduction in the productivity growth rate of the economy[2]. This monetary policy allows minimizing both the output gap and the inflation rate over a given period of time, because it allows minimizing the distortions in the structure of productive capacity.

A policy dilemma is typical of economies that follow out-of-equilibrium paths (Amendola *et al.*, 2004b). Innovation requires "to transmute the capital that was embodied in the late stages of old production processes into capital embodied in the early stages of new processes, that is a disruption of other activities which is 'bound to be a strain'" (Hicks, 1989, p. 535). Then inflationary pres-

2. The reasons why an initial fall in productivity is usually associated with a process of structural change will be explored in detail in the next section.

sures (and/or deficits in the trade balance in open economies) necessarily appear "because the goods in which the wages (...) will be spent (...) cannot be provided out of the product of the labour which is newly employed, for that is not yet ready" (*ibid.*). Central banks can try to bring inflation back to the target level as soon as possible, with the consequence of exacerbating the initial negative impact of the shock on output and employment. They can, alternatively, decide an accommodating monetary policy bringing inflation back to the target more slowly with the consequence of simultaneously reducing inflation and unemployment (Amendola, Gaffard and Saraceno, 2004a). The latter policy consists in accepting a transitory inflation in the perspective of reducing unemployment. Later on, when and if co-ordination of the production process and the flow of internally generated financial resources are re-established, a restrictive monetary policy may be required to hamper the arising of inflationary pressures. These results hence call for a conduct of monetary policy substantially more articulated than a simple inflation-targeting rule, or even a Taylor rule. Monetary policy needs by its very nature to be discretionary, because it needs to accommodate the changing needs of the economy during the transition process.

3.3. Appropriating the potential gains of technology: The productivity paradox

The standard representation of production and technology, forces to consider unemployment as an equilibrium phenomenon. Its natural rate is determined by 'fundamentals' in a wide (*i.e.* not only referred to labor market features) sense. Unemployment is seen "as shaped by the structure of the economy rather than its recent history: technology, individual preferences, social values and institutions" (Phelps and Zoega, 1998, p. 783). Shocks—including temporary productivity shocks—may in the short run produce disequilibrium transitory effects on employment due to adjustments failures or lags; but in the long run only changes in the fundamentals—e.g. in the productivity growth rate—may explain changes in the natural rate. This, among other things, should account for the fact that "unemployment rates viewed over the very long run...appear to be un-trended in most nations, despite tremendous increases in productivity" (Blanchard and Katz,1997, p.56).

In the equilibrium framework technological advances should be instantaneously mapped into increases in productivity, and the only way to explain the 'productivity paradox'—a fall in productivity, we remember, notwithstanding the introduction of a superior technique in terms of production coefficients—is to assume adoption delays (Amendola, Gaffard and Saraceno, 2005). This happens because in the standard representation of technology, productivity is built into the production function as a given relationship between inputs and output. Such a representation needs an equilibrium framework, in which the ratios between the factors and output are constant and corresponding to those dictated by the production function coefficients.

If productivity is seen instead as the outcome of an out-of-equilibrium process, triggered by a technological shock, then the potential gains of a superior technology may only be appropriated if agents succeed in reshaping the productive capacity (whose distinguishing feature is to be temporally articulated), and in recovering the intertemporal co-ordination disrupted by the introduction of the new technique. Physical, human, and financial capital are complementary in this process of reshaping, and may constrain each other. The outcome of the disequilibrium process depends then on the interaction of accumulation choices, learning, and money supply rules.

The out-of-equilibrium analysis, we have seen, makes it possible to show how a shock of any kind brings about first and foremost a distortion of the existing productive capacity due to a breaking of the intertemporal complementarity of the production process. This implies the appearance of disequilibria, and hence of problems of co-ordination that extend to all aspects of economic activity (resulting, for example, in inflation, unemployment, and so on). Reactions to these disequilibria stimulate a process of adjustment sketched out by sequentially interacting disequilibria, which will amplify or dampen the original deformation of the structure of productive capacity—and hence create or eliminate viability problems—according to the working of the co-ordination mechanisms along the way. If co-ordination is not re-established, this will result in particular in increasing levels of unemployment, and decreasing levels of productivity and real wages (Amendola *et al.*, 2005).

Consider the case of the introduction of a new technology characterised by higher construction costs, as it is typical of the new information and communication technologies. The costs come earlier, and hence cannot be financed out of current production This causes a distortion of productive capacity and the dissociation in time of inputs from output, and of costs from receipts, which puts a financial constraint on investment in capacity. The availability of financial resources *at the right time* is then essential to build a bridge over time between costs and revenues, so as to render the required restructuring of productive capacity viable while it is still on the way and does not yet deliver output and revenues. If these resources are not available, the necessary investment cannot be realised, which will further reduce final output and postpone (or even cast doubts on the effective obtainment of) the expected increases in productivity. What we shall have in the meantime is less production and less labor demand. Unemployment, lower revenues and the subsequent fall in final demand will further reduce receipts and financial resources. Insufficient investments will paradoxically result in excess supply, excessive productive capacity and in the scrapping of production processes. And so on, in a process that is a threat to the viability of the change undertaken. Viability that, therefore, calls for the kind of discretionary monetary policy described in the preceding section.

This process also occurs if the new technology requires a different gamut of skills. We shall immediately have the appearance of a human resource constraint, taking the form of a labour mismatch, which implies the co-existence of unemployment and unfilled vacancies (for lower and higher skills respectively). Once again this will result in lower investment and hence in a subsequent fall in revenues and final demand. Unemployment thus reveals the existence of co-ordination problems at the economy level. It cannot be reduced to a matching problem, to be solved thanks to appropriate changes in the regulations of the labour market or appropriate actions that would allow workers to learn new competencies.

Of course, with a fully rational behaviour making available the financial resources aimed at covering balanced investment expenses, equilibrium will be maintained or mechanically re-established. But this simply means wiping out by assumption the coordination problems that arise along a process of qualitative change and the implied requirement of a macroeconomic management of the process itself.

The above scenario illustrates the productivity paradox, that is, a fall in productivity notwithstanding the introduction of a superior technique in terms of production coefficients. There is a divorce between the productivity of the technique, which can only be verified in an economy already in the equilibrium state associated with the technique itself, and the effective productivity of the economy resulting from how the out-of-equilibrium process of transition takes place. This divorce has nothing to do with the specific character of the technique concerned; it depends on the co-ordination problems that arise in the transition process from an old technique to a new one.

3.4. Trade and domestic distortions

Comparative advantages postulate that an increase of exchanges between countries is generally beneficial to all partners. Importing new goods and services, even when these goods were previously domestically produced, creates new opportunities and allows the use of productive resources in a more efficient way. The loss of manufacturing jobs due to the growing import penetration is generally offset by the job creation effect of growing exports. International trade is thus a positive sum game and cannot be held responsible for increasing unemployment, waste of resources, and low growth.

However, old as well as more recent analyses demonstrate the possibility of losses for some participants to the exchange. These losses would be essentially due to differences in productivity gains among countries, which result in differences in real income (Hicks, 1953; Krugman, 1985; Gomory and Baumol, 2000; Samuelson, 2004). These models deal with the welfare effects for a country when a part of domestic production is taken over by its trading partner, generally a less advanced country. Usually, changes in international trade result in widespread gains if there are no obstacles to prevent the redistribution of productive resources among sectors that allows the convergence toward the full employment equilibrium. Within the standard analytical framework, these considerations lead to focus on the role played by wage adjustments and distortions associated with them. For the gains from trade and relocation to materialize, it is essential that no domestic distortion prevents the necessary adjustment (*i.e.* the convergence towards the full employment equilibrium). Changes in fundamentals (technology and preferences) must be accommodated by relative prices (in particular wages). In this case, relocation and outsourcing only correspond to a better allocation of resources at the international level without harmful consequences on employment. Increasing imports will be matched by increasing exports.

Out of equilibrium models are not concerned by the final welfare effects of changing trading patterns, but with the positive implications of the transition process. One cannot deny that changes in international trade entail social and distributional costs: "An irony that is not sufficiently appreciated in the public debate is that the economist's case for gains from trade relies heavily on the restructuring of national economies by the forces of trade: specialization requires restructuring. While re- structuring may take different forms, in most cases it is likely to have distributional impacts-both in the short term as a consequence of adjustment costs and in the long-term as a result of permanent changes in relative factor demands. One might even say that the dislocations and distributional consequences produced by trade are the flip side of the efficiency gains. No pain, no gain!" (Rodrik, 1998, p.6).

The restructuring mentioned by Rodrik is an intrinsic feature of globalization and of relocation processes. In fact, increasing openness is a form of structural change and, hence, analytically equivalent to technical progress; as such, it entails the destruction of existing productive capacity (and of the corresponding jobs), and the construction of something new to replace it. Changes in international trade go hand-to-hand with the breaking-up of the pre-existing industrial and spatial structure of productive capacity, which results in unavoidable disequilibria between supply and demand of final goods, all along the transition towards the new adapted structure of the economy. Thus, the supply side, in particular investment, becomes crucial for the transition to a new steady state. It is therefore pointless, when not harmful, to try to bypass the transition and the associated turbulence by eliminating price distortions. Policy should rather accompany the process of change, progressively removing or softening the constraint faced by the economy.

The out-of-equilibrium approach allows to push Rodrik's argument even further, by arguing that this process of restructuring does not necessarily converge to the new equilibrium: the *ex ante* benefits from increased openness may *ex post* fail to materialize, if something goes wrong with the co-ordination process. Thus, the process of restructuring needs not to be successful. The viability of the transition and the recovery of coordination crucially hinge upon the right mix of institutional and policy factors, notably in access to credit.

A two sector oligopolistic model that is subject to an external demand shock (Gaffard and Saraceno, 2012). Firms can migrate between sectors, following relative profits, but need to adapt their productive capacity to the new sector of activity. The natural tendency of the system to converge to the new steady state equilibrium corresponding to a different demand structure may be hampered by excessive variations in wages and/or by too fast migration between sectors; these may trigger, via aggregate demand effects, an important drop in the investment capacity of firms. In turn, if this lack of resources is not compensated by the credit sector, the insufficient investment disrupts the productive capacity of the economy, and triggers a cumulative explosive process. Therefore, re-establishing the coordination between investment and consumption and reabsorbing unemployment requires an accommodating credit policy, and a sufficiently slow change in wages on the one side and on migration rates on the other. Excessive rigidity, on the other hand, will result in a new equilibrium permanently characterized by unemployment. The paper concludes therefore that there is a sort of “optimal” degree of flexibility for the economy.

4. The current crisis from an out-of-equilibrium perspective

As already stressed, any structural change is a process of development defined as “disturbance of equilibrium, which forever alters and displaces the equilibrium state previously existing” (Schumpeter, 1934, p. 64). The on-going crisis is clearly a moment of such a process, and should be analysed in this perspective. Of course, financial (mis)behaviours have played an essential role in triggering the crisis. But, its roots are real. Technological shocks, growth strategies carried out by emergent countries, and, most important, dramatic changes in the income distribution, have generated large distortions in the structure of productive capacity in several countries. These have resulted in the involuntary accumulation of real and financial stocks.

Thus, the on-going crisis should have led to understand and address the policy mistakes that prevented the world economy from fully adjusting to the unavoidable structural changes, rather than to propose the same recipes that prevailed before the crisis. As a matter of fact, capitalism is submitted to recurrent structural changes and its survival depends on the way co-ordination takes place. Private (market) or public (policy) co-ordination will be successful when leading to the harmonisation between supply and demand at each moment of time and over time, that is, when smoothing adjustment processes. This co-ordination consists in arbitrages between conflicting objectives, but also requires a harmonisation of interests, which in turn is possible only through a fair distribution of income.

The sequence of events that have been observed can be explained with reference to the analytical construct described above, *i.e.* by focusing on the divorce between investment and consumption that characterized most countries.

The US crisis can be interpreted as the consequence of the way a deep structural change, mostly linked to technological advances, has been managed. During a first phase, financial markets and monetary policy have allowed investments in new technologies to be easily financed (Amendola *et al.*, 2005). But delayed reaction to the building up of disequilibria led to overinvestment and to the emergence of a stock market bubble that eventually burst. In the second phase, the indebtedness of households belonging to poor and middle classes compensated the increased income inequality that should have had a negative effect on final demand and on the potential growth rate (Fitoussi and Saraceno, 2010; 2011). Indebtedness would have created inflationary pressure and would have led the Federal Reserve to apply a tight monetary policy, if the gap between domestic supply and demand for final goods had not been filled by imports from emerging countries. To summarize, imbalances in the structure of productive capacity have opened the way to cumulative disequilibria (both real and financial, domestic and foreign) that have resulted in a great contraction.

The EU crisis is a different story. During two decades large European countries have experienced a lower growth rate that has ended in increasing budget deficits. This has been the result of a restrictive monetary policy that compressed inflation, but also prevented adequate investments in new technologies. In other words, the transition to the new productive capacity that would have incorporated new technologies has not been completed (Amendola *et al.*, 2005). Nevertheless there were no strong imbalances between investment and consumption that would have rendered unviable this slow growth path. A serious problem arose during the 2000's, when Germany adopted a strategy that has consisted in stimulating exports while constraining its domestic demand with labor market reforms (Carlin and Soskice, 2008). This has resulted in a divorce between Germany and other European countries, that is, between a country with a current-account surplus and countries with deficits, that contributed to the crisis of the Euro in 2011. A strong imbalance between domestic investment and consumption in Germany required a high level of consumption in some other developed countries. Decreasing interest rates in the euro zone periphery and available funds in particular from German banks have fuelled housing bubbles, specifically in Spain, where a symmetric distortion has arisen: domestic consumption was no longer in line with investment in new productive capacity.

In China, excessive inequality, and an insufficient provision of welfare (in particular health care and pensions), led to the necessity of an export led growth and to the ensuing accumulation of assets. Given the large (and sometimes excessive) investment driving the fast growth carried out, only increased exports of goods have allowed absorbing the resulting supply. If this pattern of growth is to be reversed, the growth of investment must fall well below that of GDP and consumption must be dramatically augmented. In our framework, the economy should re-establish a balance between the construction and the utilisation of production processes at the domestic level. This transition to greater reliance on internal consumption might be quite bumpy, and, it should be managed smoothly to be successful. The government response to the crisis, that took the form of increased public investment, and incentive to private capital accumulation, sustained the economy in the short run, but widened the imbalance between consumption and investment, making the long run adjustment harder and more necessary at the same time.

The global imbalances that result from the prevailing relations between advanced and emerging countries (mainly between the US and China), but also among advanced countries (e.g., between Germany and Spain), take the form of national current-account surpluses and deficits offset by net capital inflows. According to the international intertemporal trade agreement, these global imbalances should create no problems since surplus countries are foregoing goods and services today but expect, in return, to receive net goods and services tomorrow. This is what Corden (2011) calls 'the return journey'. In this scenario, borrowing is supposed to be aimed at financing sound investment, and to provide for the return journey. However, what has happened is that financial resources thus made available have actually been devoted to finance increased current consumption and unproductive investment (housing). As a consequence, a large imbalance has appeared between consumption and sound investment, which is not sustainable. This is an example of the paradox of thrift. We must stress that we are not only concerned with the divorce between saving and (sound) investment, but also with the imbrications of successive disequilibria that push economies out of their stability corridor.

Focusing on the distortions between investment and consumption as the engine driving the evolution of the economy, and identifying them as one of the main causes that pushed economies out of their stability corridor, helps to better understand the intrinsic complexity of the situation. This also reveals how difficult is to elaborate exit strategies for macroeconomic policies. Re-establishing a better co-ordination between investment and consumption will take time. Governments should be able both to smooth short-term fluctuations, and, at the same time, to favour a restructuring of the economy. It would then be a mistake to focus on fiscal consolidation and to ask to implement structural policies as if the only problem were to re-establish a balanced public budget. Nevertheless, the old Keynesian recipes are also not enough, because as the case of China teaches, they may worsen the co-ordination problems the economy faces. The transition paths should take place in such a way as to correct existing distortions, which means obtaining greater reliance on investment in some countries, on consumption in other ones, while sustaining aggregate demand in the short run. In both cases, this requires adjustments in the structure of the productive capacity and eventually in the distribution of income and wealth. Such changes take time and must be managed in a way that prevents the economy to experiment too strong fluctuations in the meantime. The real challenge is to co-ordinate and harmonize short-term and long-term policies. This may require that inflationary pressures and budget deficits are accepted for a while, when not sought for. Structural policies should not be oriented towards more flexibility on the market, but, at the opposite, they should favor rigidities that permit smoothing the necessary adjustments. Indeed, "the crisis has also put to the test long-standing dogmas that blame labor-market rigidity for unemployment, because countries with more flexible wages, like the U.S., have fared worse than northern European economies, including Germany" (Stiglitz, 2011).

References

Amendola M. and J.-L. Gaffard, 1998. *Out of Equilibrium*. Oxford and New York. Oxford University Press Clarendon Press.

Amendola M., J.-L. Gaffard and F. Saraceno, 2004a. "Technological Shocks and the Conduct of Monetary Policy". *Revue Economique*. 55(6). 1241-63.

Amendola M., J.-L. Gaffard and F. Saraceno, 2004b. "Wage Flexibility and Unemployment: The Keynesian Perspective Revisited". *Scottish Journal of Political Economy*. 51(5). 654-74.

Amendola M., J. L. Gaffard and F. Saraceno, 2005. "Technical Progress, Accumulation and Financial Constraints: Is the Productivity Paradox Really a Paradox?". *Structural Change and Economic Dynamics*. 16(2). 243-61.

Bénassy J.-P., 1982. *The Economics of Market Disequilibrium*. New York. Academic Press.

Blanchard O. and L. F. Katz, 1997. "What We Know and Do Not Know About the Natural Rate of Unemployment". *Journal of Economic Perspectives*. 11(1). 51-72.

Carlin W. and D. Soskice, 2008. "Reforms, Macroeconomic Policy and Economic Performance in Germany". in *Economic Policy Proposals for Germany and Europe*. R. Schettkat and J. Langkau, Eds., Routledge, 72-118.

Clarida R., J. Gali and M. Gertler, 1999. "The Science of Monetary Policy: A New Keynesian Perspective". *Journal of Economic Literature*. 37(4). 1661-707.

Clower R., 1965. "The Keynesian Counterrevolution: A Theoretical Appraisal". in *The Theory of Interest Rates*. F. H. Hahn and F. P. R. Brechling. Ed.. MacMillan.

Corden W. M., 2011. "Global Imbalances and the Paradox of Thrift.". *CEPR Policy Insight*. 54

Fitoussi J.-P. and F. Saraceno, 2010. "Europe: How Deep Is a Crisis? Policy Responses and Structural Factors Behind Diverging Performances". *Journal of Globalization and Development*, 1(1). Article 17.

Fitoussi J. P. and F. Saraceno, 2011. "Inequality, the Crisis and After". *Rivista di Politica Economica*. III(1-3). 9-28.

Gaffard J.-L. and F. Saraceno, 2012. "International Trade and Domestic Distortions". *Journal of Evolutionary Economics*. 22(2). 275-301.

Garegnani P., 1984. "Value and Distribution in the Classical Economists and Marx". *Oxford Economic Papers. N. S.*. 36 2, 291-325.

Gomory R. E. and W. J. Baumol, 2000. *Global Trade and Conflicting National Interests*. Cambridge. Mass. MIT Press.

Hayek F. A., 2002. "Competition as a Discovery Procedure", *Quarterly Journal of Austrian Economics*. 5(3). 9-23. (Originally published as Der Wettbewerb als Entdeckungsverfahren". Kieler Vorträge. 56. 1968).

Hicks J. R., 1939. *Value and Capital*. Oxford. Clarendon Press.

Hicks J. R., 1953. "An Inaugural Lecture". *Oxford Economic Papers. New Series*. 5. 117-35.

Hicks J. R., 1956. "Methods of Dynamic Analysis". in *Collected Essays on Economic Theory. Vol. Ii: Money. Interest and Wages* Basil Blackwell. 1982. 217-35.

Hicks J. R., 1970. "A Neo-Austrian Growth Theory", *Economic Journal*, 80(3183. 257-81.

Hicks J. R., 1973. *Capital and Time*. Oxford. Clarendon Press.

Hicks J. R., 1989. *A Market Theory of Money*. Oxford. New York. Clarendon Press.

Hommes C. H., 1998. "On the Consistency of Backward-Looking Expectations: The Case of the Cobweb". *Journal of Economic Behavior and Organizatio*. 33(3-4). 333-62.

Howitt P., 1994. "Adjusting to Technological Change". *Canadian Journal of Economics*. 27 4. 763-75.

Kaldor N., 1985. *Economics without Equilibrium*. Cardiff. U.K. University College Cardiff Press.

Keynes J. M., 1936. *The General Theory of Employment. Interest. and Money*. London. MacMillan.

Krugman P. R., 1985. "A 'Technology Gap' Model of International Trade". in *Structural Adjsutment in Advanced Economies*. K. Jungenfelt and D. Hague. Ed McMillan. 35-49.

Kydland F. E. and E. C. Prescott, 1982. "Time to Build and Aggregate Fluctuations". *Econometrica*. 50(6). 1345-70.

LeBaron B. and L. Tesfatsion, 2008. "Modeling Macroeconomies as Open-Ended Dynamic Systems of Interacting Agents". *American Economic Review*. 98 2. 246-50.

Leijonhufvud A., 1968. *On Keynesian Economics and the Economics of Keynes; a Study in Monetary Theory*. New York. Oxford University Press.

Malinvaud E., 1977. *The Theory of Unemployment Reconsidered*. Oxford. Blackwell.

Mehrling P., 2006. "The Problem of Time in the Dsge Model and the Post Walrasian Alternative". in *Post Walrasian Macroeconomics : Beyond the Dynamic Stochastic General Equilibrium Model*. D. C. Colander. Ed. Cambridge University Press. 70-9.

Merton R. K., 1936. "The Unanticipated Consequences of Purposive Social Action". *American Sociological Review*. 1(6). 894-904.

Phelps E. S., 2007. "Macroeconomics for a Modern Economy. Nobel Lecture". *American Economic Review*. 97 3. 543-61.

Phelps E. S. and G. Zoega, 1998. "Natural-Rate Theory and OECD Unemployment". *Economic Journal*. 108 448. 782-801.

Rodrik D., 1998. "Symposium on Globalization in Perspective: An Introduction". *Journal of Economic Perspectives*. 12 4. 3-8.

Samuelson P. A., 2004. "Where Ricardo and Mill Rebut and Confirm Arguments of Mainstream Economists Supporting Globalization". *Journal of Economic Perspectives*. 18(3). 135-46.

Saraceno F., 2004. "Wage Regimes, Accumulation, and Finance Constraints: Keynesian Unemployment Revisited." Observatoire Français des Conjonctures Économiques, *Document de Travail*. 2004-01. January.

Schumpeter J. A., 1934. *The Theory of Economic Development; an Inquiry into Profits. Capital, Credit, Interest, and the Business Cycle*. Cambridge. Mass. Harvard University Press.

Stiglitz J. E., 1987. "The Causes and Consequences of the Dependence of Quality on Price". *Journal of Economic Literature*. 25(1). 1-48.

Stiglitz J. E., 2011. "The Ideological Crisis of the Western Capitalism". *Project Syndicate*. July 6th.

Woodford M., 2003. *Interest and Prices : Foundations of a Theory of Monetary Policy*. Princeton. Princeton University Press.

Appendix:
Analytical Elements of an Out-Of-Equilibrium Model

A.1. Technology

The two elements of complementarity and time-to-build can be introduced via an Hicksian representation of technology (Hicks, 1970; 1973): Consider a sequential economy of a neo-Austrian type which uses a homogeneous labour resource. Labour is inputted for *n* periods to build the productive capacity, and used for the following *N* to operate it and obtain a final output. An elementary process of production is defined by input coefficients such that:

$$\begin{aligned} \mathbf{a}^c &= \left[a_i^c\right] \qquad \forall i = 1,\dots,n \\ \mathbf{a}^u &= \left[a^u\right], \quad \forall i = n+1,\dots,n+N, \end{aligned} \tag{A.1}$$

and output coefficients

$$\mathbf{b} = \left[b_i\right], \qquad \forall i = n+1,\dots,n+N. \tag{A.2}$$

We usually assume that $a_i^c = a^c$, $a_i^u = a^u$, and $b_i = b$, even if any time profile can be modeled through the appropriate choice of vectors. The productive capacity of the economy is given by the number of processes in use at the moment *t*, in construction, $x^c(t)$ and in utilization, $x^u(t)$:

$$\mathbf{x}(t) = \left[\mathbf{x}^c(t), \mathbf{x}^u(t)\right]. \tag{A.3}$$

This capacity is subject to ageing and to modifications due to investment and scrapping of processes in case of financial constraints.

Alternatively, especially if the construction phase is short enough, it can be assumed that the production function takes the form of a Leontief function with dated labor input

$$q_t = min[\kappa l_{t-1}, \lambda l_t] \tag{A.4}$$

with q denoting quantity produced, and l denoting labour imput either at time t or at t-1. Thus, dated and current labor (l_{t-1} and l_t) concur in fixed proportions to the determination of production q ; this formulation is equivalent to assuming capital, built by labor in the previous period, that fully depreciates.

Both with an Hicksian and a Leontief technology, past "investment" may constrain current production: if firms don't possess the appropriate amount of capital/dated labor, they will not be able to produce as much as they wish. The Leontief representation is analytically more treatable, whereas the Hicksian representation allows a finer description of the time structure of production.

A.2. Expectations

We discussed at length, above, why short term expectations (also called *intraperiod*) need to be anchored in past behaviour. A general formulation, for a generic variable x (usually expected demand) contains three terms:

$$x_t^e = \phi x_{t-1} + \gamma x^* + \delta\left(\frac{x_{t-1} - x_{t-1}^e}{x_{t-1}}\right) \tag{A.5}$$

The first is the past value of the variable; the second is the 'normal' value, and the third is an error correction term. An appropriate choice of the coefficients ϕ γ and δ allows to describe a wide range of adaptive behaviours. Steady state/ equilibrium values may anchor long term or *interperiod* expectations (for example those affecting variables like investment, in human and physical capital). In this case, in eq. expect, we would have $\phi = \delta = 0$ and $\gamma = 1$ so that $x^e = x^*$.

A.3. Prices

Out-of-equilibrium models borrow from the fix-price literature (Hicks, 1939; Malinvaud, 1977; Bénassy, 1982), the idea that disequilibria are absorbed by quantity adjustments (short-side rule), while prices only change discretely over time. Analytically, this is obtained by having wages and prices fixed within periods, and adjustments that take place only between the periods (the Hicksian 'weeks'). The adjustment can simply follow previous excess demand (D-S), for example

$$p_t = p_{t-1}\left(1 + \omega\frac{D_{t-1} - S_{t-1}}{S_{t-1}}\right). \tag{A.6}$$

ω_j is an indicator of price flexibility that nevertheless, as the equation clarifies, has nothing to do with market clearing behavior. Alternatively, one can have a formulation that echoes the Calvo (1983) partial adjustment scheme:

$$p_t = \psi p^* + (1-\psi) p_{t-1}, \tag{A.7}$$

with ψ denoting the fraction of firms adapting their price to its optimal value at each period.

A.4. Quantity Adjustments

The short side rule applies to obtain equilibrium in most markets. The actual value of a variable is thus computed as

$$x_t = min[x_t^d, x_t^s]. \tag{A.8}$$

The markets subject to quantity adjustments may be goods, labour or financial markets.

A.5. The Sequence

We said before that the time structure of the model is generally obtained through a sequence of periods linked by state variables such as the quantity of (dated) labor embedded in production processes, the stocks that result from past disequilibria, and past prices and wages. The interperiod sequence is complemented by an intraperiod sequence that allows the emergence of disequilibria:

Prices and wages change in response to market disequilibria, even if we do not let them clear markets (eqs. A.6 or A.7).

Firms form expectations (eq. A.5), Given expectations, the technology (eqs. A.1, A.2 and A.3 or A.4), and the stock of dated labor l_{t-1}, firms decide desired demand (for labor and external funds, in case the internal funds from previous periods are not sufficient) and supply (of goods).

The first market that opens is the financial market, in which demand for external funds may be rationed (eq. A.8). Financial constraints cause a rescaling of labor demand.

Total labor employed is determined once the second market, the labor market opens, where once again eq. A.8 determines whether unemployment or a human resource constraint appears. Then wages are paid, and production is carried over. Households form their demand based on the actual wage perceived.

Finally, the product market opens; as in the other markets, rationing may appear. Rationing in the goods and labor market will determine the change in prices and wages between periods, as well the stocks carried on from period to period.

STRUCTURAL INTERACTIONS AND LONG RUN GROWTH
AN APPLICATION OF EXPERIMENTAL DESIGN TO AGENT BASED MODELS[1]

Tommaso Ciarli
SPRU, University of Sussex, UK

We propose an agent-based computational model defining the following dimensions of structural change—organisation of production, technology of production, and product on the supply side, and income distribution and consumption patterns on the demand side—at the microeconomic level. We define ten different parameters to account for these five dimensions of structural change. Building on existing results we use a full factorial experimental design (DOE) to analyse the size and significance the effect of these parameters on output growth. We identify the aspects of structural change that have the strongest impact. We study the direct and indirect effects of the factors of structural change, and focus on the role of the interactions among the different factors and different aspects of structural change. We find that some aspects of structural change—income distribution, changes to production technology and the emergence of new sectors—play a major role on output growth, while others—consumption shares, preferences, and the quality of goods—play a rather minor role. Second, these major factors can radically modify the growth of an economy even when all other aspects experience no structural change. Third, different aspects of structural change strongly interact: the effect of a factor that influences a particular aspect of structural change varies radically for different degrees of structural change in other aspects. These results on the different aspects of structural change provide a number of insights on why regions starting from a similar level of output and with initial small differences grow so differently through time.

Keywords: Structural change, Long run growth, ABM; DOE.

1. The paper builds on previous work and discussions with André Lorentz, Maria Savona and Marco Valente, to whom I am indebted for their advice, suggestions, and support. Insightful comments from a referee helped to improve significantly the paper. This work is based on an intensive use of simulations that would have not been possible (or would have taken a few years) without the access to the computer cluster of the Max-Planck-Institut für Mathematik in den Naturwissenschaften in Leipzig, facilitated by Rainer Kleinrensing, Ronald Kriemann, and Thomas Baumann. Financial support by the Max Planck Institute of Economics in Jena, where I have conducted a substantial part of the present research as a research fellow in the Evolutionary Economics group, is gratefully acknowledged. Any error is my own responsibility.

The dramatic increase in output and consumption following the industrial revolution was accompanied by substantial changes in the structure of the economies involved. Countries of late industrialisation and current transition countries are also experiencing dramatic changes (Dasgupta and Singh, 2005). Economists usually refer to structural change as the reshuffling in the share of employment or value added in the three main sectors: agriculture, manufacturing and services (Clark, 1940; Fisher, 1939; Dietrich and Krüger, 2010; Baumol, 2010) which has led to these grand economic shifts to be described as "industrialisation" and "tertiarisation" of advanced economies. However, structural changes encompass more than shifts in labour and value added from one sector to another; they include complex adjustments in the structure of production, consumption, labour organisation and income distribution, which interact in a continuous evolutionary process. For instance, industrialisation is accompanied by the concentration of production in large capital intensive firms and firm size growth (Desmet and Parente, 2009), an increase in the number of goods available for final consumption (Berg, 2002), closer involvement of science in technological change (Mokyr, 2002), increased use of capital in agriculture and especially manufacturing accompanied by an improvement in the technology embedded in new machines and overall increases in productivity (Kuznets, 1973), greater urbanisation usually accompanied by increased income inequality and changes in social class composition (McCloskey, 2009), and so on. In other words, industrialisation leads to transformations of economies and societies. Thus the definition proposed by Matsuyama, that structural change is "complementary changes in various aspects of the economy, such as the sector compositions of output and employment, the organisation of industry, the financial system, income and wealth distribution, demography, political institutions, and even the society's value system" (Matsuyama, 2008).

To be sure, some changes precede income growth, others unfold as a consequence of income growth, and there are interactions among the different aspects of structural change. For instance, changes in the distribution of income are related to changes in class composition and patterns of consumption. Changes to class composition, in their turn, are related to the accumulation of capital and the different organisation of labour. The accumulation of capital induces the search of new technologies embedded in more efficient capital goods, and so on.

Ideally we would like to explain the changes in each aspect of structural change, their co-evolution and their effect on the direction of economic growth and on other dimensions of structural change. We believe that such an investigation is fundamental to shed light on the determinants and dynamics of long-run growth, and to derive policy implication that consider different aspect of economic change. This is especially relevant since traditional explanations of the relation between structural change and growth point to opposing dynamics (Matsuyama, 2008): i) exogenous changes in productivity in the manufacturing sector—which somehow emerge in the economy—induce labour migration from agriculture to industry; and ii) an increase of productivity in agriculture reduces demand for labour and induces migration to the manufacturing sector where capital investment—characterised by higher increases in productivity per unit of investment—spurs growth; the more investment that is concentrated in manufacturing, the greater manufacturing productivity increases. Both these mechanisms are plausible. However, taken in their basic version they do not acknowledge the wide array of "complementary changes" they are conducive to, and which help in solving their contradiction. We believe that a more accurate explanation should include the various economic aspects that accompany the transformation of an economy.

In this paper we heed Matsuyama (2008) definition of structural change and model complementary changes in various aspects of the structure of an economy, namely organisation of production, technology of production, and product on the supply side, and income distribution and consumption patterns on the demand side. However, we also follow Saviotti and Gaffard (2008) suggestion and investigate the microeconomic sources of structural changes. Saviotti and Gaffard (2008, p. 115), in line with Matsuyama (2008), define structural change as a "change in the structure of the economic system, that is, in its components and in their interactions. Components are [...] particular goods or services, and other activities and institutions, such as technologies, types of knowledge, organisational forms etc.". However, departing from Matsuyama (2008), they ask: "What does it mean for a system to be in equilibrium when its composition keeps changing due to the emergence of qualitatively different entities?" [p. 116].

We take on board these remarks and propose a model of the microeconomic dynamics of structural change as processes that never reach equilibrium, because of the continuous changes to the underlying dimensions of the economy. In order to model these microeconomic interactions and study the emergent structural change and aggregate output, we use computational models and solutions (Colander *et al.*, 2008; LeBaron and Tesfatsion, 2008; Dosi *et al.*, 2010; Leijonhufvud, 2006; Buchanan, 2009; Delli Gatti *et al.*, 2010; Dawid and Semmler, 2010).

We propose an agent-based computational model defining the following dimensions of structural change—organisation of production, technology of production, and product on the supply side, and income distribution and consumption patterns on the demand side—at the microeconomic level. We model their co-evolution in terms of the interactions among the different agents on the supply and demand sides, and the changing behaviour promoted by changes to income and structure. We contribute to the traditional literature on structural change by accounting for 'complementary changes' and in a micro to macro framework, which can be treated exhaustively using agent based computational models.

The model includes two types of firms: capital and final goods producers. Final goods producers produce goods that satisfy different consumption needs, serving different markets. New markets emerge as an outcome of firms' investments in innovation. Consumer goods differ also with respect to their quality. A firm includes many layers of employees (workers and managers at different levels), with each layer earning a different wage. This creates consumers with unequal income distribution. Consumers are grouped into classes that demand different varieties of goods, affecting firm demand. Among other things, this implies that the larger the number of organisational layers required in the firm (organisational complexity), the higher are the differences across consumers, *ceteris paribus*. Each class distributes its consumption differently across the different markets. These consumption shares evolve endogenously as new classes emerge in the economy, representing Engel curves. Growth results from demand expansion, which is a joint outcome of firm selection and technology investment.

The structure of the model is based on Ciarli *et al.* (2010) and Ciarli and Lorentz (2010), which discuss the micro economic dynamics that lead to growth in output *via* endogenous changes in different aspects of economic structure. Ciarli *et al.* (2012) discuss the non-linear effects of organisational complexity, production technology and product variety on income growth and distribution. They show that output is negatively related to initial product and demand variety, organisational complexity and faster technological change in capital goods increase output despite higher inequality, and this last, in the form of large earning disparities, leads to lower output growth.

In this paper we build on existing results and assess the relative importance of all the factors that, in the model, determine the initial conditions of structural change and also the pace at which the different aspects of the economic structure evolve. The organisation of production is defined by the structure of labour and earnings disparities. Production technology is defined by the speed of change in capital innovation, the share of resources invested in R&D, and its success. Product technology is defined by the ability of firms to explore new sectors for a given level of R&D investment, improved quality of a new product, and share of resources invested in R&D. Income distribution is studied in relation to profits in capital and final goods firms. Consumption patterns are defined by the speed at which consumption shares change with increases in income and class differentiations and changes in consumer preferences promoted by the emergence of different income classes. Whilst we define each aspect of structural change based on specific factors, most of these factors induce structural change in several aspects of the economy. For instance, the organisation of labour has an impact on the evolution of income classes and, therefore, also on patterns of consumption; the resources invested in R&D reduce the profits available to be shared among firm managers, which affects income distribution; and so on.

We use a full factorial experimental design (DOE) to analyse the size and significance of the impact of the parameters that define structural change, on output growth. We decompose and identify the aspects of structural change that have the strongest impact on growth. We study the direct and indirect effects of the factors of structural change, where indirect effects are those that occur through those variables that also have an impact on income growth. We focus on the role of the interactions among the different factors and different aspects of structural change.

Interactions among factors are of particular interest here, since the early steps in the analysis show that in most cases the effect of one specific factor that influences a particular aspect of structural change varies radically for different levels of the other factors. In many cases, the main effect of a factor defining the economic structure is inverted under different structural conditions defined by other factors. Second, we find that some aspects of structural change, such as income distribution, changes to production technology and the emergence of new sectors, play a major role on output growth, while the roles of others, such as changes in consumption shares, preferences, and the quality of goods, play a rather minor role. Related to this, we find that some factors can radically modify the growth of an economy even when all other aspects experience no structural change, whereas most factors, on their own, do not affect outcomes if all other aspects change rapidly. In other words, one single factor that induces rapid changes in one particular aspect of the economy can induce changes that lead to large growth in output; however, in economies already undergoing structural change in most aspects, slow changes in most other factors have little influence. Finally, we find that, when controlling for other model variables, the effect of most factors on output growth is significantly reduced, showing large indirect effects.

The arguments are organised as follows. First, we describe the model focussing on the main micro dynamics and the main aspects that are mostly affected by the factors that define structural change (Section 1). Next we describe the methodology and briefly present the model initialisation and design of experiment (DOE) (Section 2). Section 3 is divided in four subsections. First, we describe the general properties of the model, compare the model's output with some empirical evidence, and show how the distribution of world income across countries can be explained by different initial factors, with some caveats. Second, referring to the model, we show how each factor is suited to analysing one or more aspects of structural change. Third, we use analysis of variance to determine the significance of the main effects of factors and of the interaction effects between pairs of factors. Fourth, we show results from an econometric analysis of the factors to quantitatively assess their relevance in the model, to distinguish direct from indirect effects, and to assess the relevance and direction of the first order interaction between each pair of factors. Section 4 discusses the results and concludes the paper.

1. Model

1.1. Final good Firms

We model a population of $f \in \{1,2,...,F\}$ firms producing final goods for the consumer market. Each good satisfies one consumer need $n \in \{1,2,...,N\}$. Or, equivalently, each firm produces in one of the $n \in \{1,2,...,N\}$ sectors. For simplicity we refer interchangeably to needs and sectors.[2] The firms produce an output addressing a consumer need n with two characteristics i_{j,f_n}: price $p_{f,t} = i_{1,f_n}$ and quality $q_{f,t} = i_{2,f_n}$.

1.1.1. Firm output and production factors

Firms produce using a fixed coefficients technology:[3]

$$Q_{f,t} = min\ \left\{Q^d_{f,t}\ ;\ A_{f,t-1} L^1_{f,t-1}\ ;\ BK_{f,t-1}\right\} \tag{1}$$

where $A_{f,t-1}$ is the level of productivity of labour $L^1_{f,t-1}$ embodied in the firm's capital stock $K_{f,t-1}$. $Q^d_{f,t}$ is the output required to cover the expected demand $Y^e_{f,t}$, past inventories $S_{f,t-1}$, and the new inventories $\bar{s}Y^e_{f,t}$: $Q^d_{f,t} = (1+\bar{s})\ Y^e_t - S_{t-1}$. The capital intensity $1/B$ is constant.[4]

2. In referring to the same good, we prefer to refer to firm innovation in terms of sectors and consumer demand in terms of needs. Establishing a mapping between the two is not one of the aims of this paper and, ultimately, depends on the definition of *sectors*.

3. For the sake of readability we omit the sector/need index *n*.

4. This assumption is supported by evidence from several empirical studies, starting with Kaldor (1957). The capital investment decision ensures that the actual capital intensity remains fixed over time.

Firms form their sales expectations in an adaptive way to smooth short term volatility (Chiarella *et al.*, 2000): $Y^e_{f,t} = a^s Y^e_{f,t-1} + (1 - a^s) Y_{f,t-1}$, where ($a^s$) defines the speed of adaptation. We assume that the level of demand faced by a firm is met by current production ($Q_{f,t}$) and inventories ($S_{f,t-1} \geq 0$), or is delayed ($S_{f,t-1} < 0$) at no cost. Following Blanchard (1983) and Blinder (1982), production smoothing is achieved by means of inventories $\bar{s}Y^e_{f,t}$ —where s is a fixed ratio.[5]

Given $Q^d_{f,t}$, labour productivity $A_{f,t-1}$ and an unused labour capacity (u^l) to face unexpected increases in final demand, firms hire shop-floor workers:

$$L^1_{f,t} = \varepsilon L^1_{f,t-1} + (1-\varepsilon)\left[\left(1+u^l\right)\frac{1}{A_{f,t-1}} min\left\{Q^d_{f,t}; BK_{f,t-1}\right\}\right] \tag{2}$$

where ε mimics labour market rigidities. Following Simon (1957) firms also hire "managers": every batch of ν workers requires one manager. Each batch of ν second tier managers requires a third level managers, and so on. The number of workers in each tier, given $L^1_{f,t}$ is thus

$$\begin{aligned} L^2_{f,t} &= L^1_{f,t}\nu^{-1} \\ &\vdots \\ L^z_{f,t} &= L^1_{f,t}\nu^{(1-z)} \\ &\vdots \\ L^{\Lambda_f}_{f,t} &= L^1_{f,t}\nu^{\left(1-\Lambda_f\right)} \end{aligned} \tag{3}$$

where Λ_f is the total number of tiers required to manage the firm f.

Consequently, the total number of workers is $L_{f,t} = L^1_{f,t}\sum_{z=1}^{\Lambda_f}\nu^{1-z}$

The firm's capital stock is:[6]

$$K_{f,t} = \sum_{h=1}^{V_f} k_{h,f}(1-\delta)^{t-\tau_h} \tag{4}$$

5. We assume adaptive rather than rational expectations. Here we assume an inventory/sales ratio that corresponds to the minimum of the observed values (e.g., Bassin *et al.*, 2003; U.S. Census Bureau, 2008).

6. Following Amendola and Gaffard (1998) and Llerena and Lorentz (2004) capital goods define the firm's production capacity and the productivity of its labour.

where V_f is the number of capital vintages purchased, $k_{h,f}$ and τ_h respectively the amount 3of capital and date of purchase of vintage h, and δ the depreciation rate. The firm's productivity embodied in the capital stock then is the average productivity over all vintages purchased:

$$A_{f,t} = \sum_{h=1}^{V_f} \frac{k_{h,f}(1-\delta)^{t-\tau_h}}{K_{f,t}} a_{g,\tau_h} \tag{5}$$

where $a_{g,\tau h}$ is the productivity embodied in the h vintage.

Capital investment is driven by market outcomes and depends on the expected demand

$$k^e_{f,t} = (1+u)\frac{Y^e_{f,t}}{B} - K_{f,t-1}$$

where u is the unused capital capacity. This is equivalent to assuming that if the firm faces a capital constraint (because of an increase in demand or a depreciation of the current stock) it purchases new capital, accessing profits or an unconstrained financial market. Investment then defines the demand for capital good firms: $k^d_{g,f,t} = k^e_{f,t}$. Each firm selects one of the capital producers $g \in \{1;...;G\}$ with a probability that depends positively on g's output embodied productivity ($a_{g,t-1}$), and negatively on its price ($p_{g,t-1}$) and cumulated demand of capital g still has to produce. The delivery of the capital investment may take place after one or more periods, during which the firm cannot make a new investment.

1.1.2. Wage setting, pricing and the use of profits

We model an aggregate minimum wage (w_{min}) as an outwards shifting wage curve (Blanchflower and Oswald, 2006; Nijkamp and Poot, 2005), where unemployment is derived following a Beveridge curve from the vacancy rate (Wall and Zoega, 2002; Nickell *et al.*, 2002; Teo *et al.*, 2004), endogenously determined by firms' labour demand. The minimum wage setting (Boeri, 2009) is related to changes in labour productivity and the average price of goods.[7] The wage of first tier workers is a multiple of the minimum wage, $w^1_{f,t} = \omega w_{min,t-1}$. For the following tiers the wage increases expo-

7. For a detailed description of the computation of the minimum wage see Ciarli *et al.* (2010).

nentially by a factor b which determines the skewness of the wage distribution (Simon, 1957; Lydall, 1959):

$$\begin{aligned} w_{f,t}^{2} &= b\omega w_t^1 = b\omega w_{min,t-1} \\ &\vdots \\ w_t^z &= b^{(z-1)}\omega w_{min,t-1} \\ &\vdots \\ w_t^{\Lambda_f} &= b^{(\Lambda_f-1)}\omega w_{min,t-1} \end{aligned} \tag{6}$$

Price is computed as a markup on unitary production costs (Fabiani *et al.*, 2006; Blinder, 1991; Hall *et al.*, 1997), *i.e.* the total wage bill divided by labour capacity:[8]

$$p_{f,t} = (1+\mu)\frac{\omega w_{min,t-1}}{A_{f,t-1}}\sum_{z=1}^{\Lambda_f} b^{(z-1)}\nu^{(1-z)} \tag{7}$$

The tier-wage structure implies diseconomies of scale in the short-run, which is in line with the literature on the relation between firm size and costs (e.g. Idson and Oi, 1999; Criscuolo, 2000; Bottazzi and Grazzi, 2010).

The profits ($\pi_{f,t}$) resulting from the difference between the value of sales, $p_{f,t-1}Y_{f,t}$, and the cost of production,

$$\omega w_{min,t-1}L_{z,t}^{1}\sum_{z=1}^{\Lambda_f} b^{(z-1)}\nu^{(1-z)},$$

are distributed between (i) investment in new capital ($k_{f,t}^{e}$), (ii) product innovation R&D ($R_{f,t}$) and (iii) bonuses to managers ($D_{f,t}$). For simplicity we assume that firms always prioritise capital investment when they face a capital constraint, while the parameter ρ determines the allocation of the remaining profits between R&D and bonuses:[9]

8. This is in line with evidence that firms revise prices once a year, mainly to accommodate inputs and wage costs (Langbraaten *et al.*, 2008).

9. We are aware of recent empirical evidence which suggest that R&D growth is caused by growth in sales rather than profits (Coad and Rao, 2010; Moneta *et al.*, 2010; Dosi *et al.*, 2006). Indeed, assuming a fixed markup, in our model profits are a constant share of sales. In other words, we would maintain that R&D is related to sales figures but since the model does not include a credit market we prefer to constrain R&D investment by the available resources, *i.e.* profits. Moreover, the model accounts for the case where profits are distributed to managers and not invested in R&D, for a very small ρ, as suggested in some of the cited literature.

$$R_{f,t} = max\left\{0;\rho\left(\sum_{\tau=1}^{t}\pi_{\tau} - \sum_{h=1}^{V_f} k_{h,f}\, p_{g,h}^{K} - \sum_{\tau=1}^{t-1}\left(R_{f,\tau} - D_{f,\tau}\right)\right)\right\} \tag{8}$$

$$D_{f,t} = max\left\{0;(1-\rho)\left(\sum_{\tau=1}^{t}\pi_{\tau} - \sum_{h=1}^{V_f} k_{h,f}\, p_{g,h}^{K} - \sum_{\tau=1}^{t-1}\left(R_{f,\tau} - D_{f,\tau}\right)\right)\right\} \tag{9}$$

This amounts to assuming that (i) firms invest in R&D to seek new sources of revenues, *i.e.* when they perceive a reduction in competitiveness—as no new capital is required; (ii) respond to an increase in demand reducing the resources constraint; and (iii) distribute profits only when this does not interfere with the positive momentum—increase in demand, capital investment, increase in productivity. We assume that bonuses are distributed proportionate to wages, to the manager tiers $\left(z \in \{2;..;\Lambda_{f,t}\}\right)$. The overall earnings of an employee in tier z is then $w^{z}_{f,t} + \psi^{z}_{f,t}$, where ψ^{z}_{t} is the share of redistributed profits to the managers of each tier z.[10]

1.1.3. Product innovation

Firms innovate in two stages: first new products are discovered through R&D, second they are introduced into the market. The R&D activity has two phases: research, *i.e.* the choice of consumer need/market n' in which to focus the innovation effort, and development, *i.e.* the production of a prototype of quality $q'_{f,t}$.

The range of sectors $\left\{n_{f,t}^{min};...;n_{f,t}^{max}\right\}$ that a firm can search is centred on the knowledge base of the current sector of production $n \in N^{+}$ (Nelson and Winter, 1982) and depends on R&D investment and a parameter ι:

$$\begin{aligned} n_{f,t}^{min} &= max\left\{1; n - nint\left(\frac{N}{2}\left(1 - e^{-\iota R_{f,t}}\right)\right)\right\} \\ n_{f,t}^{max} &= min\left\{N; n + nint\left(\frac{N}{2}\left(1 - e^{-\iota R_{f,t}}\right)\right)\right\} \end{aligned} \tag{10}$$

where *nint* is the nearest integer function. Within this set a firm selects the sector with the largest excess demand $Y^{x}_{n,t}$.

10. This assumption is inspired by evidence that the exponential wage structure of a hierarchical organisation is not sufficient to explain earnings disparities (Atkinson, 2007).

The quality of the new prototype developed in sector n' is extracted from a normal distribution where the mean is equal to the quality currently produced by the firm and the variance is negatively related to the distance between the old and the new sectors and positively related to a parameter ϑ:

$$q_{n',f,t} = max\left\{0; q_{n',f,t} \sim N\left(q_{f,t}; \frac{\vartheta}{1-|n-n'|}\right)\right\} \tag{11}$$

If the innovation occurs in n the new good is maintained only if it is of higher quality than the currently produced good and if it represents an incremental innovation in the market n. Otherwise, the new product is discarded. If it is maintained the new good is introduced in a set Φ of prototypes $q'_{\phi,f,t-1}$. If Φ includes less than three prototypes the new one is added. If $\Phi = \{0;...;3\}$ the new prototype replaces the one with the lowest quality as long as its own quality is higher. Otherwise, the new product is discarded.

A firm introduces a new prototype in its market with a probability negatively related to the growth of sales.[11] We assume that a firm introduces in the market its highest quality prototype. We assume also that if a firm's prototype is for a different sector from the one in which it is currently producing, it will be introduced in this other sector only if the number of firms in that sector is lower than in the current sector of production. In other words, a firm moves to a new sector where there is less competition, or introduces a higher quality product in the current sector of production.

1.2. Capital suppliers

The capital goods sector is formed of a population of $g \in \{1,2,\ldots,G\}$ capital suppliers that produce one type of capital good characterised by vintage τ_h and an embodied productivity a_{τ_h}.

1.2.1. Output and production factors

In line with the empirical evidence (e.g. Doms and Dunne, 1998; Cooper and Haltiwanger, 2006) we assume that production

11. For positive growth the probability is 0. We follow the well know Schumpeterian argument that firms innovate to seek new sources of revenues. The probabilistic behaviour captures firms' limited forecasting capacity and distinguishes between temporary falls in sales from long term structural downturns which are more likely to require an innovation.

is just-in-time. Capital suppliers receive orders k^d_{g,f,τ_f} from firms in the final good sectors—where τ_f refers to the date of order—and fulfil them following a first-in first-out rule. The total demand

$$K^D_{g,t} = \sum\nolimits_{f=1}^{F} k^d_{g,f,t} + U^K_{g,t-1}$$

for a capital supplier is then the sum of current orders and past unfulfilled orders

$$U^K_{g,t-1} = \sum\nolimits_{\tau=1}^{t} \sum\nolimits_{f=1}^{F} k^d_{g,f,\tau} - \sum\nolimits_{j=1}^{t} Y_{g,j} \ .$$

For simplicity, we assume that capital producers employ labour as the sole input, with constant returns to scale: $Q_{g,t} = L^1_{g,t-1}$; in each period firms sell the orders manufactured:

$$Y_{g,t} = min\{Q_{g,t}; K^D_{g,t}\} \tag{12}$$

Similar to final goods firms, capital suppliers hire a number of workers necessary to satisfy the demand plus a ratio u of unused labour capacity:

$$L^1_{g,t} = \varepsilon L^1_{g,t-1} + (1-\varepsilon)\left[\left(1+u\right)K^D_{g,t}\right] \tag{13}$$

where ε mimics labour market rigidities. To organise production capital suppliers hire an executive for every batch of v_k production workers $L^1_{g,t-1}$, and one executive for every batch of v_k second-tier executives, and so on. The total number of workers in a firm therefore is:

$$L_{g,t} = L^1_{g,t} + \ldots + L^z_{g,t} + \ldots + L^{A_g}_{g,t} = L^1_{g,t} \sum_{z=1}^{A_g} v_k^{1-z} \tag{14}$$

1.2.2. Process innovation

Capital firms use a share ρ_k of cumulated profits $\Pi_{g,t}$ to hire R&D engineers. The maximum number of engineers is constrained to a share v^K of first tier workers:[12]

$$L^E_{g,t} = min\left\{\nu_k L^1_{g,t}; \rho_k \frac{\Pi_{g,t}}{w^E_{g,t}}\right\} \tag{15}$$

12. See footnote 9 for a discussion of profits and R&D.

The outcome of R&D is stochastic (e.g. Aghion and Howitt, 1998; Silverberg and Verspagen, 2005), and the probability of success depends on the resources invested in engineers and a parameter ζ (Nelson and Winter, 1982; Llerena and Lorentz, 2004):

$$P_{g,t}^{inn} = 1 - e^{-\zeta L_{g,t-1}^{E}} \tag{16}$$

If the R&D is successful[13] a firm develops a new capital vintage with productivity extracted from a normal distribution centred on its current productivity:

$$a_{g,\tau_h} = a_{g,\tau_{h-1}}\left(1 + max\{\varepsilon_{g,t}^{a};0\}\right) \tag{17}$$

where $\varepsilon_{g,t}^{a} \sim N(0;\sigma^{a})$ is a normally distributed random function. The higher is σ^{a} the larger are the potential increases in productivity. The new level of productivity enters the capital good produced by the firm for the following period and sold to the final good firms.

1.2.3. Wage setting, price and profits

The price of capital goods is computed as a markup (μ^{k}) over variable costs (wages divided by output ($Q_{g,t}$)):

$$p_{g,t} = (1+\mu^{K})\omega w_{min,t-1}\left(\sum_{z=1}^{\Lambda_g} b_k^{z-1}\nu_k^{1-z} + \frac{\omega^{E} L_{g,t-1}^{E}}{Q_{g,t}}\right) \tag{18}$$

where $w_{g,t}^{E}$ is the wage of engineers. The first tier wage is a multiple of the minimum wage $w_{min,t}$, such as the wages paid to the engineers ($w^{E}w_{min,t-1}$). For simplicity we assume no layer/manager structure among the engineers. Wages increase exponentially through the firm's tiers by a factor b identical to the final goods firms.

Profits resulting from the difference between the value of sales $p_{g,t}\,Y_{g,t}$ and the costs for workers and engineers

$$\omega w_{min,t-1}\left(L_{g,t-1}^{1}\sum_{z=1}^{\Lambda_g} b_k^{z-1}\nu_k^{1-z} + w_{g,t}^{E}L_{g,t-1}^{E}\right)$$

13. R&D is successful when a random number from a uniform distribution [0; 1] is smaller than $P_{g,t}^{inn}$.

are cumulated ($\Pi_{g,t}$). The share not used for R&D ($1-p_k$) is distributed to managers as bonuses, proportionate to their wages:

$$D_{g,t} = max\left\{0;(1-\rho_k)\,\Pi_{g,t}\right\} \tag{19}$$

where $$\Pi_{g,t} = \sum_{\tau=1}^{t-1}\pi_{g,\tau} - \sum_{\tau=1}^{t-1} w_\tau^E L_\tau^E - \sum_{\tau=1}^{t-1} D_{g,\tau}$$

1.3. Demand

The composition of demand depends directly on the structure of production (product technology, firm organisation and labour structure, and production technology) acting as the endogenous transmission mechanism through which structural changes on the supply side affect changes to consumption.

We assume that each tier of employees in the hierarchical organisation of firms defines one (income) class of consumers with the same income (W_z), consumption share ($c_{n,z}$), and preferences (υ_z^j). This is a restrictive assumption, but also an improvement with respect to models that assume two fixed classes (rural and urban) or homogeneous consumers.

1.3.1. Income distribution and consumption shares

The income of each consumer class $z \in \{0,1,\ldots,\Lambda_t\}$[14] is the sum of wages ($W^w_{z,t}$), distributed profits ($W^\psi_{z,t}$) and an exogenous income ($\bar{W}_{z,t}$):

$$W_{z,t} = b^{z-1} w_{min,t-1}\left(\sum_{f=1}^{F} L_{f,z,t} + \sum_{g=1}^{G} L_{g,z,t}\right) + \sum_{f=1}^{F}\psi_{f,z,t} + \sum_{g=1}^{G}\psi_{g,z,t} + \bar{W}_{z,t} \tag{20}$$

Consumers react to changes in total income, changing total current consumption by a small fraction $\gamma \in [0;1]$ and postponing the remaining income for future consumption (Krueger and Perri, 2005):

$$X_{z,t} = \gamma X_{z,t-1} + (1-\gamma) W_{z,t} \tag{21}$$

14. Where Λ_t is the number of tiers in the largest firm in the market, and $z = 0$ is the class of engineers in capital sector firms.

Consumers divide total consumption across different needs $n \in \{1;...;N\}$, each satisfied by a different sector, and allocate to each need a share $c_{n,z}$. The *desired* consumption per need then is simply $C^d_{n,z,t} = c_{n,z} X_{z,t}$ (where $\sum_{n=1}^{N} c_{n,z} = 1$).

Following the empirical literature on Engel curves we allow these expenditure shares to vary endogenously across income classes, representing a different income elasticity for different income classes and different consumption goods (needs in this model). As we move from low to high income classes the expenditure shares change from "primary" to "luxury" goods at a rate η:

$$c_{n,z} = c_{n,z-1}\left(1 - \eta\left(c_{n,z-1} - \overline{c}_n\right)\right) \tag{22}$$

where $\overline{c}_n$ is an 'asymptotic' consumption share of the richest theoretical class, towards which new classes of workers (with higher income) emerging endogenously tend (see equations 3 and 14). The "asymptotic" distribution is defined as the consumption shares of the top income centile in the UK in 2005 for the ten aggregate sectors (Office for National Statistics, 2006)—which we assume satisfy ten different needs—ordered from smallest to largest (Figure 1).[15] For reasons of simplicity (and lack of reliable data) we assume that the consumption shares of the first tier class, 2000 periods before—the initial period in the model, are distributed symmetrically (Figure 1).[16]

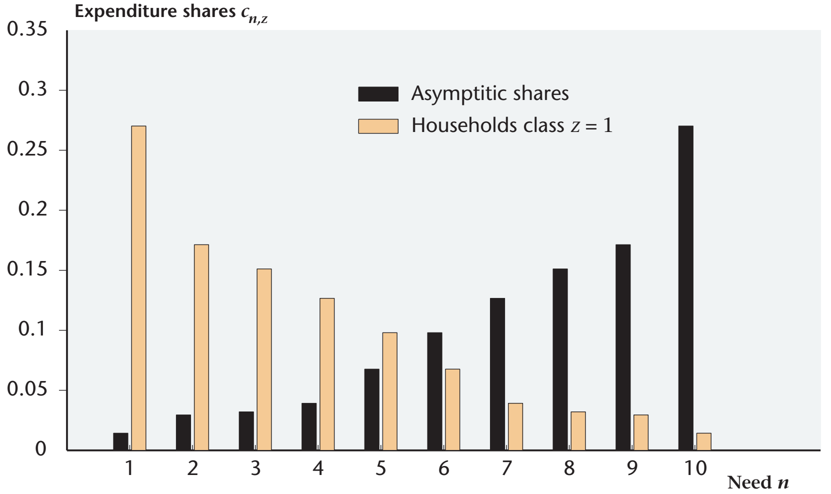
If the goods available on the market satisfy only a limited number of needs—since new goods are discovered through firms' R&D—consumers adapt consumption shares accordingly, redistributing the shares for non available needs to the needs that are available, proportional to the consumption shares of their existing needs. The demand for non available needs is defined as excess demand, which works as the signal for final goods firms to choose the sector in which to innovate:

$$Y^x_{n,t} = \sum_{f_n} Y_{f_n,t} p_{f_n,t} - \sum_z c_{n,z} X_{z,t} \tag{23}$$

15. We thank Alessio Moneta for these data.
16. Madisson (2001) provides qualitative evidence to support this assumption about changes in household expenditure shares.

Figure 1. Expenditure shares: initial and asymptotic

The distribution of the asymptotic level of shares corresponds to current expenditure shares for the highest percentile of UK consumers. For simplicity, initial shares are assumed to be distributed symmetrically



1.3.2. Consumer behaviour and firm sales

We model consumers who purchase a number of goods in each of the available markets with lexicographic preferences. In line with the experimental psychology literature (e.g. Gigerenzer, 1997; Gigerenzer and Selten, 2001) we assume also that consumers have imperfect information on the characteristics of goods, and that they develop routines to match a satisficing behaviour, leading to the purchase of goods equivalent to the optimal good.

Consumer classes access the market in sequence and demand a non negative quantity of goods from each firm. Firm demand is defined as follows. Consumers in a class z are divided into $m = \left\{1, H \in N^{+}\right\}$ identical groups with an equal share of the class income $\dfrac{X_{z,t}}{H}$.

First, a consumer group m screens all the goods on offer from all the firms in the market (need) and observes their characteristics

$$i^{*}_{j,f_n,t} \sim N\left(i_{j,f_n,t}, \sigma^{i}_{j}\right), \forall j = \{p;q\},$$

where σ_j^i measures the extent of incomplete information, which differs for quality and price (Celsi and Olson, 1988; Zeithaml, 1988).

Consumer preferences are modelled here as degree of tolerance over shortfalls with respect to the best good available in the market in terms of its characteristics $\hat{i}^*_{j,f_n,t}$. That is, given the tolerance level $\upsilon_{j,z} \in [0,1]$ a consumer is indifferent towards all of the goods that have a quality above $\upsilon_{2,z}\hat{i}^*_{2,f_n,t}$ and a price below $\upsilon_{1,z}\hat{i}^*_{1,f_n,t}$. In other words, for a very large $\upsilon_{j,z}$ a consumer buys only from the best firm in the market, while a small $\upsilon_{j,z}$ indicates indifference towards a large number of goods that differ in terms of price and quality. We assume also that preferences change across income classes: first tier workers have a high tolerance towards quality differences ($\upsilon_{2,1} = \upsilon^{min}$) and very low tolerance towards price differences ($\upsilon_{1,1} = \upsilon^{max}$). As we move to higher income classes, tolerance towards price differences increases and tolerance towards quality differences reduces by a factor ς:

$$\begin{aligned} \upsilon_{1,z+1} &= (1-\varsigma)\upsilon_{1,z} - \varsigma\upsilon^{min} \\ \upsilon_{2,z+1} &= (1-\varsigma)\upsilon_{2,z} + \varsigma\upsilon^{max} \end{aligned} \tag{24}$$

Then, a consumer group selects the subset of firms that matches its preferences:

$$\hat{F}_{n,m,t} \mid (1-\upsilon_z)\hat{i}^*_{j,f_n,m,t} > \mid i^*_{j,f_n,m,t} - \hat{i}^*_{j,f_n,m,t} \mid, \forall j = \{p;q\} ,$$

and purchases are equally distributed among selected firms. Then, the total demand of a firm in market n is the sum of sales across all groups and classes:

$$Y_{f_n,t} = \sum_{z=1}^{\Lambda_t} \sum_{m=1}^{H_{n,z}} \frac{1}{F^*_{n,m,t}} \frac{c_{n,z} X_{z,t}}{H} \tag{25}$$

2. Methodology

The main aim of this paper is to assess the relative effects of the parameters that define the different aspects of structural change. The model is agent-based and has no analytical solution, but we can study its properties with a systematic numerical analysis. We do so using a simple experimental design. We describe the initialisation of the model, and then the method of analysis.

Table 1 presents the initial conditions and the value of the parameters not included in the DOE.[17] For these parameters we also

Table 1. Parameters setting

Parameter's (1) name, (2) description, (3) value, and (4) empirical data range

Parameter	Description	Value	Data
$\underline{i}_2$	Initial min quality level	98	Analysed
$\overline{i}_2$	Initial max quality level	102	Analysed
a^s	Adaptation of sales expectations	0.9	// [a]
$\overline{s}$	Desired ratio of inventories	0.1	[0.11 - 0.25] [b]
u^l	Unused labor capacity	0.05	0.046 [c]
u	Unused capital capacity	0.05	0.046 [c]
δ	Capital depreciation	0.001	[0.03, 0.14]; [0.016, 0.31] [d]
$\frac{1}{B}$	Capital intensity	0.4	B = [1.36, 2.51] [e]
ε	Labor market friction (final firms)	0.9	0.6; [0.6, 1.5]; [0.7, 1.4]; [0.3, 1.4] [f]
ω	Minimum wage multiplier	2	[1.6, 3.7] [g]
$1\text{-}\gamma$	Smoothing parameter	0.2	[.04, .14]; [.06, .19] [h]
σ^i_j	Error in the consumer's evaluation of characteristics	j = 1: 0.05; j = 2: 0.1	// [i]
ω^E	Engineers' wage multiplier	1.5	[1.2, 1.4] [j]
$\upsilon^{min}=\upsilon_{2,1}$	Highest = first tier quality tolerance	0.1	//
$\upsilon^{max}=\upsilon_{1,1}$	Lowest = first tier quality tolerance	0.9	//
F	Final good firms	100	//
G	Capital good firms	10	//
H_z	Consumer samples	100	//
N	Number of needs	10	//

a) Empirical evidence not available: the parameters has no influence on the results presented here.
b) U.S. Census Bureau (2008); Bassin *et al.* (2003).
c) Coelli *et al.* (2002), with reference to the `optimal' unused capacity.
d) Nadiri and Prucha (1996); Fraumeni (1997) non residential equipment and structures. We use the lower limit value, (considering 1 year as 10 simulation steps) to avoid growth in the first periods to be determined by the replacement of capital.
e) King and Levine (1994).
f) Vacancy duration (days or weeks) over one month: (Davis *et al.*, 2010; Jung and Kuhn, 2011; Andrews *et al.*, 2008; DeVaro, 2005.
g) Ratio with respect to the average (not minimum) wage in the OECD countries (Boeri, 2009).
h) Krueger and Perri (2005); Gervais and Klein (2010).
i) No empirical evidence available to the best of our knowledge. Parameters set using the qualitative evidence in Zeithaml (1988).
j) Relative to all College Graduates and to accountants (Ryoo and Rosen, 1992).

17. The remaining factors are presented in Table 7.

report the data ranges available from empirical evidence. While we are not calibrating the model to any specific economy, all parameters are within the ranges observed across countries and over time.

In $t = 0$ firms produce goods in the first two sectors, and consumers can satisfy only those two needs.[18] Final goods firms differ only with respect to the quality of the good produced, which is extracted from a uniform distribution ($i_2 \sim U\,[\underline{i}_2, \overline{i}_2]$). All capital goods firms are identical. All firms are small, requiring only one manager; capital good firms also hire engineers. This labour structure defines three initial classes of consumers: engineers, first tier workers, and one manager tier.

2.1. Experimental Design

To analyse the effect of the parameters that define the structure of the economy (Table 7) we make use of the simplest DOE, the 2^k full factorial design. It consists of analysing k factors at two different levels (typically High and Low), simulating all possible combinations of both levels (Montgomery, 2001; Kleijnen *et al.*, Summer 2005). 2^k factorial designs are appropriate for the purposes of this paper: to study the main effects of a large number of factors, and to identify the factors that are more influential on the model behaviour from those that are less relevant; to study a large number of interactions of different orders, between factors; and to, at the same time, minimise the number of simulation runs required to study a large number of factors in a complete design (Montgomery, 2001).

In particular, we analyse the effect of the ten factors that define the initial structure of the economy and the scale at which it changes through time. To each parameter we assign "Low" and "High" values (Table 7), which we consider to be the theoretical extreme values (observed infrequently). In Appendix A we provide evidence for the choice of the extreme values.

We test all 2^{10} combinations of Low and High values of the $i = 1, ..., I$ factors. We run 20 replicates for each combination for 2000 periods.[19] We then totalise a sample of factor responses (*i.e.* output variables) y_{ijlt} where $j = \{1, ..., 1024\}$ is the number of

18. The remaining sectors may emerge as a result of firms' product innovation.

designs—combinations of the different parameters, $l = \{1, \ldots, 20\}$ is the number of replicates, and $t = \{1, \ldots, 2000\}$ is the time periods.[20]

We focus on aggregate outputs and analyse the responses using various methods, taking into account the violations of normality and constancy of variance in the responses (Kleijnen, 2008).[21] First, we assess the significance of the factors effect and of their first level interactions with an analysis of variance. Then, we study the relative importance of factors and their interactions, controlling for the effect of a number of variables and using Least Absolute Deviation (LAD) regressions. A graphical description of the impact of each factor on output can be found in the working paper version (Ciarli, 2012).

3. Results

Using the baseline configuration (simulated for 200 replicates) the model generates long term endogenous growth in output with a transition from linear growth to exponential growth (Figure 2 (a))—occurring here around $t = 1400$ —(Maddison, 2001; but also Galor, 2010). Output growth is preceded by an increase in aggregate productivity. The linear growth is characterised by very low investment rate, that induce slow changes in productivity, and is driven by the final demand—*via* slow grow in firm size, and demand for labour.

The transition to the second phase occurs as heterogeneous firms emerge—due to the acquisition of slightly different vintages and their own innovation—and the linearly increasing working population selects the best firms. Selection induces large changes in the demand for few firms, that require large investment in new, more productive, capital. Demand for new capital, in turn, spurs innovation in the capital sector, which supply even more produc-

19. Our model is a non terminating simulation, which requires us to choose a cut-off point when the simulation enters a "normal, long run" regime (Law, 2004). For some responses, such as output, under a large number of factorial combinations the model does not reach a steady state. For others, such as output growth and market concentration, the model reaches a "long-run steady state" before 2000 periods.

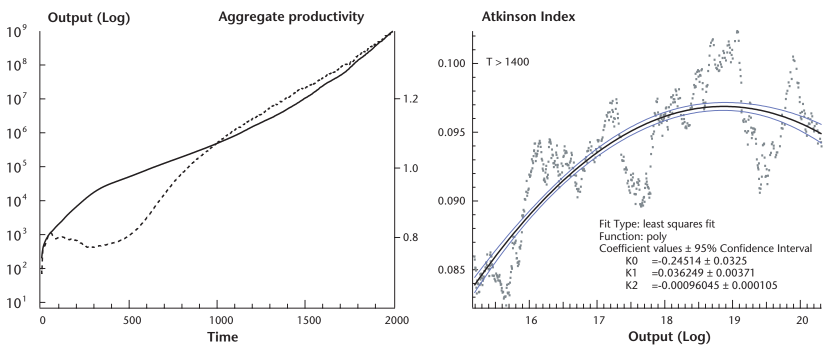
20. 20,480 simulation runs and 40,960,000 observations.

21. Our model and simulation procedure satisfy the remaining properties outlined in Kleijnen (2008). See also Montgomery (2001) for a comprehensive treatment of the analysis of experiments in simulations.

tive capital goods. This starts a cumulative causation process characterised by decreasing prices and increasing consumption, profits, investment, and tiers of workers, which induce more demand heterogeneity, at the higher inequality (Ciarli *et al.*, 2010). Indeed, we provide evidence of the often observed non-linear relation between inequality and income (Kuznets curve) for the period from before take off to the end of the simulation (Figure 2 (b)).[22]

Figure 2. Model properties: output (log), productivity and Kuznets curve

In graph (a) we plot the time series of output in log scale, and aggregate productivity in linear scale. In graph (b) we plot the relation between output (log) and the Atkinson index (dotted line), and the polynomial curve fit (full line) with confidence intervals (blune lines), for the period 1400-2000



(a) Output and productivity **(b) Kuznets curve**

For given values of the extent of exploration of new sectors, the model qualitatively reproduces the s-shaped curve characterising the growth in sectoral output from birth to diffusion in the economy (Figure 3). Both figures show that at this level of aggregation, sectors are not expected to disappear (as it would be the case with goods). Both figures also show that the emergence of new sector is concentrated in a relatively short time span. However, the simulated results show a higher concentration of emergence, due

22. Inequality is computed using the Atkinson index:

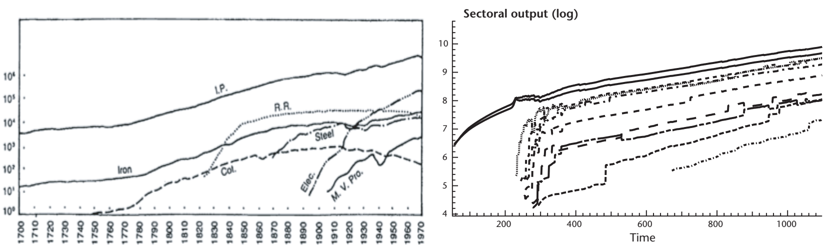
$$\mathcal{A}T_t = 1 - \frac{1}{\sum_{z=1}^{Z} \frac{W_{z,t}}{L_{z,t}}} \left[\frac{1}{\sum_{z=1}^{Z} L_{z,t}} \sum_{z=1}^{Z} \left(\frac{W_{z,t}}{L_{z,t}} \right)^{1-\varrho} \right]^{\frac{1}{1-\varrho}}$$

where $W_{z,t}$ is the total income of consumer class z, $L_{z,t}$ is the total number of workers in class z and ϱ is the measure of inequality aversion. As we are not measuring an empirical level of inequality, we use an intermediate value of $\varrho = 0.5$.

to the fact that we have a fixed number of needs—unlike real sectors, that firms attempt to satisfy as soon as they manage to search all sectors: a lower ι would imply a less clustered emergence (see equation 10).

Figure 3. Sectoral output: industrial production in Britain and simulation results

(a) Sectoral output (log scale) computed by Rostow (1978) (cited in Aoki and Yoshikawa, 2002); (b) sectoral output (log) from the model results with ι = 0.005



(a) Observed **(b) Simulated**

Table 8 in the Appendix reports the results of a Vector Autoregressive (VAR) analysis on 10 period growth rates, and coefficients estimated using LAD—with bootstrapped standard errors. The VAR shows the relations between output (1), aggregate productivity (2), average price (3), the inverse Herfindahl index (4) and the Atkinson index of inequality (5). Results are in line with the expected macro dynamics. All variables show a strong cumulative process with one lag. Inequality growth has an immediate positive effect on output (1 lag) which becomes negative after three lags. Similarly, an increase in market concentration has an immediate negative effect on output (1 lag), which becomes positive after two lags. Market concentration also determines an increase in prices and inequality. The effect of productivity on output in this short-run analysis is captured through price reduction, which has an immediate positive effect on output (1 lag). A detailed discussion on the short and long run dynamics of a previous version of this model can be found in Ciarli *et al.* (2010) and Ciarli *et al.* (2012).

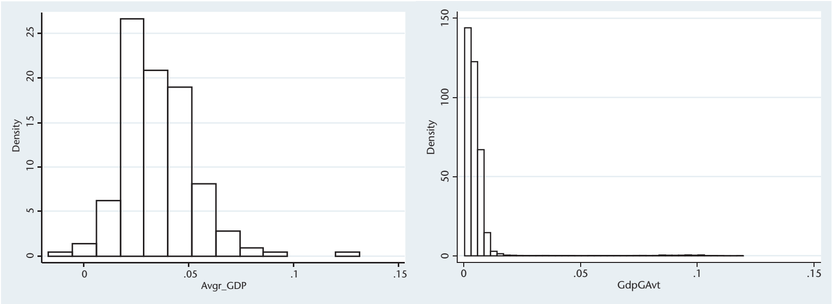
3.1. Distribution of income across countries

We now move to the analysis of the model for the 2^k combinations of factors. Each combination of factors in the model can be interpreted as a different country with different initial conditions.

We thus compare the distribution of the average growth rate of GDP from 1980 to 2010 across countries using IMF data, with the distribution of average growth rate of output from period 1 to 2000 across factorial designs using the simulated data (Figure 4). The distribution is definitely more skewed in our simulations across different combination of factors than across world economies. There are three main reasons for this. First, a trivial one: we look at 2000 periods, which includes long periods of stagnation that precede take-off (see above), while the IMF data refer to the the period between 1970 and the present.

Figure 4. Income growth distributions—world Vs simulation

Graph (a) plots the distribution of income growth (averaged over 1980-2010) across world economies (Source IMF); graph (b) plots the distribution of output growth (averaged for 1-2000) across different combinations of factors

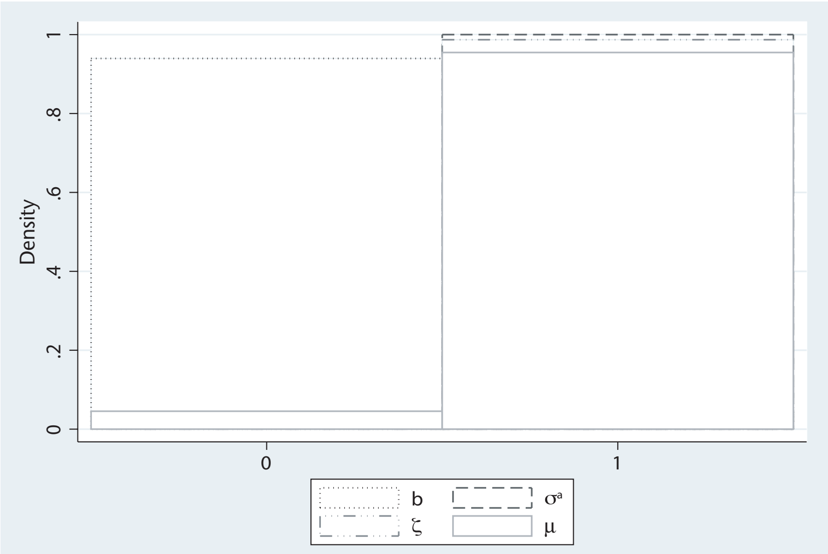


(a) Observed **(b) Simulated**

Second, we are analysing the model under extremely "stressful" conditions, *i.e.* for extreme values of the parameters not generally observed in the real world (see Table 7). For example, for some factor combinations no investment occurs, and the economy stagnates over the 2000 periods. To show the relevance of these extreme conditions on the distribution of output we show the probability that a low or a high value of the parameters occurs for designs with very high income. We analyse the following parameters: the variance of the distribution that determine an increase in the productivity of capital (σ^a), the wage differentials between classes (b), the probability of process innovation (ζ), and the markup (μ). Figure 5 plots the density of these four parameters when the level of output (log) is larger than 36, the top bin of the world income distribution according to IMF data. For these extremely high levels of output we observe almost one single combination of factors: σ^a is high with probability 1,[23] b is low with probability around .95, ζ is high with probability nearly 1 and μ is high with probability .95.

Figure 5. Density of some parameters when Log output > 36

0 denotes a low value of the parameter and 1 denotes a high value. For extremely high values of output the parameters take almost always the same value



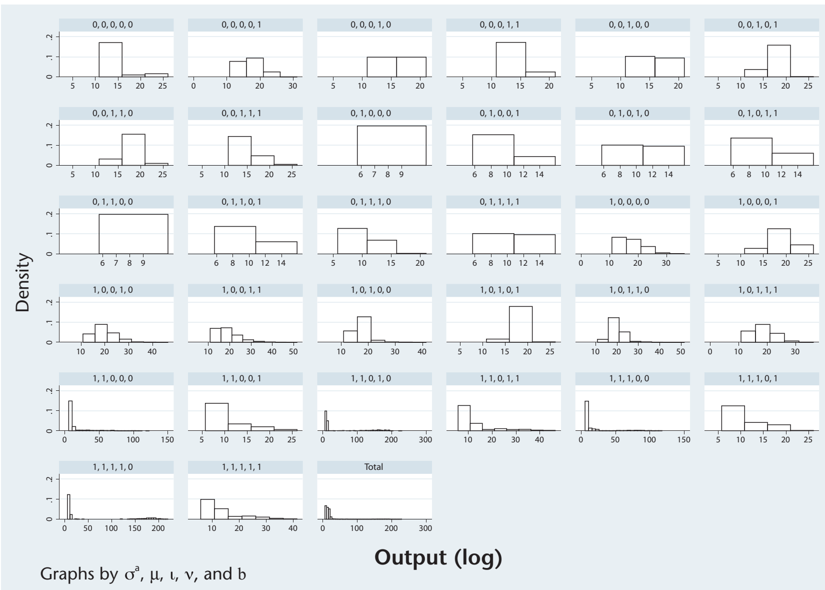
The third related reason for the atypical distribution of output growth in our model simulated across the 2^k different factor combinations is due to the DOE: in Figure 4 (b) we are overlapping distributions from different data generation processes, where each combination of the High and Low values of the parameters represent one process. We show below that some parameters have a dramatic effect on the output variables. The distribution of output variables differs enormously when these parameters switch from one state to another. We show this again by comparing the distributions in the simulated data. Figure 6 plots the distribution of output for different combinations of some influencing parameters with High and Low values. It is sufficient to compare the support

23. We denote the low level of the parameters as 0 and the high level of the parameter as 1.

of the total distribution (the last one in the figure, for all values of the parameters) with the support of any other distribution which represents a combination of different High and Low values (1,0) of σ^a, μ, ι, v, b. In only one case the support is the same. In most cases, the support is radically different (lower by a factor 10 or 30).

Figure 6. Distribution of output (log) for different High and Low values of the parameters

σ^a, μ, ι, v, b ; 0 denotes a low value and 1 denotes a high value



To sum up, the overall distribution of output variables, such as output in the final period, and the average rate of output growth over periods, cannot be approximated satisfactorily by any theoretical distribution that we know of: the closest would be the Pareto distribution.

The above discussion suggests that economies endowed with different factors that determine the initial structural conditions and the way in which structural changes in different aspects of the economy unfold (or not), experience very different growth paths. By testing extreme values of these conditions we see that a limited number of economic aspects—different from the beginning—

produce dramatic differences in growth. Also the way that these different aspects of structural change interact seems to be relevant.

The rest of the paper provides a detailed analysis of the (main, interactive, direct and indirect) effects of these factors on the final distribution of output across economies with very different starting conditions.

3.2. The factors of structural change

Before analysing factor responses we briefly summarise the effect of the different factors (parameters) that define the initial structure of the economy and the dynamics of structural change.[24] We group them with respect to the aspects of structural change they capture directly: product technology, production technology, organisation of production (which refer mainly to the structure of employment), income distribution and consumption patterns. Table 7 summarises the "Low" and "High" values, the main aspects affected, and the equation where they appear. Appendix A provides detail for the choice of the "Low" and "High" values.

Table 2. Effect of parameters on structural change

A "+" indicates that the High value of the parameter induces relatively more structural change

Factor	Equation	High/Low	Low	High	Main Economic aspect [a]	Main indirect aspects [a]
ι	10	+	.001	.3	3	_
$\boldsymbol{v}$	3	_	3	50	1	4, 5
$\boldsymbol{b}$	6	+	1	3	1	4
σ^a	17	+	.01	.2	2	1
η	22	+	.1	3	5	_
$\rho = \rho^k$	8, 15	+	.05	.95	3, 2	4
ϑ	11	+	.01	10	3	_
ζ	16	+	.1	1000	2	1
ς	24	+	.05	.9	5	_
$\mu = \mu^K$	7, 18	+	1.01	2	4	2,3, 1

[a] 1: Organisation of production; 2: Production technology; 3: Product technology; 4: Income distribution; 5: Consumption patterns

24. The terms factors and parameters are used interchangeably.

Product technology: ι, ϑ, ρ. All these factors have an effect on the variety in the final goods market, ι determining the pace at which new goods are discovered, ϑ influencing the rate of change in the quality of new goods and ρ altering the resources employed by the firm for the exploration of new goods. ι and ϑ play no other role in structural change; ρ influence the distribution of income through the share of profit not redistributed as bonuses and used for R&D.

Production technology: σ^a, ζ, ρ^k. Large values of σ^a and ζ directly modify the capital structure of the economy, determining the pace at which innovation occurs in the capital goods sector; ρ^k has the same effect as ρ in altering the resources devoted to R&D. All three factors have a number of indirect effects on other aspects of structural change. Similar to ρ, ρ^k influences the income distribution. σ^a and ζ in addition to altering the productivity of the final goods firm, modify the demand for production factors (labour and capital), affecting firms' labour structure (through changes in size). Also, given the different pace at which different firms change capital vintages, σ^a and ζ change the distribution of prices in the final goods market, allowing consumers to select based on their price preferences.

Organisation of production: v and b. Both parameters define the way in which a firm is organised: a very low v means that a corporation needs a large number of tiers to organise a small pool of workers, whereas for a large v a single manager can deal with a large production unit (few changes as the size of the corporation increases). b tells us simply how the different levels of workers and managers are paid (and bonuses are distributed). Both parameters have a strong bearing on the distribution of income as they implicitly determine the number of income classes (v) and their wage income. Indeed, v indirectly influences also at least one other aspect of the economy—changes in consumption patterns—by altering the pace at which new classes with different consumption styles endogenously emerge.

Income distribution: μ, μ^k. For a small ρ and low capital investment a large μ implies a redistribution of income from all consumers to higher classes. Indeed, an increase in μ also increases the resources available for investment in R&D—thus it increases the pace at which product and production technology change.

Finally, differences in markups indicate different market structures, from competitive to oligopolistic.

Consumption patterns: η, ς. High level of both factors induce faster change in consumption behaviour. A high η implies a very fast change in expenditure shares from basic needs to the asymptotic distribution that of the top income centile of UK consumers in 2005. ς changes the consumer preferences in a given classes, for a given expenditure share: a large ς implies that the tolerance for relatively lower quality (higher price) goods decreases (increases) at a faster rate moving towards the high income classes.

Main and cross effects (without normalising the scale). A simple graphical analysis explaining the main and the cross effects of the factors is detailed in Ciarli (2012). Given the scale effect that underlies these results, in this paper we focus on the analysis of variance (next section). However, a summary of the entity of the effect of each parameter—which, including its scale, reflects the fact that the parameters represent very different dimensions of structural conditions—is a useful complement to an analysis of variance that informs on the significance of the effects, but not on their magnitude (more in Section 3.4).

For simplicity, we explored two extreme cases, out of the thousand possible states of the world analysed in this paper: we analysed the impact on total output of each factor when all other factors are either Low (L) or High (H). We found that, in the case where parameters induce low structural change in all dimensions of the economy, a single factor inducing high structural change is sufficient for a strong effect on output. However, this does not apply to all factors and especially not to those that determine changes in the structure of consumption: the wage multiplier (b) and a higher variance of the productivity shock (σ^a) have the strongest positive effects, followed by the exploration of new goods (ι); while ρ (share of profits invested in R&D) and μ (markup) have the strongest negative impact. On the other hand, if all the parameters induce high structural change in all dimensions of the economy, then just two of the parameters inducing low structural change will have an effect on output (ρ and μ).

We found also that the effect of each parameter in many cases is not monotonous: the signs of the main effects change if some of the other factors change from inducing low to inducing high structural change. For example, we analysed the effect of a more or less complex organisational structure (v) under varying structural conditions, such as wage regimes (b) the likelihood of inventing a new product (ι), and increases in the productivity of capital vintages (σ^a). We found that, while a few factors do not interact with v – η, ϑ, ς, and ζ, some induce only a level effect—ρ and μ—, ι, b, and σ^a change the sign of the effect of v: when they are Low, an increase in v has a mild positive effect on output; when they are High, an increase in v has a strong negative effect on output. This seems to suggest that highly complex organisations (which require many organisational layers, *i.e.*, many employees receiving different levels of remuneration, for a given number of workers in the first tier) have a positive impact on output growth when markets diversify quickly (High ι), when firms can recover the higher organisational costs (reflected in higher consumer prices) through increased productivity growth (High σ^a), and when wages differ between organisational layers (b). The rapid vertical growth of firms in fact creates classes of workers with different consumption shares and different preferences, *i.e.* consumers that buy more goods from markets that firms still need to discover (with a High ι) and that are ready to buy goods at higher prices. However, the higher organisational costs translate into lower aggregate demand. The net effect on output is positive only if either product innovation brings results in rapid time to market for goods to satisfy the emerging classes of consumers, or when rapid change in product technology compensates for increasing prices (or possibly if both conditions hold).

3.3. Analysis of variance: the significance of factors' interactions

In order to assess the statistical significance of the effect of each factor, and the joint significance of the different factors on output, we run an ANOVA on 20 replicates for each combination of parameters. The results in Table 3 show that apart from η and ς—respectively the speed of convergence of the expenditure shares and of the change in the preferences of consumers for a good's characteristics—all parameters have a significant main effect, even when tested jointly—when considering the effect of each parameter for all possible states of the world (High and Low values of the other factors).

Table 3. ANOVA – main effects

Source	Partial SS	df	MS	F	Prob>F
Model	1.258e+06	9	139790	414.1	0.00
ι	4840	1	4840	14.34	0.00
v	101546	1	101546	300.8	0.00
b	114912	1	114912	340.4	0.00
σ^a	260782	1	260782	772.6	0.00
η	691	1	691	2.05	0.15
ρ	481609	1	481609	1427	0.00
ϑ	1399	1	1399	4.150	0.04
ζ	240068	1	240068	711.2	0.00
ς	0.821	1	0.821	0	0.96
μ	52954	1	52954	156.9	0.00
Residual	6.909e+06	20469	337.5		
Total	**8.168e+06**	**20479**	**398.8**		

Number of obs = 20480; Root MSE = 18.37; R-squared = 0.154; Adj R-squared = 0.154

Due to the blatant departure from normality of the output variable, we check the robustness of the results of the ANOVA by testing one way differences between the samples defined by the parameters High and Low values with a Kruskal-Wallis (KW) equality of population rank test. The results (see Table 9 in the Appendix B) differ from the ANOVA only with respect to η, which turns out to have a small but significant negative effect.[25] The one way mean test confirms the results from the graphical analysis: high values of ι, v, b, σ^a, ϑ and ζ are associated with a higher output; high values of ρ and μ are associated with a lower output; and η and ς have a negligible effect.

This similarity makes us confident that the results of the ANOVA for our large sample are informative, and we proceed to analyse the significance of the first order interactions between all factors. For instance, as discussed above, ι, b, and σ^a modify the effect of the organisation structure on output. We analyse these interactions more systematically in Table 4, which summarises the

25. However, note that the KW test is one way.

results of the ANOVA that includes all the main effects and first order interactions (*i.e.* all possible interactions between two different factors).

Table 4. ANOVA – first order interactions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	ι	v	b	σ^a	η	ρ	ϑ	ζ	ς	μ
ι	0									
v	**	***								
b	***	***	***							
σ^a	0	***	***	***						
η	**	0	0	0	0					
ρ	***	***	***	***	0	***				
ϑ	0	0	0	0	0	**	0			
ζ	**	***	***	***	0	***	0	***		
ς	0	***	*	0	0	0	*	0	0	
μ	**	***	***	***	0	***	0	***	**	***

Note: Values on the diagonal refer to the factor main effect.
***Prob > F < 0.01; **Prob > F < 0.05; *Prob > F < 0.1; 0: Prob > F > 0.1

Table 4 confirms the intuition—from the analysis of distributions of output (Figure 6)—that most factors induce structural changes that have a significant effect which differs (in size or direction) for different combinations of the other factors, that is, which is subject to the structural changes induced by the other factors. In other words, the different dimensions of structural changes induced by the factors, significantly interact in determining the aggregate behaviour of the economy.

For example, what would be the effect on growth of increasing the opportunities for R&D in the capital sector (σ^a)? As shown in Table 3, σ^a alone has an apparent impact on output. However, Table 4 shows that the role of the production technology crucially depends on many other structural aspect of the economy, such as the organisation of production (v, b), and especially the share of profits invested in R&D and its effectiveness on the innovation result (ρ, ζ). Its strong effect on output is relatively independent of the introduction of product variety in the consumer market (ι, ϑ) and of the structure of demand for more variety (η, ς).

To better discriminate among the different aspects of the structure of an economy, in the next section we perform a regression analysis that allows us to quantify the effect of factors across designs.

3.4. The relative influence of the different aspects of structural change

We run quantile regressions (with bootstrapped standard errors) to estimate the relative impact and significance of each factor and their first order interactions on output. We distinguish between direct and indirect impact: Table 5 reports estimates for the factors (1), for a number of control variables, most of which are correlated to the factors (2), and for the parameters when the least correlated control variables are included—a sort of reduced form of the model (3).[26] Table 6 reports the estimates for the parameters and their first order interactions, with and without control variables (respectively bottom-left and top-right triangular matrix).

On average, when abstracting from the different structural change regimes (column 2), the model shows that aggregate labour productivity ($\mathcal{A}$)—measured as output per worker—is strongly and positively correlated to output as well as average expenditure on R&D (across firms and time periods) (R). As noted elsewhere (Ciarli *et al.*, 2010; Ciarli and Lorentz, 2010), product variety (averaged over the full period)— σq and σp, respectively for quality and price—and the selection that they enable also positively affect output; but their effect is weak and non-significant when controlling for inequality and productivity, which is the prime cause of price differences. Inequality ($\mathcal{AT}$), averaged over the whole period, has an overall negative effect on output.

With reference to factors (column 1), ρ and μ determine structural changes with the strongest (negative) effect on output, followed by σ^{α}, ι and b. Related to σ^{α}, ζ also has a positive and significant effect. The least relevant are the structural changes induced by v, ς (positive) and η (negative).

The factors determining structural change also influence the dynamics of a large number of variables. Therefore, the estimated effect of a variable on output (Table 5) is likely to differ for different levels of the parameters. Likewise, the use of control variables in the estimation of parameters allows us to estimate the direct effect of the factors on output, depurated from the indirect effect through the control variables.

26. The estimated sample is the result of the simulations for the last period for 20 replicates of each combination of parameters.

Table 5. The relative impact of factors and main variables on output

LAD estimates with s.e. obtained from bootstrapping (400); the dependent variable is (Log) output

Variables	(1) Factors	(2) Contr Var	(3) F & CV
ι	0.692***		1.063***
	(0.056)		(0.071)
v	0.009***		-0.012***
	(0.000)		(0.000)
b	0.107***		-0.061***
	(0.008)		(0.007)
σ^a	3.242***		0.966***
	(0.088)		(0.083)
η	-0.023***		-0.016***
	(0.006)		(0.004)
ρ	-4.900***		-3.947***
	(0.024)		(0.036)
ϑ	0.013***		0.003**
	(0.002)		(0.001)
ζ	0.001***		0.000***
	(0.000)		(0.000)
ς	0.040**		0.021*
	(0.019)		(0.011)
μ	-9.330***		-9.510***
	(0.018)		(0.021)
A		1.201***	2.900***
		(0.071)	(0.057)
$\mathcal{A}T$		-3.809***	3.523***
		(1.109)	(0.109)
σp		0.119*	-0.092***
		(0.065)	(0.004)
σq		0.001	0.000***
		(0.001)	(0.000)
R		0.779***	
		(0.048)	
Constant	28.301***	12.944***	29.424***
	(0.043)	(0.380)	(0.075)
Observations	20,480	20,480	20,480
Pseudo R^2	0.43	0.09	0.48

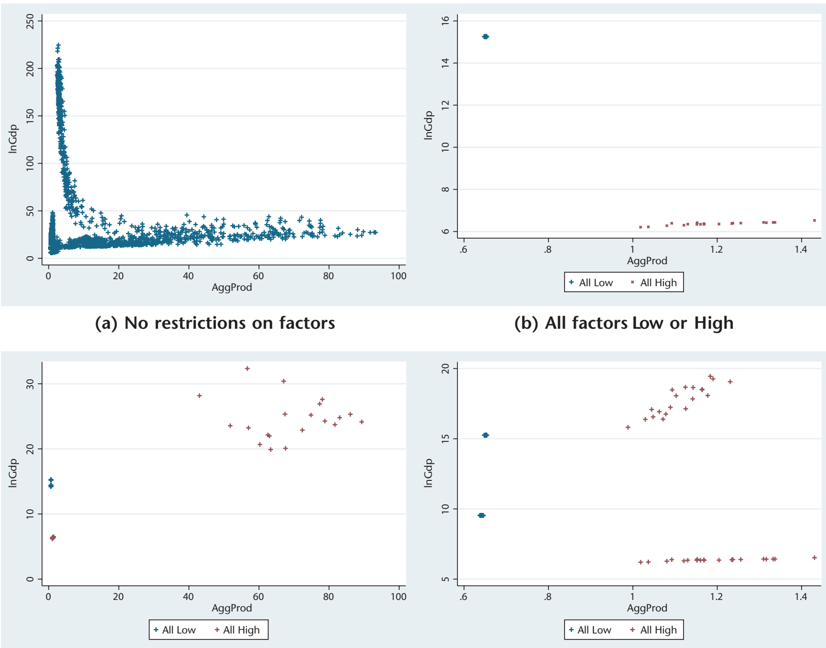
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

To better show this interaction between factors and economic variables, Figure 7 plots the relation between an independent variable, aggregated productivity, and output (Log), for different values of the parameters. In panel *a* no restrictions on parameters are imposed, and the relation between the two variables is distinctly non-linear and non monotonic. More importantly, panel *a* suggests that the relation between productivity and output could be reduced to different functional forms, depending on the combination of factors. In panel *b* we restrict all factors to either the high or the low value: aggregate productivity does not show any significant effect on output under either condition. Results are different if we allow the parameters that are strongly related to aggregate productivity to fluctuate. In panel *c* all factors are either high or low, except for ρ, which takes both values. Although the relation between the two variables is not so clear cut, the scale is radically different—the small dot in the bottom left of panel *c* is the flat relation that we observe at the bottom right side of panel *b*. Alternatively, for different levels of markup (μ) the relation turns from null to positive, for high values of all other parameters (panel *d*).

Returning to Table 6, column (3) shows estimates for the direct effects of the factors and for the effect of the least correlated variables. First, the sign of the two factors defining the organisation of production are inverted. A high v means a lower number of (organisational) workers per good produced, increasing labour productivity. When we control for labour productivity, though, the lower number of tiers for a given number of shopfloor workers reduces the pace at which firms grow in size and diversify by adding different levels of workers and managers (see Eq. 3 and 14). The effect on structural change is a slower increase in the aggregate demand and its variety, and a negative impact on output growth. While large wage differentials (b) increase inequality, which has a negative effect on output, *ceteris paribus*, but which in our model is associated also with larger aggregate demand (Eq. 20); thus the inverted sign of $\mathcal{AT}$ when controlling for the factors. Second, the direct effect of σ^{α} is strongly reduced when controlling for aggregate productivity. Third, the estimated effect of a few variables change their sign and significance as we control for different combinations of the factors determining structural changes.

Figure 7. Aggregate productivity Vs output (Log)

T=2000, aggregate productivity on the x-axis and output on the y-axis



(a) No restrictions on factors

(b) All factors Low or High

(c) All factors Low or High, except ρ

(d) All factors Low or High, except μ

As already mentioned, the effect that each of the factors induces on structural change depends also on other structural aspects. The relevance of the interaction among several factors is established in the results of the quantile regression where all first order interactions are estimated together with the main effects, with and without the control variables Table 6. We estimate the effect of the high value of factors, the low value being the reference case. Estimates without control variables are reported in the top-right triangular matrix and those that include control variables are reported in the bottom-left triangular matrix. On the diagonal we report the main effects (when controlling for interactions). In the following we discuss the estimates obtained when including the control variables (bottom-left triangle). First, the results in Table 6 show that the main effects are strongly significant, despite the inclusion of the interaction terms. Second, most of the first order interaction terms are also significant, particularly when they include a factor shown to have a significant main effect. We proceed by discussing the effect of the different aspects of structural change, following the classification in Table 2.

Table 6. LAD regression – the effect of first order interactions on output

The top-right triangular matrix shows estimates without control variables; the bottom-left triangular matrix shows estimates with control variables.
The dependent variable is output (Log)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	ι	ν	b	σ^a	η	ρ	φ	ζ	ς	μ
ι	1.51	0.16	-0.09	-0.31	-0.01	-0.55	0.03	-0.18	0.04	-0.97
	1.38									
ν	0.28	0.33	-0.95	0.69	0.23	-0.56	-0.08	0.62	0.12	0.20
		-0.81								
b	-0.10	-0.16	1.14	-0.24	-0.15	0.11	-0.05	-0.34	-0.28	0.10
			0.10							
σ^a	-0.37	0.68	-0.34	1.51	0.00	-1.58	0.09	2.26	-0.12	-0.73
				1.39						
η	0.04	0.28	-0.21	-0.03	-0.18	0.00	-0.02	-0.04	0.08	0.10
					-0.26					
ρ	-0.52	-0.35	0.31	-1.00	0.00	-0.56	-0.57	-1.42	0.01	-2.51
						-0.98				
ϑ	-0.05	0.01	-0.08	0.18	0.00	-0.61	0.55	0.10	0.20	-0.17
							0.58			
ζ	-0.14	0.42	-0.34	1.18	0.03	-0.88	0.08	1.24	-0.14	-0.39
								1.05		
ς	0.05	0.23	-0.20	-0.03	0.11	-0.06	0.20	-0.12	0.17	-0.37
									0.10	
μ	-0.98	-0.24	0.37	-1.27	0.12	-1.61	-0.56	-0.96	-0.31	-6.06
										-6.22

Note: Values on the diagonal refer to the factor main effect. Standard errors computed with 400 bootstraps. Reference case is the low value of factors.

$p<0.01$	$p<0.05$	$p<0.1$

Product technology: ι, ϑ, ρ. The effect on output of rapid emergence of new sectors is large and positive on average, but is always negative when other factors induce strong structural change (with the exception of the factors defining consumption patterns, η and ς). That is, for fast changes in consumption shares and/or preferences the high level of ι has a (weakly) significant positive effect. We note also that the interaction between the two main factors of product technology, ι and ϑ—respectively needs and quality—is not significant.

Production technology: σ^a, ζ, ρ^K. The main factors that affect changes in production technology, σ^α and ζ, also have a pervasive effect, interacting with all other aspects of structural change, except those that define changes to consumption patterns. In particular, the relation with organisational factors is negative: for an increase in organisational complexity—*i.e.* reduction in v—or wage differences, fast changes in technology reduce output growth. Taken together with the negative sign of the interaction with ρ, this result suggests that rapid changes in the productivity component of capital have a negative impact on output in the presence of high markups, large wage differences and multiple organisational tiers which amplify wage differences. In other words, economies that experience fast changes in productivity grow at a lower pace if the structure imposes as well an unequal distribution of the productivity gains. The interaction with product technology is negative with respect to the discovery of new sectors, and positive with respect to changes in product quality. Finally, the two main factors affecting changes in production technology— σ^a and ζ—are strongly and positively related to increased output.

Organisation of production: v and b. A large number of layers required to manage an organisation (low v) and large wage differences (b) between these layers have a negative impact in the presence of strong structural change in all other aspects except those affecting income distribution, μ. Even in the presence of rapid changes in consumption patterns, higher wage inequality reduces output growth.

Income distribution: μ, μ^k. A large markup generating high profits in our model has a negative impact under almost all aspects of strong structural change, except for the organisational aspects discussed above, and changes in expenditure shares.

Consumption patterns: η, ς. As already noted, the factors determining structural changes in consumption play a minor role in output growth. The interactions with many other aspects of structural change are not significant: the main effects of consumption patterns are visible when structural change occurs in product technology and the distribution of income. Indeed, the emergence of new products is a requirement for changes in consumption to have a role on output. The two main aspects of consumption patterns, η and ς interact positively.

4. Discussion and conclusions

A large wealth of research has investigated why regions starting from a similar level of output grow so differently through time. Which are the initial apparently small differences that diverge so markedly? In this paper we build on the idea that the initial differences that determine growth divergence are those that define the structure of an economy and the way in which this evolves through time. We maintain that structural change occurs in different aspects of the economy, such as the structure of production, consumption, labour organisation and income distribution. These different dimensions of structural change interact in a continuous evolutionary process, and are not independent from the growth pattern. For example, one could not think at the industrial revolution without considering aspects such as changes in knowledge and technology, changes in the patterns of consumption, changes in trade patterns and extraction of resources through colonial power, and institutional changes. In turn, changes in knowledge and technology may be driven by changes in the demand, as well as by changes in labour relations (e.g. increase in the cost of labour). Similarly, changes in transportation and military technologies are not independent from changes in technological knowledge. Changes in consumption patterns are also related to increased trade. And so on.

We propose a model of the microeconomic dynamics of structural change and their interactions, abstracting from the institutional aspects. The model defines the following related aspects of structural change: organisation of production, technology of production and the emergence of new sectors on the supply side, and income distribution and consumption patterns on the demand side. Ten different parameters (factors) in the model account for these five dimensions of structural change.

In this paper we simply investigate the relative importance of each aspect of structural change in economic growth. We also acknowledge that some of these aspects are likely to have a sizeable impact on economic growth only when complemented by other aspects of structural change. For example, changes in consumption patterns may not occur if there is no change in the sectoral composition of the economy, or if there is no change in wages. We therefore use the model to study the interaction among the different aspects of structural change, analysing factors' interactions.

To do so, for each of the ten parameters defining the different aspects of structural change we assign a high and a low value, respectively identifying large and small structural change as the economy grows. We then define a DOE that accounts for all possible combinations of the parameters' high and low values.

In other words, we define 2^{10} different economies, all starting from the same initial conditions except for one of the aspects of structural change. On one extreme we have an economy that experiences negligible structural changes in all economic aspects, and on the other extreme an economy that experiences large structural changes in all economic aspects. In between are all other possible combinations. We study the impact of the five different aspects of structural changes with an analysis of variance, and running quantile regressions on the cross-design—cross-country—sample.

We find that almost all aspects of structural change are significant determinants of the differences in the growth rate of output across designs, but their magnitude varies substantially. Income distribution, rate of change in production technology and the emergence of new sectors, explain a great deal of output growth differences, changes in consumption patterns—shares and preferences—are barely significant, and changes in the organisation of production—organisation and compensation of labour—lay in between. Moreover, the most relevant factors of structural change play a determinant role even in the presence of negligible structural changes in all other economic aspects. While the opposite case is rarely true: the output growth of economies that are set to experience large structural changes in all aspects except for one, is not affected if one of the changes is only negligible.

Concerning the complementarity among different aspects, we find that most factors of structural change strongly interact. This is an extremely relevant result suggesting that we should always account for a large number of economic aspects to understand long term patterns of across-countries divergence. For example, we find that changes in the composition of sectors available in the market become more relevant when consumption patterns also experience strong structural changes. Or, technological progress is strongly relevant, but quick technological advances have a negative effect on output growth when they are accompanied by structural changes in the organisation of labour that leads to large inequality among a large number of emerging classes. In other words, economies that experience fast changes in productivity, grow at a lower pace if the structure imposes as well an unequal distribution of the productivity gains among workers.

Taken together, the model and the analysis of the experimental design we proposed in this article clearly show that to explain long term growth we need to look at the way in which a large number of structural changes interact at the microeconomic level. It is not unlikely to find that some aspects of structural change were determinant for some regions, and detrimental for others (such as the pay structure at the beginning of the industrial revolution in Europe). Moreover, while some of the interactions between the different aspects are relatively known, many interactions instead suggest avenues for future research on the relation between growth and structural change.

For example, the results show that inequality is related to output growth in a number of different ways through the different aspects of structural change. This is in line with the lack of agreement, on theoretical and empirical grounds, on the relation between inequality and GDP growth (Aghion *et al.*, 1999; Eicher and Turnovsky, eds, 2007). In our model the effect of inequality on output is transmitted through the aggregate demand. For example, higher wage inequality per se increases the overall amount of resources allocated to consumers, as well as more complex organisation of firm labour (low v). But the relation between inequality and output when we do not control for structural parameters (a result biased by omitted variables) is negative. This correlation simply says that in stagnating economies, due to a lack of demand, there is no investment in capital goods and all profits are redistributed to managers: low growth induces inequality. Although institutions have no role in our model, this result is in line with the literature that shows the relevance of elite behaviour in countries long run development (Acemoglu *et al.*, 2005). While an economy growing rapidly also generates some inequality, due to the hierarchical structure of firms, this same non-linear increase in vacancies that accompanies an increase in demand generates even more growth and demand. This is amplified by the size of wages, *i.e.* consumable income. However, as noted elsewhere (Ciarli *et al.*, 2010), it is the overall demand that counts in our model, and not how it is redistributed, given that we assume that all classes have the same propensity to consume their income. Here we simply observe that growing firms generate more demand; and different wage classes consume a different set of goods. However, given the minor role played by demand side factors in our model, we can make no judgement on what would be the best way to redistribute salaries. What seem to be essential from analysing the model dynamics is the selection of firms by consumers, either because they differentiate on price (low income) or on quality (top incomes). And because of the pyramidal structure of the firm, the price selection by the large population of shop-floor workers seems essential in facilitating the transition from low to exponential growth. Finally, large profit shares in the form of large mark-up have the opposite effect on real demand, through an increase in prices, with a negative impact on output growth. We plan to focus more on the relation between the different aspects of inequality and output growth in future research, exploiting the panel structure of the simulation data.

Among the implication for future research with this and similar models, two limitations in the analysis of this paper suggest two future steps. First, here we do not consider the deeper determinants of the initial differences in the factors that determine structural change, which are mainly related to institutional aspects—in a broad sense—and to the intricate relation between knowledge, technology, and institutions. In future work we want to build on the sizeable literature that uses the case of the industrial revolution(s) to investigate the relations between knowledge, technology, and institutions to shed more light on the origins of the structural difference that we model here. Second, having studied here which are the most relevant aspects in determining the output response with a simple DOE, the next step is to focus on those aspects and analyse whether their effect on output is linear and monotonous, as assumed in the DOE. Previous work suggests that it is not (Ciarli and Lorentz, 2011; Ciarli *et al.*, 2012). Of particular importance is the relation between the various components of inequality, labour organisation and output growth. Moreover, we plan to study how different aspects of structural change modify their effect through time, exploiting the panel structure of the data. Finally, one can exploit these same results to simplify the model reducing the number of parameters and behavioural details. With reference to the model presented here, we plan to simplify purchase behaviour to a completely adaptive demand behaviour. More in general, it seems like a good strategy to start with highly complex agent based models, thoroughly analyse the most relevant aspects with an appropriate DOE, and learn from the analysis how to simplify the model, stylising the aspects that have little influence on the outputs.

References

Acemoglu D., S. Johnson, and J. A. Robinson, 2005. "Institutions as a fundamental cause of long–run growth." In Philippe Aghion and Steven N. Durlauf, editors, *Handbook of Economic Growth*, volume 1A of *Economic handbooks*, chapter 6, pages 385–472. Elsevier.

Aghion P. and P. Howitt, 1998. *Endogenous Growth Theory*, Vol. 1 of MIT Press Books, Cambridge, MA: The MIT Press, December.

Aghion P., E. Caroli, and C. García-Peñalosa, 1999. "Inequality and economic growth: The perspective of the new growth theories." *Journal of Economic Literature*, 37(4):1615–1660, December.

Amendola M., and J.-L. Gaffard, 1998. *Out of Equilibrium*. Oxford: Oxford Clarendon Press.

Andrews M. J., S. Bradley, D. N. Stott, and R. Upward, 2008. "Successful employer search? an empirical analysis of vacancy duration using micro data." *Economica*, 75(299):455–480.

Aoki M., and H. Yoshikawa, 2002. "Demand saturation–creation and economic growth." *Journal of Economic Behavior & Organization*, 48(2): 127–154.

Atkinson A. B., 2007. "The distribution of earnings in oecd countries." *International Labour Review*, 146(1-2):41–60.

Bassin W. M., M. T. Marsh, and S. Walitzer, 2003. "A macroeconomic analysis of inventory/sales ratios." *Journal of Business & Economics Research*, 1(10):37–45.

Baumol W. J., 2010. “The two-sided cost disease and its frightening consequences.” In F. Gallouj and F. Djellal, editors, *The handbook on Innovation and Services*. Cheltenham: Edward Elgar.

Berg M., 2002. “From imitation to invention: creating commodities in eighteenth-century Britain.” *Economic History Review*, 55(1):1–30.

Blanchard O. J., 1983. “The production and inventory behavior of the american automobile industry.” *Journal of Political Economy*, 91(3): pp. 365–400.

Blanchflower D. G. and A. J. Oswald, 2006. *The wage curve: An entry written for the new palgrave, 2nd edition*. *IZA Discussion Paper* 2138, Bonn, May.

Blinder A. S., 1982. “Inventories and sticky prices: More on the microfoundations of macroeconomics.” *American Economic Review*, 72(3): 334–348.

Blinder A. S., 1991. “Why are prices sticky? preliminary results from an interview study.” *American Economic Review*, 81(2):89–96.

Boeri, T., 2012. “Setting the minimum wage.” *Labour Economics* 19(3): 28–290.

Bottazzi, G. and M. Grazzi, 2010. "Wage-size relation and the structure of work—force composition in Italian manufacturing firms," *Cambridge Journal of Economics*, 34 (4):649–669.

Buchanan M., 2009. “Economics: Meltdown modelling. could agent-based computer models prevent another financial crisis?.” *Nature*, 460: 680–682.

Celsi R. L., and J. C. Olson, 1988. “The role of involvement in attention and comprehension processes.” *Journal of Consumer Research*, 15(2):210–224.

Chiarella C., P. Flaschel, G. Groh and W. Semmler, 2000. *Disequilibrium, growth, and labor market dynamics: macro perspectives*. Springer, Heidelberg.

Ciarli T., A. Lorentz, M. Savona, and M. Valente, 2010. “The effect of consumption and production structure on growth and distribution. A micro to macro model.” *Metroeconomica*, 61(1):180–218.

Ciarli T., A. Lorentz, M. Savona, and M. Valente, 2012. “The role of technology, organisation, and demand in growth and income distribution.” *LEM Working Papers* 2012/06, Laboratory of Economics and Management, San’Anna School of Advanced Studies, Pisa.

Ciarli T. and A. Lorentz., 2010. “Product variety and changes in consumption patterns: The effects of structural change on growth.” *Working paper mimeo*, Max Planck Institute of Economics.

Ciarli T., and A. Lorentz, 2011. “Product variety and economic growth. trade off between supply and demand dynamics.” *Working paper mimeo*, Max Planck Institute of Economics.

Ciarli T., 2012. “Structural interactions and long run growth: An application of experimental design to agent based models.” *Papers on Economics and Evolution 1206*, Max Planck Institute of Economics, Evolutionary Economics Group, Jena.

Clark C., 1940. *The Conditions of Economic Progress*. MacMillan, London.

Coad A. and R. Rao, 2010. “Firm growth and r&d expenditure.” *Economics of Innovation and New Technology*, 19(2):127–145.

Coelli T., E. Grifell-Tatjé, and S. Perelman, 2002. “Capacity utilisation and profitability: A decomposition of short-run profit efficiency.” *International Journal of Production Economics*, 79(3):261–278.

Colander D., P. Howitt, A. Kirman, A. Leijonhufvud, and P. Mehrling, 2008. “Beyond DSGE models: toward an empirically based macroeconomics.” *American Economic Review*, 98(2):236–240.

Cooper, Russel W and John C. Haltiwanger, "On the Nature of Capital Adjustment Costs," *Review of Economic Studies*, 2006, 73(3):611–633.

Criscuolo C., 2000. "Employer Size - Wage Effect: A Critical Review and an Econometric Analysis." SSRN eLibrary.

Dasgupta S., and A. Singh, 2005. “Will services be the new engine of indian economic growth?” *Development and Change*, 36(6):1035–1057.

Davis S. J., R. J. Faberman, and J. C. Haltiwanger, 2010. “The establishment-level behavior of vacancies and hiring.” *Working Paper* 16265, National Bureau of Economic Research.

Dawid H. and W. Semmler, 2010. “Introduction to the special issue: Computational perspectives in economics and finance: Methods, dynamic analysis and policy modeling.” *Journal of Economic Dynamics and Control*, 34(9):1529–1530.

Delli Gatti D., E. Gaffeo, and M. Gallegati, 2010. “Complex agent-based macroeconomics: a manifesto for a new paradigm.” *Journal of Economic Interaction and Coordination*, 5:111–135.

Desmet K., and S. L. Parente, 2009. “The evolution of markets and the revolution of industry: A quantitative model of england’s development, 1300-2000.” *Working Papers* 2009-06, Instituto Madrileño de Estudios Avanzados (IMDEA) Ciencias Sociales.

DeVaro J., 2005. “Employer recruitment strategies and the labor market outcomes of new hires.” *Economic Inquiry*, 43(2):263–282.

De Loecker J., and F. Warzynski, 2009. “Markups and firm-level export status.” *NBER Working Papers* 15198, National Bureau of Economic Research, Inc.

Dietrich A., and J. J. Krüger, 2010. “Numerical explorations of the ngai-pissarides model of growth and structural change.” *Working Paper mimeo*, Darmstadt University of Technology.

Doms M. and T. Dunne, 1998. "Capital Adjustment Patterns in Manufacturing Plants," *Review of Economic Dynamics*, 1, 409–429.

Dosi G., G. Fagiolo, and A. Roventini, 2010. “Schumpeter meeting keynes: A policy–friendly model of endogenous growth and business cycles.” *Journal of Economic Dynamics and Control*, 34:1748–1767.

Dosi G., Marengo L., and C. Pasquali, 2006. “How much should society fuel the greed of innovators?” : On the relations between appropriability, opportunities and rates of innovation. *Research Policy*, 35(8):1110 – 1121.

Eicher T. S. and S. J. Turnovsky, editors, 2007. *Inequality and Growth. Theory and Policy Implications*. CESifo Seminar Series. MIT Press.

Fabiani S., M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Mathä, R. Sabbatini, H. Stahl, and A. Stokman, 2006. “What firms’ surveys tell us about price-setting behaviour in the euro area.” *International Journal of Central Banking*, 2:3–47.

Fisher A. G. B., 1939. “Production, primary, secondary and tertiary.” *Economic Record*, 15(1):24–38.

Fraumeni B., 1997. “The measurement of depreciation in the u.s. national income and product accounts.” Survey of current business, Bureau of Economic Analysis, July.

Galor O., 2010. “The 2008 lawrence r. klein lecture comparative economic development: Insights from unified growth theory.” *International Economic Review*, 51(1):1–44.

Gervais M., and P. Klein, 2010. “Measuring consumption smoothing in cex data.” *Journal of Monetary Economics*, 57(8):988–999.

Gigerenzer G., 1997. "Bounded Rationality: Models of Fast and Frugal Inference." *Swiss Journal of Economics and Statistics*, 133 (2).

Gigerenzer G., and R. Selten (editors), 2001. *Bounded rationality: The adaptive toolbox*, Cambridge, MA: MIT Press.

Hall S., M. Walsh, and A. Yates, 1997. “How do uk companies set prices?” *Working Paper* 67, Bank of England.

Idson T. L. and W. Y. Oi, 1999. "Workers Are More Productive in Large Firms." *American Economic Review*, 89 (2, Papers and Proceedings), 104–108.

Oliveira M. J., S. Scarpetta, and D. Pilat, 1996. “Mark-up ratios in manufacturing industries: Estimates for 14 oecd countries.” *OECD Economics Department Working Papers 162*, OECD Publishing.

Jung P., and M. Kuhn, 2011. “The era of the u.s.-europe labor market divide: what can we learn?” *MPRA Paper 32322*, University Library of Munich, Germany, July.

Kaldor N., 1957. “A model of economic growth.” *Economic Journal*, December:65–94.

King R. G. and R. Levine, 1994. "Capital fundamentalism, economic development, and economic growth." Policy Research Working Paper Series 1285, *The World Bank*, April.

Kleijnen J. P. C., S. M. Sanchez, T. W. Lucas, and T. M. Cioppa, 2005. "State-of-the-art review: A user's guide to the brave new world of designing simulation experiments." *INFORMS Journal on Computing*, 17(3):263–289, Summer.

Kleijnen J. P. C., 2005. "Simulation experiments in practice: statistical design and regression analysis." *Journal of Simulation*, 2(1):19–27.

Krueger D. and F. Perri, 2005. "Understanding consumption smoothing: Evidence from the u.s. consumer expenditure data." *Journal of the European Economic Association*, 3(2/3):340–349.

Kuznets S., 1973. "Modern economic growth: Findings and reflections." *The American Economic Review*, 63(3):247–258.

Langbraaten N., E. W. Nordbo and F. Wulfsberg, 2008. "Price-setting behaviour of norwegian firms - results of a survey." *Norges Bank Economic Bulletin*, 79(2):13–34.

Law A. M., 2004. "Statistical analysis of simulation output data: The practical state of the art." In *Winter Simulation Conference*, pages 67–72.

LeBaron B. and L. Tesfatsion, 2008. "Modeling macroeconomies as open-ended dynamic systems of interacting agents." *American Economic Review*, 98(2):246–50, May.

Leijonhufvud A. 2006. "Agent-based macro." Volume 2 of *Handbook of Computational Economics*, pages 1625–1637. Elsevier.

Llerena P. and A. Lorentz, 2004. "Cumulative causation and evolutionary micro-founded technical change: On the determinants of growth rates differences." *Revue Economique*, 55(6):1191–1214.

Lydall H. F., 1959. "The distribution of employment incomes." *Econometrica*, 27(1):110–115.

Maddison A., 2001. *The World Economy: A Millennial Perspective*. OECD, Paris.

Marchetti D. 2002. "Markups and the business cycle: Evidence from italian manufacturing branches." *Open Economies Review*, 13(1):87–103.

Matsuyama K., 2008. "Structural change." In Steven N. Durlauf and Lawrence E. Blume, editors, *The New Palgrave Dictionary of Economics*. Palgrave Macmillan, Basingstoke.

McCloskey D. N., 2009. "Science, bourgeois dignity, and the industrial revolution." *MPRA Paper* 22308.

Mokyr J., 2002. *The gifts of Athena: historical origins of the knowledge economy*. Princeton: Princeton University Press.

Moneta A., D. Entner, P. Hoyer, and A. Coad, 2010. "Causal inference by independent component analysis with applications to micro- and macroeconomic data." *Jena Economic Research Papers* 2010-031, Jena: Friedrich Schiller University Jena and Max Planck Institute of Economics.

Montgomery D. C. 2001. *Design and Analysis of Experiments*. John Wiley & Sons, New York, 5 edition.

Nadiri M. I., and I. R. Prucha, 1996. "Estimation of the depreciation rate of physical and r&d capital in the u.s. total manufacturing sector." *Economic Inquiry*, 34(1):43–56.

Nelson R. R., and Winter S.G., 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press.

Nickell S., L. Nunziata, 2002. Wolfgang Ochel, and Glenda Quintini. "The beveridge curve, unemployment and wages in the oecd from the 1960s to the 1990s." *Working paper, CEPR*, LSE, London.

Nijkamp P., and J. Poot. 2005. "The last word on the wage curve?" *Journal of Economic Surveys*, 19(3):421–450.

Office for National Statistics, 2006. *Expenditure and food survey, 2005-2006*. Technical report, Office for National Statistics, Department for Environment, Food and Rural Affairs.

Rostow W.W., 1978. *The World Economy: History and Prospect*. Austin: University of Texas Press.

Ryoo J., and S. Rosen, 1992. "The market for engineers." Working Paper 83, University of Chicago - George G. Stigler Center for Study of Economy and State, Chicago.

Saviotti P. P., and J.-L. Gaffard, 2008. "Preface for the special issue of jee on 'innovation, structural change and economic development'." *Journal of Evolutionary Economics*, 18(2):115–117.

Silverberg G. and B. Verspagen, 2005. "Evolutionary Theorizing on Economic Growth." in Kurt Dopfer, ed., The Evolutionary Foundations of Economics, Cambridge: *Cambridge University Press*, chapter 16, pp. 506–539.

Simon H. A., 1957. "The compensation of executives." *Sociometry*, 20(1):32–35, March.

Teo E., S. M. Thangavelu, and E. Quah, 2004. "Singapore's beveridge curve: A comparative study of the unemployment and vacancy relationshio for selected east asian countries." *Economic Survey of Singapore Second Quarter*, Ministry of Manpower.

U.S. Census Bureau, 2008. *Manufacturing and trade inventories and sales*. http://www.census.gov/mtis/www/mtis.html

Wall H. J. and G. Zoega, 2002. "The british beveridge curve: A tale of ten regions." *Oxford Bulletin of Economics and Statistics*, 64(3):257–276.

Zeithaml V. A., 1988. "Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence." *The Journal of Marketing*, 52(3):2–22.

Appendix

A. Experimental design

The choice of limit values

The "Low" and "High" values of the factors analysed in this paper were chosen following three criteria: (i) the minimum or maximum possible values (when a limit exists, such as in shares): in this case we selected values that approximate the extremes. (ii) We made reference to the empirical evidence, and chose values below the minimum observed for "Low", and values above the maximum observed for High. Or (iii) when the factors have been tested in previous work (for a large number of values) we used the values below/beyond which the factor did not show to have a significant effect. The mix of these criteria makes the choice of the "Low" and "High" values in this paper quite robust and meaningful—we do not see many other options for this type of work.

Table 7. Choice of limit values of factors

The "Low" and "High" values of the factors are chosen with respect to available empirical evidence and/or within theoretically possible and meaningful values.

Factor	Low	High
ι	Approximates the minimum	No significant effect found above 0.2 when we analyse a large number of values between 0 and 10 (Ciarli and Lorentz, 2011)
ν	Minimum postulated in Lydall (1959); Simon (1957)	Far above the max postulated in Lydall (1959); Simon(1957) and very marginal effect observed beyond 15 (Ciarli *et al.* 2012)
$\boldsymbol{b}$	The minimum	Above the maximum postulated by Lydall (1959); Simon (1957)
σ^a	Approximates the minimum	Maximum valued used in Ciarli *et al.*(2012): an extremely large number for the time frame postulated here
η	Close to the minimum (to allow for some changes in consumption patterns)	Well beyond the theoretical maximum
$\rho = \rho^k$	Approximates the minimum share	Approximates the maximum share
ϑ	Approximates the minimum value	No significant effect found above 2 when we analyse a large number of values between 0 and 10 (Ciarli and Lorentz, 2011)
ξ	Close to the minimum	Close to max probability (also for small level of R&D investment)
ς	Approximates the minimum share	Approximates the maximum share
$\mu = \mu^K$	Approximates the minimum	Well beyond observed evidence (De Loecker and Warzinski, 2009; Joaquim Oliveira *et al.*, 1996; Marchetti, 2002)

B. Tables

Table 8. Vector autoregression analysis of the main macro variables

Results from LADestimates of 10 periods growth rates, and bootstrapped standard errors (400). (1) ΔY: output growth; (2) ΔA: Aggregate productivity (3) ΔP: price; (4) ΔIHI: inverse Herfindhal Index (5) ΔAT: Atkinson inequality index. The numbers in parenthesis indicate the number of lags

Variables	(1) ΔY	(2) ΔA	(3) ΔP	(4) ΔIHI	(5) ΔAT
ΔY (1)	**0.986*** **	**-0.021** **	**0.014**	**0.189**	**0.247*** **
	(0.037)	(0.010)	(0.009)	(0.145)	(0.076)
ΔA (1)	**0.016**	**0.735*** **	**-0.045*** **	**-0.284**	**-0.273* **
	(0.026)	(0.029)	(0.013)	(0.237)	(0.161)
ΔP (1)	**-0.697*** **	**0.049**	**1.058*** **	**0.718**	**0.165**
	(0.063)	(0.058)	(0.062)	(0.746)	(0.405)
ΔIHI (1)	**0.010*** **	**0.019*** **	**-0.011*** **	**0.624*** **	**-0.249*** **
	(0.003)	(0.004)	(0.003)	(0.043)	(0.026)
ΔAT (1)	**0.020*** **	**0.001**	**-0.001**	**0.065**	**0.941*** **
	(0.005)	(0.006)	(0.003)	(0.049)	(0.035)
ΔY (2)	**0.081**	**0.024**	**-0.012**	**-0.384** **	**-0.165* **
	(0.051)	(0.016)	(0.011)	(0.192)	(0.088)
ΔA (2)	**0.049* **	**0.044**	**0.025**	**-0.085**	**0.228**
	(0.026)	(0.039)	(0.017)	(0.272)	(0.228)
ΔP (2)	**0.771*** **	**-0.001**	**0.054**	**-1.045**	**-2.385*** **
	(0.092)	(0.082)	(0.072)	(0.981)	(0.690)
ΔIHI (2)	**-0.014*** **	**-0.026*** **	**0.005* **	**0.237*** **	**0.247*** **
	(0.004)	(0.004)	(0.003)	(0.053)	(0.032)
ΔAT (2)	**-0.001**	**0.006**	**-0.000**	**-0.032**	**0.013**
	(0.007)	(0.008)	(0.003)	(0.057)	(0.045)
ΔY (3)	**-0.069** **	**-0.006**	**0.000**	**0.213* **	**-0.080**
	(0.035)	(0.011)	(0.006)	(0.127)	(0.052)
ΔA (3)	**0.000**	**0.085*** **	**0.005**	**0.138**	**-0.015**
	(0.023)	(0.029)	(0.014)	(0.271)	(0.177)
ΔP (3)	**-0.052**	**-0.083**	**-0.131*** **	**0.247**	**2.104*** **
	(0.056)	(0.066)	(0.051)	(0.775)	(0.528)
ΔIHI (3)	**0.006* **	**0.005**	**0.005**	**0.098* **	**-0.010**
	(0.003)	(0.004)	(0.003)	(0.054)	(0.029)
ΔAT (3)	**-0.014*** **	**-0.008**	**0.002**	**0.018**	**-0.019**
	(0.005)	(0.005)	(0.002)	(0.042)	(0.029)
Constant	**-0.000**	**0.000*** **	**-0.000*** **	**-0.001**	**-0.000**
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
Observations	1,950	1,950	1,950	1,950	1,950
Pseudo R^2	0.98	0.61	0.84	0.71	0.67

Standard errors in parentheses

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 9. Kruskall-Wallis rank test. Main parameter effects

Parameter	Obs	Rank Sum	chi-squared	df	Prob
ι					
0.001	10240	1.010e+08	76	1	0.00
0.3	10240	1.090e+08			
ν					
3	10240	1.020e+08	34.94	1	0.00
50	10240	1.070e+08			
b					
1	10240	1.030e+08	17.54	1	0.00
3	10240	1.070e+08			
σ^a					
0.01	10240	9.120e+07	1048	1	0.00
0.2	10240	1.190e+08			
η					
0.1	10240	1.060e+08	8.621	1	0.00
3	10240	1.040e+08			
ρ					
0.05	10240	1.310e+08	3744	1	0.00
0.95	10240	7.900e+07			
ϑ					
0.01	10240	1.030e+08	15.88	1	0.00
10	10240	1.070e+08			
ξ					
0.1	10240	9.220e+07	902.3	1	0.00
1000	10240	1.180e+08			
ς					
0.05	10240	1.050e+08	0.334	1	0.56
0.9	10240	1.050e+08			
μ					
1.01	10240	1.480e+08	10452	1	0.00
2	10240	6.160e+07			

ON THE CO-EVOLUTION OF INNOVATION AND DEMAND: SOME POLICY IMPLICATIONS[1]

Pier Paolo Saviotti

INRA-GAEL, Université Pierre Mendès-France - GREDEG CNRS, Sophia Antipolis, Valbonne

Andreas Pyka

University of Hohenheim, Economics Institute, Germany

Long term economic development is characterized not only by increasing efficiency of economic activities but also by qualitative change within industries and increasing variety concerning the existence of different industries. Traditional economic growth models do not cope with the complex amalgam of these three trajectories of economic development nor could comprise the interactions among them. Furthermore, economic development is not a process which is spurred by supply-side effects but driven by the co-evolutionary interplay of supply and demand side forces. With our TEVECON model we analyze economic development driven by efficiency and quality improvements together with structural change and the co-evolution between innovation and demand. The first part of the paper introduces to the basic model and some general results. The second part of the paper deals with policy experiments which are undertaken by comparing different numerically analyzed scenarios.

Keywords: Economic Development, Qualitative Change, Co-Evolution, Simulation.

The main objective of this paper is to establish that innovation could not have contributed to economic development unless a demand for the goods and services created by innovation existed. We explore the conditions required for such a demand to exist and argue that the process which gave rise to the observed path of economic development was the co-evolution of demand and innovation. Furthermore, we explore how the co-evolution of demand

---

1. The research leading to these results has received funding from the European Union Seventh Framework ProgrammeFP7/2007-2013 under grant agreement n° SSH-CT-2010-266959.

and innovation changed the capitalist economic system from one in which most people could afford only bare necessities to one in which most people have a highly and increasingly varied pattern of consumption, including a growing proportion of items which cannot be judged necessities, and which are of higher quality than in the past. Finally, we study the possible impact of economic policies on the above co-evolutionary process. We carry out these explorations by means of an extension of our TEVECON model of economic development, which is described in the following part of the paper.

1. 2. Conceptual background

1.1. Co-evolution and economic development

The concept of co-evolution has recently been used in the innovation literature to analyze the co-evolution of technologies and institutions. In this section we make a brief reference to this literature and propose a more general concept of co-evolution. Technologies cannot develop in an institutional vacuum but need appropriate institutions (Nelson, 1994). Such institutions are required to support the collective interests of a new technology and of the corresponding industry, to lobby the industry, to regulate it, to establish intellectual property rights, to create the required infrastructures etc. Examples of such co-evolution are mass production in the United States car industry, the emergence of synthetic dyes in Germany (Murmann, 2003), biotechnology in the USA (Nelson, 2008).

The need for new institutions becomes evident when new technologies emerge. There are even institutions which might be appropriate at a level of aggregation higher than that of an individual industry. For example, a set of interconnected technologies sharing common resources could require a set of institutions appropriate to the whole set. Perez (1983) used the related concept of techno-economic paradigm (1983) to encompass a technological paradigm (Dosi, 1982) and the institutions appropriate to it. She maintained that the creation of the appropriate institutions was likely to be a longer and more complex process than the initial creation of a given technology, or technologies, corresponding to a given technological paradigm. Other scholars stressed that a country which had been successful in creating a given set of technologies and the appropriate institutions might be unable to do the same for a subsequent set of technologies. Veblen (1915) had already remarked how British industry, which was very successful in the early part of the industrial revolution, could not adopt the institutions appropriate to the new technologies that Germany developed much more successfully. Lazonick (1990) maintained that the organization of work and the institutions for training labor which had underpinned successful development of Britain in the late 18th and early 19th century became a handicap in the 20th century. At a higher level of generality Polanyi (1944) maintained that capitalism would require the creation of institutions which were capable to compensate the harsh if efficient nature of capitalist societies.

A more general interpretation of the concept of co-evolution can be proposed at a system level. A system is constituted by different and interacting components. Co-evolution exists when two different components (C_1 and C_2) interact in such a way that changes in one of them, say C_1, affect C_2 and that changes in C_2 affect C_1. Typically, for co-evolution to exist this relationship of mutual interaction must last for several periods, giving rise to a sustained feedback loop.

The dynamics we can imagine for an economic system consists of the early emergence of an innovation in a pre-institutional form, that is, without institutions specific to the new technology. This would be followed by the creation of institutions which, for example, would provide the rules for the new technology to be used for the advantage of society at large avoiding as much as possible negative side effects, and of infrastructures which would allow the market for the new technology to grow. A clear example is given by cars and roads: the scope for cars has been considerably enhanced by the construction of roads. Thus, the more the new technology develops, the more the appropriate institutions need to grow giving rise to a feedback loop which would slow down only when the market(s) for the new technology were completely saturated.

The type of co-evolution we are going to be concerned with in this paper is between innovation and demand. Thus, it would be the co-evolution of two different economic variables, mutually influencing each other during the course of economic development. In TEVECON this co-evolutionary process occurs because sectoral search activities, which increase with sectoral demand, affect output price, quality and differentiation, which in turn affect demand. A positive feedback loop can be established which can give rise to a faster growth of both demand and innovation than it would have been possible if the two variables had not been influencing each other. In this sense co-evolution works as autocatalysis (Nicolis and Prigogine, 1989) by using the output of one stage of the process as the input of the following stage.

1.2. Trajectories and patterns of economic development

Models of economic development need to be able to explain patterns of long run development and growth. Here long run is intended to indicate a period such as the one from the industrial revolution to the present. Focusing on such a period requires understanding the broad features which occurred in it. First, we had the emergence of manufacturing industry. Second, within manufacturing there was a progressive differentiation, beginning with sectors such as textiles, energy (steam engine), railways, steel and following with chemicals, electricity, cars, planes etc. During this process manufacturing industry became increasingly differentiated, with newer sectors coexisting with older ones. Third, the employment share of services overtook that of manufacturing.

Any model of economic development which is in principle capable of interpreting events which occurred since the industrial revolution needs to explain why such structural change occurred. The fundamental ingredient which gives rise to growth and development in our work is innovation. The emergence of innovations is due to search activities, which provide the knowledge required to create and modify innovations. Innovations affect economic development because entrepreneurs fund new firms to exploit the outcomes of search activities and because consumers and users purchase the products and services embodying such innovations. In this process the economic system becomes increasingly differentiated. The addition of new sectors to the economic system not only contribute to structural change but to a structural change occurring in a particular direction, that of increasing differentiation.

The above processes can be described in terms of three trajectories and of two periods.

Trajectory 1: The efficiency of productive processes increases during the course of economic development. Here efficiency must be understood as the ratio of the inputs used to the output produced, when the type of output remains *constant*.

Trajectory 2: The output variety of the economic system increases in the course of time. Here such variety is measured by the number of distinguishable sectors, where a sector is defined as the set of firms producing a common although highly differentiated output.

Trajectory 3: The output quality and internal differentiation of existing sectors increases in the course of time after their creation. This means that if during the period of observation the type of output changes what we will observe is a combination of growing productive efficiency and of quality change.

From now on we will use the term variety as a synonym of diversity, although the two are in principle not identical. Such variety can exist at the inter-sectoral as well as at the intra-sectoral level. Thus, two sectors will produce completely different types of output while one sector will produce a diversified output. In the literature these two types are often described as vertical and horizontal differentiation respectively. Such long run trajectories do not emerge separately but exist due to a complex pattern of interactions within the economic system.

These trajectories occur at a level of aggregation higher than that of an individual industrial sector or a technology. The ratio can be calculated in value or in volume terms. Growing productive efficiency is the oldest and, until the industrial revolution, the most developed of the three trajectories. For example, the efficiency of food production increased with the transition from hunters and gatherers to settled agriculture (Diamond, 1997). However, any such increases in productive efficiency were very slow and not necessarily cumulative. Productive efficiency started growing in a cumulative fashion only after the beginning of the industrial revolution (Maddison, 2007). Simple recent examples of this trajectory can be found in the falling number of workers required to produce a unit of output in the steel, chemicals or car industries. Of course, these are just examples and the phenomenon is far more general. Growing productive efficiency is certainly one of the factors which contributed to economic growth since the industrial revolution. However, the observed patterns of economic development could not have been produced by growing productive efficiency alone. In this case we would produce today Ford Model T like cars with much smaller quantities of all the inputs required. As even the most casual observer would have noticed, today's cars are not only produced much more efficiently than those of the early 20th century but they are also of a much higher quality. Hence, growing productive efficiency and growing output quality were combined in the patterns of economic development which we can observe today.

During the industrial revolution output differentiation (trajectory 2) was very limited. At the beginning it occurred mostly at the level of capital goods (new textile and engineering equipment, railways equipment etc.) and only considerably later at the level of consumer goods. The increasing internal differentiation and output quality of consumer goods and durables started increasing during the 19th century and in particular after the beginning of the 20th century. Growing output variety can be observed at the inter-sector level. A clear example of this is the large number of completely new sectors which emerged during the 20th century, such as cars, aircraft, television, computers, telecommunications etc. All of these not only constituted completely new sectors but underwent a very high degree of internal differentiation.

These three trajectories are not independent. None of them could have occurred taken place alone without the other two. Thus, a continuous increase in productive efficiency, if not accompanied by the emergence of new sectors and by their internal differentiation and rising quality could have led the economic system to a bottleneck in which all demanded output could have been produced by a declining proportion of the labor force (Pasinetti, 1981). Such a bottleneck, determined by the imbalance between continuously increasing productive efficiency and saturating demand, could have been overcome by the emergence of new sectors (Pasinetti, 1981). While the assumption of demand saturation and the neglect of the internal differentiation of sectors limited the possible generalization of Pasinetti's approach, we have shown (Pyka and Saviotti, 2012) that both the emergence of new sectors and their increasing quality and internal differentiation provided additional scope for further growth and allowed its continuation in the long run. In this context, full demand saturation is unlikely to occur within any sector as long as new sectors keep being created (Saviotti and Pyka, 2010). Furthermore, both the emergence of new sectors and their growing quality and internal differentiation can compensate the diminishing capability to create employment of incumbent and maturing sectors.

In the previous sections we described the period from the industrial revolution to the present as the transition from necessities to imaginary worlds. This description emphasizes that until the end of the 19th century most people, even in countries which were for the standards of the time relatively rich, could not purchase anything but bare necessities. All throughout the 19th century British working class households spent about ninety percent of their income on food, clothing and housing. Only during the 20th century, and in particular after the 1930s, the share of income spent on the above three categories started falling (Hobsbawm, 1968, diagrams 45 and 46). By the 1950s the share of necessities fell to about 60 percent, leaving about forty percent to be spent on other, presumably higher, goods and services. The compression of the combined expenditure on necessities (trajectory 1) created the disposable income required to buy the new goods and services which were gradually being created. Starting from the beginning of the 20th century new goods and services emerged (trajectory 2) and their quality and differentiation increased constantly (trajectory 3). This combination of trajectories contributed to a mechanism which allowed the capitalist economic system to create growing wealth for most of the population of industrialized countries.

1.3. TEVECON

1.3.1. Modeling philosophy

Our model, which we call TEVECON, can be considered an Agent Based Model (ABM) for a number of reasons. First, it is not an analytical model in the same sense as the more orthodox models, because it lacks *closure conditions*. The most important of such conditions is the presence of general equilibrium. Our model has an endogenously varying number of sectors, and thus an endogenously variable composition. In these circumstances, as Kaldor (1957) had already well understood, there can be no general equilibrium. However, we do have sectoral equilibrium in the form of a feedback mechanism ensuring that demand does not deviate too much from supply. Also, TEVECON agents are not optimizers but only improvers possessing bounded rationality (Pyka and Fagiolo 2007), since learning mechanisms (mainly learning by searching) play a central role in TEVECON.

TEVECON has a number of agents, but sometimes they are implicitly or lightly represented only. The central agents of TEVECON are sectors, defined as the collection of firms producing a unique though highly differentiated type of output. Firms are present and one of the most important modeling outcomes of TEVECON is the evolution of the number of firms in time. Although reduced, such a presentation of firms gives rise to the very interesting prediction of the existence of an Industry Life Cycle (ILC) under a very wide range of conditions. However, the representation of firms can be considerably expanded by including firm characteristics, internal structure and distributional properties. An agent which is present only implicitly is the Schumpeterian entrepreneur, who is creating new firms by exploiting important innovations induced by the expectation of a temporary monopoly. The role of the entrepreneur is extremely important in TEVECON but its representation at the moment is reduced to the action of open up new sectors. Thus, the central agents of TEVECON are industrial sectors as previously defined.

Another important feature of Agent Based Models (ABM) is the reconstruction of the macro-economic states of the system from its micro-economic ones (Pyka, and Fagiolo, 2007). In this sense TEVECON is best defined as providing aggregation from micro to meso and from meso to macro. Firms (micro) are aggregated to sectors (meso) and sectors are aggregated to the macroeconomic state of the system. In the present version of the model the meso to macro aggregation is better specified than the micro to meso one.

Sectors are very considerably heterogeneous in TEVECON. They can differ on a very large number of dimensions, such as expected market size, technological opportunity, investment patterns, wage rates etc. Furthermore, TEVECON satisfies most of the conditions required to be considered an evolving complex system (ECS) (Pyka and Fagiolo, 2007, p. 474) since it is a highly interactive model in which new interactions are continuously being introduced between existing variables. One such interaction that was present from the very early versions of the model is that between search activities and demand, where there is a feedback mechanism from rising demand to rising search activities to further rising demand in following periods. More such interactions are continuously being introduced. Again, these interactions contribute to the emergence of complex properties out of repeated interactions among simple entities (Kirman, 1998).

TEVECON shows Endogenous and Persistent Novelty (Pyka and Fagiolo, 2007, p. 475). It is non stationary in the sense that its composition is continuously changing. New sectors produce outputs that are qualitatively different from the pre-existing ones. This means that in principle the outputs of different sectors should not be substitutable. In reality our model includes two types of competition, intra- and inter-sector. The latter exists if different sectors produce comparable services out of non comparable internal structures (Saviotti and Metcalfe, 1984; Saviotti, 1996). Thus, the qualitative difference lies mostly in the internal structure of sectoral outputs and in the sector's knowledge base.

As a consequence of the above, TEVECON shows 'true dynamics' (Pyka and Fagiolo, 2007, p. 475). Some form of dynamics is present in orthodox models simply because they include equations which show the time paths of the system. This form of dynamics does not take into account qualitative change and is not affected by the emergence of new entities. One of the most important differences between evolutionary and ABM models on the one hand, and orthodox models on the other hand, is the emergence of new entities, qualitatively different from pre-existing ones. The true dynamics which is more difficult to represent and yet vital to understand the long run evolution of the economic system is the one including qualitative change.

If the above considerations allow us to consider TEVECON an ABM model, we can still situate it within the wide range of modeling techniques which are in principle compatible with the ABM definition. TEVECON bears a close similarity to dynamical systems since its basic framework is constituted by a set of simultaneous difference equations. Although complete closure conditions such as general equilibrium are absent, the equations used are in most cases similar or identical to those which are used in orthodox analytical models. Given the absence of closure and the nature of the equations involved, TEVECON cannot be analytically solved but needs to be simulated. Thus, amongst all ABM modeling techniques TEVECON could be described as having a partly analytical, not entirely computational, structure but needing simulation to find solutions. This gives TEVECON both advantages and disadvantages. With respect to orthodox analytical models it has the advantage of allowing us to include a greater number of variables and interactions while having a greater similarity to orthodox analytical models than purely computational ABM models. TEVECON's disadvantage with respect to purely computational ABM models is its lower adaptability to model institutions and policies.

1.3.2. The model

In TEVECON the economic system is composed of an endogeneous variable number of sectors. The emergence of new sectors is due to the dynamics of the incumbent ones and the main source of economic growth consists in the emergence of new sectors. Each sector is created on the basis of an important, pervasive, innovation taken up by entrepreneurs who start new companies and thereby provide the basis for a new industry. The innovation creating the sector gives rise to an adjustment gap AG_i, a variable intended to capture the size of the potential market established by the innovation. However, this market is initially empty because neither the production capacity nor a structured demand for the new products exists. Both the production capacity and the evolution of the demand will take place during a (possibly long) period of time, by means of a gradual interaction of producers and users. Thus, the adjustment gap measures the extent to which the market is far from saturation. When the market becomes saturated, the adjustment gap is reduced to zero or to a small and constant value. The adjustment gap is very large right after the creation of the sector, and later it decreases gradually, although not continuously. It is in fact possible for the adjustment gap to grow during certain periods if innovations, following the one creating the sector, improve either the performance of the product or the efficiency with which it is produced, or both.

Each sector has a dynamics given by the entry and exit of new firms. Schumpeterian entrepreneurs create new firms to exploit a pervasive innovation induced by the expectation of a temporary monopoly. The following bandwagon of imitators raises the intensity of competition and gradually eliminates any further inducement to enter. Thus, the once innovative sector is transformed into a part of the circular flow (Schumpeter, 1912) or into one additional routine of the economic system. This happens when the incumbent sector saturates, a condition which in TEVECON is attained when the adjustment gap AG_i, becomes zero or reaches a very low and constant value (Saviotti, Pyka 2004a, 2008). The saturation of incumbent sectors induces entrepreneurs to search for new niches which could subsequently become new markets. The dynamics briefly outlined above provides a mechanism for the endogenous generation of new sectors which allows the process of economic development to continue in the long run.

A very important role is played in TEVECON by search activities, a general analogue of R&D (Nelson, Winter, 1982). Search activities can be defined as all the activities which try to better understand our external environment and which can provide the basis for the emergence of new routines. Thus, search activities are the source of new innovations and we can expect a positive relationship between the resources allocated to such activities and the rate of creation of innovations. In TEVECON the resources allocated to search activities are expected to increase with accumulated demand:

$$SE_i^t = SE^0 + k_4^i \cdot [1 - \exp(-k_5 \cdot Dacc_i^t)] \tag{1}$$

The combination of the emergence of new sectors and of their increasing quality and internal differentiation leads to an increasing differentiation of the economic system during the process of development. However, this combination can occur in many different proportions giving rise to many development paths. The analysis of the paths is one of the objectives of the present paper. A more detailed description of our TEVECON model can be found in Pyka and Saviotti (2011) and in previous papers (Saviotti and Pyka, 2004a, 2004b, 2008).

Here we describe an extension of our TEVECON model having two objectives. First we want to study the co-evolution of demand and innovation in the process of economic development; second, we want to study the effect of output variety and of output quality and differentiation on economic development paths. Most existing models of growth, including the endogenous growth ones (Aghion and Howitt, 1992; Romer, 1990, Grossman and Helpman, 2001), are supply based and they pay no attention to demand. However, innovation would not have had any impact on economic development if the products embodying specific innovations had not been purchased by consumers and users. Even evolutionary economics, which owed its origin to the difficulties encountered when attempting to use neoclassical economic theory to explain the nature and impact of innovation on economic development, is until predominantly concerned with the supply side. On the other hand, models which focus on demand tend to stress structural change and to belong to a neo-Keynesian approach (Kaldor, 1957; Pasinetti, 1981; Aoki and Yoshikawa 2002). Recently a growing attention has been paid to demand in models of economic growth, both orthodox (Murphy, Shleifer and Vishny, 1989; Matsuyama, 2002; Foellmi and Zweimuller, 2006) and evolutionary (Bianchi, 1998; Andersen, 2001, 2007; Aversi *et al.*, 1999; Metcalfe, 2001; Saviotti, 2001; Witt, 2001; Ciarli *et al.*, 2010). An even more recent paper by Nelson and Consoli (2010) makes the brave attempt to sketch a broad outline of such a demand theory. They explore the use of routines by consumers to guide their choices. In this approach the mechanisms whereby routines are constructed are of crucial importance. In demand as in supply innovation creates uncertainty. Thus, consumers' knowledge is not just likely to be imperfect but to become more so when new types of goods and services completely unknown to them are introduced into the economic system. Especially at the beginning of the life cycle of the emerging goods and services very few consumers are likely to be able to overcoming this uncertainty. In fact, in these circumstances consumers can be expected to act as innovators but to require a threshold level of human capital to do that (Saviotti, 2001).

With respect to these papers ours differs for a number of aspects. First, this paper is part of a research program, the initial objective of which was to prove that economic development has occurred by means of a growing differentiation of the economic system. This objective placed our model not only within evolutionary economics but also with the research tradition of structural change. Furthermore, from the very beginning we were interested in long range patterns of economic development. The relationship between demand and innovation was always present in our model as the potential imbalance between saturating demand and continuously growing productive efficiency (Pasinetti, 1981). However, the specification of demand changed considerably in subsequent versions of TEVECON by first incorporating product quality and differentiation (Saviotti and Pyka, 2008) and becoming for the first time fully endogenous in this paper. The distinguishing features are:

— It does not share most of the assumptions of orthodox models, such as general equilibrium or optimizing behavior, but it only considers economic agents as potential improvers engaged in learning activities.

  The type of structural change that is at the center of the process of economic development leads to a growing output variety of the economic system. Thus, there is in TEVECON an arrow of time continuously raising the differentiation of the economic system. Interestingly, this feature of TEVECON finds a growing validation in recent empirical work (Acemoglu and Zilibotti, 1997; Imbs and Warcziag, 2003; Saviotti and Frenken, 2008; De Benedictis *et al.*, 2009).

— The mechanism whereby disposable income is created is closely related to the growing differentiation of the economic system.

— The growing product quality and differentiation within each sector contributes together with growing output variety to the compensation of the falling ability of mature sectors to create employment.

None of these features is present in the orthodox models referred to above. Furthermore, some of the objectives of the papers referred to above are similar to those of our paper, but they differ in a number of ways. Murphy, Shleifer and Vishny (from now on MSV) (1989) rescue the theory of the big push put forward by Rosenstain and Rodan (1943) in the 1950s by developing a multi-sectoral model in which simultaneous investment in the different sectors of the economy can lead to growth even if no sector individually breaks even. The contribution of simultaneous investment to growth comes from the pecuniary externalities generated by each sector, which increase purchasing power in all sectors. Moreover, growth occurs by each sector shifting from constant returns to scale in cottage industry to increasing returns to scale in factory production. In this sense for MSV it is a change in process technology which gives rise to growth while in TEVECON it is the emergence of new sectors which differ for the type of output they produce. Thus, in MSV neither the type of output of sectors nor the direction in which structural change can be expected to vary, for example towards growing output variety, are defined. On the other hand, we find similarity between the ways in which MSV and our paper deal with demand: in both cases it is the income generated by the investment in industrialization (MSV) or in the emergence of new sectors (TEVECON) which creates the required demand.

With Matsuyama (2002) we share the interest for a similar transition. What we call the transition from necessities to imaginary worlds and the closely related one from low to high quality are very similar to Matsuyama's rise of mass consumption societies. However, with respect to Matsuyama our model differs for (i) the types of learning mechanisms, different types of search activities (fundamental and sectoral in TEVECON) compared to only learning by doing in Matsuyama, (ii) the specification of preferences, non-homothetic for Matsuyama, differing for consumers' propensity to move up or down a hierarchical ladder of goods or services in TEVECON, (iii) the impact of income distribution on development, which is present in Matsuyama and so far not in our model. As for MSV Matsuyama does not characterize the outputs of different sectors, and only allows them to be gradually adopted by different sections of the consumer population as the effect of learning by doing reduces the output cost of each sector making it affordable for larger and larger sections of the consumer population. Thus, Matsuyama includes a form of co-evolution (he talks about two-way causality) and a mechanism which is very similar to our trajectory 1 (growing productive efficiency). However, he has neither any direction of structural change (trajectory 2, growing output variety) nor of growing output quality and differentiation (trajectory 3).

Foellmi and Zweimuller (FZ) (2006) use non-homothetic preferences, hierarchically ordered goods and investigate the effect of income distribution on growth. Their paper differs from Matsuyama (2002) for its learning mechanism, learning by doing in Matsuyama and industrial R&D in Foellmi and Zweimuller, and from MSV due to their claim to apply a more general nature of a preference system and also due to the more dynamical character of their model.

All the three above papers investigate the effect of income distribution on growth but they reach different and sometimes opposing conclusions. For example, FZ find that falling income inequality reduces growth for MSV whereas it increases growth for FZ.

In summary, our paper is part of a research program, one of whose most important objectives is to investigate the process of progressive differentiation which accompanies, and we maintain partly determines, economic development. None of the above papers shares this objective. The extent of differentiation is given. Change occurs by a transition in process technology (MSV), by learning by doing (M), or by industrial R&D technology (FZ). Given this difference in objective, TEVECON is the only model in which the number of sectors is endogenously variable, thus stressing the direction of structural change. From the very beginning the interaction between demand and supply has been at the center of TEVECON in the form of the imbalance between saturating demand and continuously increasing productive efficiency. Aoki and Yoshikawa (2002) share part of this approach. Yet our specification of the co-evolution of demand and innovation has been completed only in recent versions of TEVECON by including disposable income in the sectoral demand function. The goods and services of TEVECON are hierarchically ordered, but what determines the order is the action of entrepreneurs creating new sectors in the expectation of a temporary monopoly. Consumers do not have the ability to anticipate the emergence or nature of future sectors but react to their existence by purchasing their goods and services to the extent that their disposable income and preferences allow them to do. In particular, the preferences of our consumers differ for their propensity to reduce or discard the consumption of older goods and services to start consuming new ones. With respect to MSV, M, and FZ we do not have included in our analysis income distribution but only calculate the average disposable income available for the consumption of new goods and services. The creation of such depends on the growing productive efficiency of older sectors (trajectory 1) and on the income created by the investment in the new sectors.

A further modeling approach which deserves to be discussed for both its similarities and differences with respect to TEVECON is that of Amendola and Gaffard (AG) (1998). AG share with TEVECON the out of equilibrium nature of the model and their emphasis on qualitative change. They include an interesting discussion of the nature of money but in the whole the sources of disequilibrium and the representation of technology are very different from TEVECON. For example, while they talk about qualitative change they do not take into account the non-comparable nature of the product and process technologies which emerge in the course of economic development.

The comparison of ours and of the above papers shows that each of these models investigates different aspects of the economic system and thus that they are not strictly comparable. Within this set of models the specificities of ours are that: (i) it is much 'lighter' in terms of its assumptions than orthodox models since it does not include closure conditions such as general equilibrium or optimizing behavior; it has a particular representation of structural change as leading to a growing output variety; (iii) it has an explicit analysis of the co-evolution of innovation and demand; (iv) it has an explicit representation of product quality and differentiation; (v) it has a more complete representation of search activities, including both fundamental research and sectoral applied research.

The previous references explored the mechanisms of creation of demand in relation to innovation at a micro economic level. In this paper we are more concerned with the joint dynamics of innovation and demand at a meso-economic level of aggregation. Two conditions are required in order for demand for new products or services to emerge:

(i) Consumers must have a disposable income which allows them to purchase the new goods and services;

(ii) Consumers must have or develop preferences which make them value positively the new goods or services.

Here the term disposable income must be understood to be the residual income, left over in a given period, after all the types of consumption of previous periods have been satisfied. A demand function had been introduced into TEVECON in a previous paper (Saviotti and Pyka, 2008). However, the demand function we used in that paper depended on output quality, on output differentiation and on price but not on income. This had the effect of overstating demand since high quality products are always preferred to low quality products irrespective of the consumer purchasing power. In this paper we use a demand function Equation (2) which depends on disposable income and on preferences in addition to product price, quality and differentiation.

$$D_i^t = k_{pref,i} \cdot D_i^0 \cdot D_{Disp,i} \frac{Y_i \cdot \Delta Y_i}{p_i} \tag{2}$$

where

$D_i^t =$ demand for product i at time t

$Y_i =$ services supplied by the product, measuring product quality

$\Delta Y_i =$ range of services supplied by the product, measuring product differentiation

$p_i =$ product price

$D_{Disp,i} =$ disposable income which can be allocated to purchase product i

$k_{pref,i} =$ parameter representing preferences

We calculate $D_{Disp,i}$ as the difference between the total income and the income required to satisfy the types of consumption of previous periods in period t.

To study how different preference systems can affect the time path of demand and of economic development we represent three very simplified preference systems which we call progressive, conservative and random. We realize that in a real economic system, preference systems of these different types would be distributed within a consumer population and that they would not be immutable. Consumers can learn and change their preferences in the course of time. Our main objective here is simply to show that consumer preferences can affect directly demand and indirectly the macroeconomic growth performance of the economic system.

Consumers with a progressive preference system value more highly new goods and services than older ones. Consumers with a conservative preference system value more highly old goods and services than newer ones. Consumers with a random preference system will have preferences randomly distributed amongst the outputs of different sectors, old an d new. These three preference systems are represented as three different parameters in the demand Equation (1). $k_{pref,i}$ is a parameter which is constant for each sector in the course of time but can vary between different sectors. The three preference systems are then represented as follows:

— Progressive preference system: $k_{pref,\,i+1} > k_{pref,i}$

— Conservative preference system: $k_{pref,\,i+1} < k_{pref,i}$

— Random preference system: $k_{pref,\,i+1} >< k_{pref,i}$

The second objective of the paper consisted of comparing the economic development paths which would be obtained when product quality (i) remained unchanged or (ii) increased during the life cycle of each sector in TEVECON. This objective is attained by modifying the values of the parameters k_{14}-k_{17} linking search activities to product quality and differentiation Equations (3), (4)

$$Y_i^t = \frac{1}{1 + exp(k_{14} - k_{15}SE_i^{\,t})} \tag{3}$$

$$\Delta Y^t{}_i = \frac{1}{1 + exp(k_{16} - k_{17}SE_i^{\,t})} \tag{4}$$

When these parameters have extremely low values product quality and differentiation remain virtually constant during the evolution of the respective sectors. Values of the parameters k_{14}-k_{17} are varied by giving them extremely low values in the low quality (LQ) scenario and considerably higher values in the high quality (HQ) scenario. Thus, in the LQ scenario the saturation of each sector is attained much more rapidly due to the absence of quality change in sectoral outputs. In other words, in the LQ scenario market saturation occurs only by volume (Saviotti, Pyka, Krafft, 2007). On the other hand, in the HQ scenario market saturation can occur much later, giving rise to longer industry life cycles (ILC) because the market can still expand after volume saturation has been attained by moving towards products of higher quality and thus of higher value.

By recalling that according to equation 1 search activities increase with accumulated demand and by combining equation 1 with equations 3 and 4 we can realize that search activities depend on demand and demand depends on search activities. This is the basis for the co-evolution of innovation and demand. The co-evolutionary loop is completed by equation 2 according to which demand is not only affected by three variables which are themselves affected by search activities (Y_i, ΔY_i and p_i) but also by the presence of a disposable income which can be used to purchase new goods and services.

Human capital is created by investment in education, which gives rise to an education capital stock (CS_{edi}), which in turn determines the quality h_i of human capital Equation (5). The parameter k_{ed} represents the effectiveness with which the investment in education is transformed into human capital. Hence, k_{ed} represents the quality of educational institutions in forming human capital. Overall human capital is obtained by multiplying sectoral labor by the quality h_i of human capital Equation (6).

$$h_i^t = k_{ed} \cdot CS_{ed\,i}^{\ t} \tag{5}$$

$$HC_i^t = labour_i^t \cdot h_i^t \tag{6}$$

Bearing in mind that sectoral output depends on human capital, we can realize that the time path of output depends on investment in education and on the effectiveness with which educational institutions improve the quality of human capital. Furthermore, the intensity of production is determined by the parameter k_{HQ} see Equation (7). Equation 7 also shows that human capital in a given period depends on investment in previous periods, which itself depends on output in previous periods. In turn, future output is affected by present human capital. Here we see some more examples of the co-evolutionary patterns included in TEVECON.

$$Q_i^t = Q^0 + \gamma \cdot (1 + \alpha_{ci}^t) \cdot \\ (1 - exp(-k_{11} \cdot SE_i^t - k_{cspq1} \cdot CSphysical_i^t - k_{11} \cdot HC_i^t) \tag{7}$$

$Q_i^t :=$ sectoral output

$\gamma :=$ scaling parameter

$\alpha_{ci}^t :=$ production adjustment

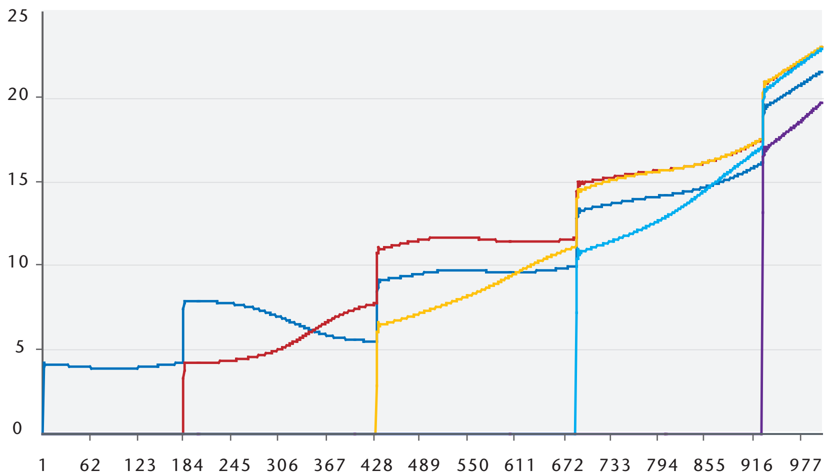
Wages depend on labor productivity and on a parameter, k_{wages}, Equation (8). The parameter k_{wages} leads to an increase or a decrease in wages at equivalent labor productivity. Thus, it could reflect the presence of particularly powerful labor unions, which would tend to raise it, or of reforms in the labor market, which could reduce it. We expect that at equivalent labor productivity a low value of k_{wages} increases the competitiveness of a sector or of a country.

$$wages_i^t = w_i^0 + k_{wages} \cdot \frac{Q_i^t \cdot P_i^t}{labour_i^t} \tag{8}$$

1.4. Disposable income for new sectors

Our calculations show that under a wide range of circumstances a disposable income can be created for new sectors, thus allowing consumers to purchase their output Figure 1.

Figure 1. Effect of product quality on the disposable income created in the economic system



Further, we can observe that while to purchase the output of sector 2 a reduction of the expenditures on sector 1 is required, such a sacrifice is not necessary for subsequent sectors. The development of the economic system manages to create enough resources in the system to allow consumers to purchase the new goods and services. The mechanisms by means of which such increasing purchasing power is created are related to the three trajectories described above. First, the growing productive efficiency in incumbent sectors (trajectory 1) reduces the cost of those sectors' goods and services and creates a surplus which can be used to fund the search activities and the investment required to produce the new goods and services. Second, the previous investment creates income for the labor employed in the production of the new goods and services. Third, as the average revenues of the population increase the possibility to make higher quality, more expensive and more profitable goods and services emerge. To the extent that such new goods and services fit consumers' preferences they will create new markets or enlarge existing ones. Thus, the growing quality and differentiation of goods and services (trajectory 3) together with the emergence of new sectors (trajectory 2) can compensate the falling ability of incumbent sectors to create employment and enable growth to continue in the long run. While this conclusion expands the range of possible growth mechanisms, in relatively wealthy economic systems it also introduces a source of uncertainty. In fact compensation can occur only if the innovations required to create new sectors are available when the saturation of pre-existing ones occurs. While this has been assumed so far in TEVECON there is no guarantee that in a real economic system this will always occur.

1.5. Preferences

The existence of an adequate disposable income is a necessary condition for consumers to be able to purchase the new goods and services which are created by innovation. However, consumers will do that only if they have an adequate set of preferences. In this section we study how the three different preference systems we suggested in the previous section can affect the time path of demand and of economic development. We realize that that these representations of a preference system are an approximation.

However, we consider that such an approximation is sufficient for our main objective here, which is to show that consumer preferences can affect directly demand and indirectly the macroeconomic growth performance of the economic system.

In different experiments we vary the degree of progressiveness or of conservativeness of our consumers by changing the Δk_{pref} between sectors i and i+1. Thus, a large and positive Δk_{pref} between sectors i and i+1 indicate strongly progressive consumers while a smaller but still positive Δk_{pref} indicate mildly progressive consumers. Likewise, a large and negative Δk_{pref} between sectors i and i+1 indicate strongly conservative consumers while a smaller negative Δk_{pref} indicate mildly conservative consumers. The results of these experiments are summarized in figures 2 and 3 by plotting the straight lines which give the rate of growth of income (Figure 2) and of employment (Figure 3). Such straight lines are the best linear fit for the income and employment curves and their slopes give us the rate of growth of income (RIG) and the rate of growth of employment (REG) respectively (see Saviotti and Pyka, 2008).

Figure 2. Influence of the different preference systems on the rate of growth of income

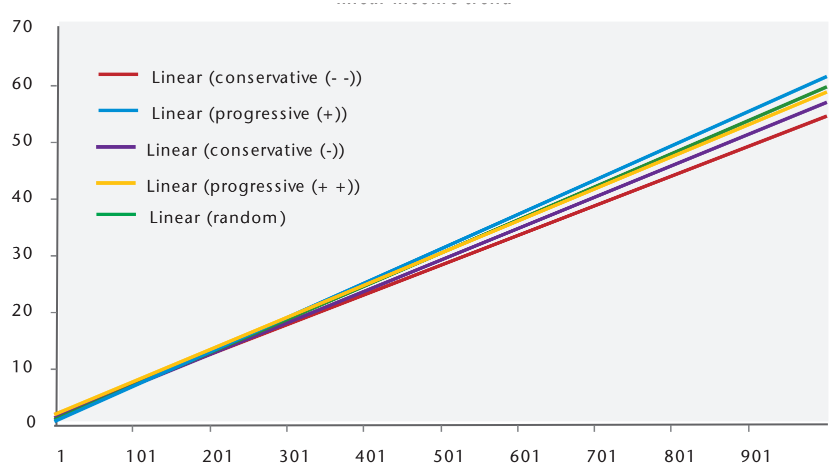
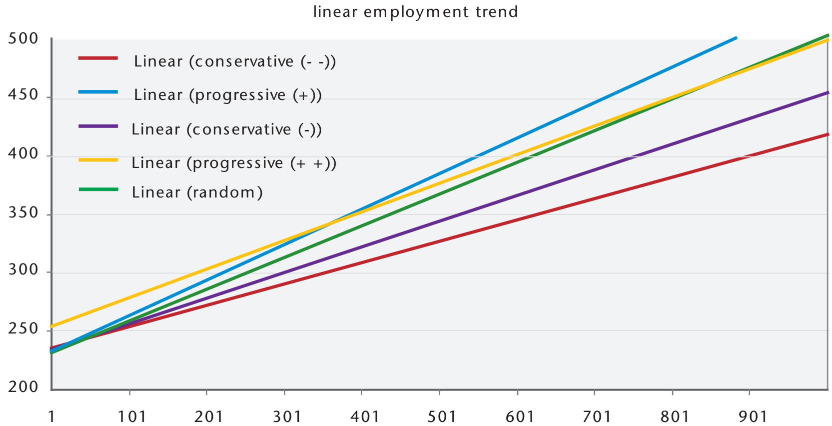


Figure 3. Influence of the different preference systems on the rate of growth of employment



These results show that both REG and RIG increase when preferences pass from conservative to neutral to progressive. However, as more preferences become more and more progressive both, REG and RIG start falling indicating the presence of a non-linear relationship between preferences on the one hand and employment or income on the other hand. Such non-linearity can be explained because the change from conservative to neutral to progressive preferences implies a transfer of resources from the purchase of old goods and services to that of emerging ones. While a moderate transfer can accelerate the emergence of new sectors, an excessive one can depress the demand for older goods and services and thereby reduce the overall growth of employment and of income.

The results of sections 1.4 and 1.5 show that (i) disposable income for new goods and services can be created by a combination of trajectories 1, 2 and 3, corresponding to the growing productive efficiency in incumbent sectors (trajectory 1), to the emergence of new sectors (trajectory 2) and to the growing quality and differentiation of goods and services (trajectory 3); (ii) consumer preferences can affect the macroeconomic performance of the economic system. We now pass to the second objective of this paper.

1.6. On the balance between the emergence of new sectors and the growing quality and differentiation of existing ones

To study this problem we define a set of parameter values which seem to give the type of regular pattern of development we had detected in previous papers. In other words, we started from a situation in which new sectors were regularly created and where the aggregate rates of growth of employment and of outcome were positive. We called this set of parameters our standard scenario.

Although many combinations of the emergence of new sectors and of the growing quality and differentiation of goods and services can be envisaged, we can in principle expect such different combinations to give rise to different development paths. To explore the relative impact of the emergence of new sectors and of the growing quality and differentiation of goods and services, we simulate two development scenarios, called high quality (HQ) and low quality (LQ) respectively. These scenarios are obtained by giving different values to the parameters k_{14}-k_{17} of Equations (3) and (4). These parameters determine the extent of product quality and differentiation corresponding to a given level of search activities. The LQ scenario is obtained by giving the parameters k_{15} and k_{17} values so low that product quality and differentiation are almost constant during the ILC of the sector. The HQ scenario is obtained by giving the same parameters considerably higher values. The results of this simulation show that the HQ and LQ scenarios give rise to very different development paths. The comparison HQ-LQ was explored by means of both micro- and macro-economic variables. In the LQ scenario, demand, human capital, wages and output remain substantially static or even declining while they increase in the HQ scenario (Figures 4a, b and c).

At an aggregate level:

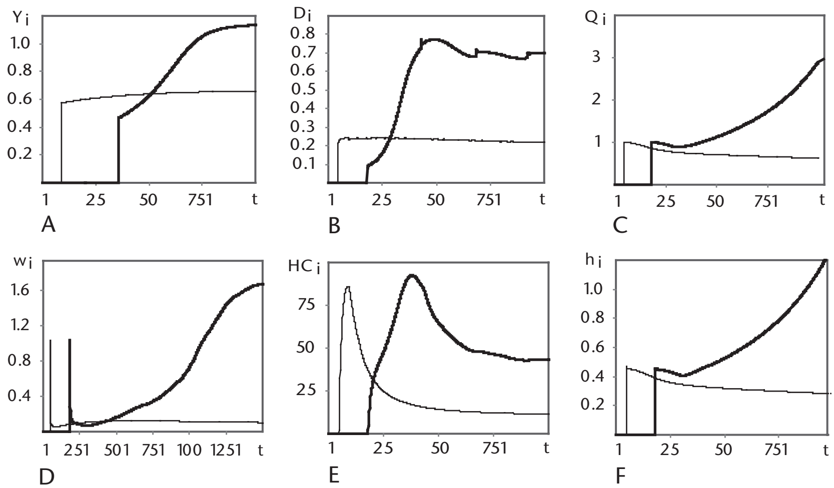
— Disposable Income grows faster in the low quality case with respect to the high quality case (Figures 5 and 1)

— Employment growth is always faster in the low quality case with respect to the high quality case (Figure 6b)

— The rate of creation of new sectors is higher in the low quality case with respect to the high quality case

— The rate of income growth (RIG) of the HQ scenario is initially lower but it overtakes that for the LQ scenario at a later time (Figure 6). We can also notice that RIG slows down in the course of economic development for the LQ scenario while it accelerates for the HQ scenario.

Figure 4. Product quality



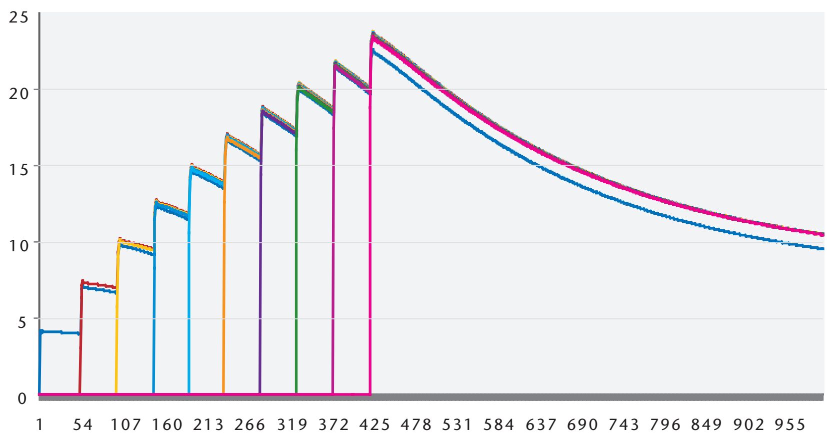
A - Product quality, as measured by the services supplied by a product (Yi) in the low quality (thin curve) or high quality (bold curve) case
B - Effect of product quality on sectoral demand
C - Effect of product quality on sectoral output
D - Effect of product quality on sectoralwages
E - Effect of product quality on thequantity of human capital used in a sector
F - Effect of product quality on thequality of human capital used in a sector

The above results can be explained as follows as follows:

Constant wages and constant human capital limit the scope for income growth in the LQ case. The absence of increases in quality and in sectoral differentiation in the LQ case, lead to shorter industry life cycles (ILC) and to a higher rate of creation of new sectors. Since the rate of employment growth (REG) is higher in the early phases of an ILC, the aggregate REG is higher for the LQ than for the HQ scenario, although such higher REG is obtained at the expense of lower wages, lower demand and lower human capital.

Initially the higher REG leads to a higher RIG for the LQ scenario. However, the rising wages and demand lead to a RIG which is not constant but increases in the course of economic development for the HQ scenario. The self-accelerating and self-limiting shapes of the RIG curves for HQ and LQ scenarios can be understood because in the former case an increase in demand leads to an increase in search activities, which in turn leads to an increase in output quality and differentiation, which is finally translated into an increase in demand. This feedback loop is considerably weakened in the LQ scenario because in this case search activities have a negligible impact on output quality and differentiation.

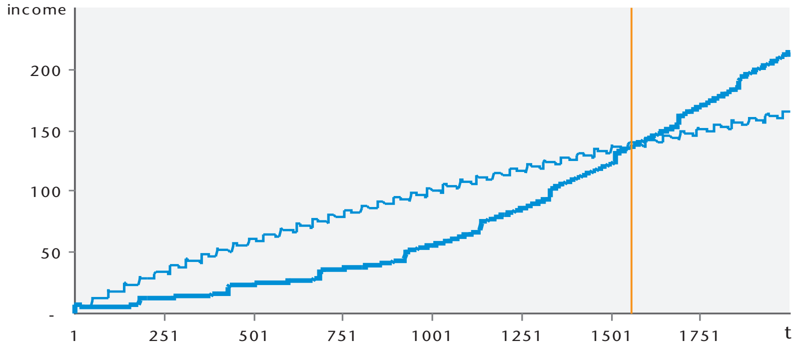
Figure 5. Effect of product quality on the disposable income created in the economic system for the low quality case



To interpret the previous results we note that empirical observations show that product differentiation started considerably after the beginning of the industrial revolution, probably towards the end of the 19th century, and initially only in relatively rich countries. Such transition proceeded by liberating a growing proportion of household income from necessities and thus making room for the purchase of new goods and services which were not necessary in the physical sense in which food or shelter are (see Hobsbawm, 1968, diagrams 45 and 46). Rather than being necessities, the result of adaptation to the external environment in which human beings live, the new goods and services shape the external environment in ways which were not necessary and along a development path which was not necessarily unique. Thus, we described the evolution of the capitalist economic system as the transition from necessities to imaginary worlds. This transition could be interpreted as the result of a continuous, linear progress which constantly improves human welfare. We think that such an interpretation would be rather simplistic. We are more interested in understanding how the mechanisms which we explore in this paper, however oversimplified, could provide us with an explanation of how the capitalist economic system managed to survive since the industrial revolution by profoundly transforming itself. Every economic system, however successful at the time it is created, brings in itself the seeds of its own destruction. Such destruction need not necessarily occur if the economic system manages to transform itself enough.

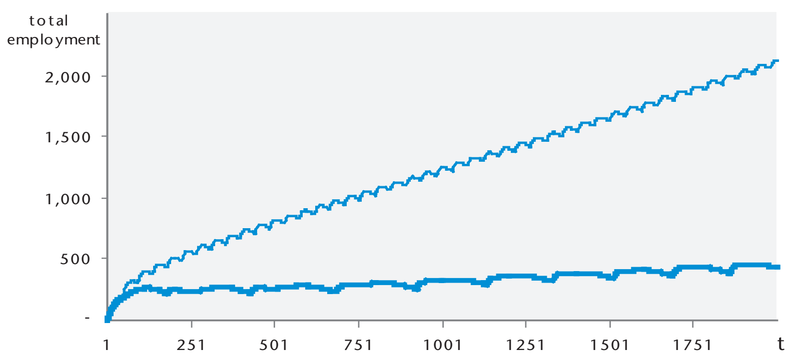
Figure 6. Effect of product quality on the aggregate rate of income growth

(a) Income



(LQ light curve, HQ bold curve); The vertical line indicates the time required for HQ income to catch up with and to overtake LQ income, which we call ICUT.

(b) Total employment



(LQ light curve, HQ bold curve)

The development mechanism we hypothesize began with the saturation of the markets for necessities, attained during the early part of the industrial revolution due to the growth of productive efficiency which occurred in that period. In turn, that saturation is likely to have induced efforts by producers to avoid it by opening new markets or by enlarging existing ones. Assuming that new technologies potentially giving rise to new markets could be created, as they were, the markets themselves would not come into being unless a large enough percentage of the population had the required purchasing power. A mechanism which could give rise to the coordinated emergence of production capabilities and of purchasing power is the following:

— The production of some of the new goods and services and the rising quality and differentiation of existing ones required higher levels of competencies and of human capital;
— such higher competencies required training and education;
— better educated workers had to be paid higher wages;
— new jobs were created in the training and education system;
— the new jobs and higher wages created the disposable income required to purchase the new goods and higher quality goods and services.

The combination of the above steps gave rise to a virtuous circle which could continue expanding the economic system as long as technologies and demand could co-evolve. This co-evolution allowed the capitalist economic system to escape the development trap which Marx and other critics of capitalism had foreseen. Of course, we think that the mechanism previously described is only a component of an overall repertoire. The capitalist economic system cannot have been saved only by an ever increasing shopping frenzy of new and more luxurious goods and services. Social innovations in pensions, unemployment benefits, health care etc. are likely to have co-evolved together with the mechanism described above to allow the capitalist economic system to transform and adapt. Thus, the real co-evolution included more mechanisms and steps than the ones we described above. However, we think our exercise is useful because it provides an analytical approach to the explanation of long range transitions in economic systems. The addition of further components to the co-evolutionary process described above can be envisaged without substantial modifications of our approach.

Let us observe that the transition from low to high quality goods and services, henceforth (LQ → HQ) transition, is not identical to that from necessities to imaginary worlds. The former is from an economic system dominated by trajectories 1 and 2 to one dominated by tranjectories 1, 2 and 3, while the latter is from an economic development dominated predominantly by trajectory 1 for consumer goods but with trajectory 2 occurring in capital goods. In its present state TEVECON cannot accurately distinguish between consumer and capital goods. In spite of these differences the transition (LQ → HQ) is very similar to that from necessities to imaginary worlds, especially for what concerns the emergence of higher quality and internally differentiated goods and services. Thus, the study of the (LQ → HQ) transition can help us understand the mechanisms of capitalist economic development.

The analysis we carried out shows that long range processes of economic development cannot be explained only by the increasing productive efficiency, or even by the increasing output quality, of a constant set of activities, but that they intrinsically involve a very high degree of structural change. In this context structural change not only means the changing weight of different sectors but also other changes in the composition of the economic system, with the inclusion of completely new institutions and organizations and of their interactions. Structural change becomes more important for the explanation of processes of economic development the longer the time horizon chosen.

We now describe a set of policy relevant experiments carried out with TEVECON.

2. Policy experiments

In these experiments we explore the effects of changes in a number of TEVECON parameters on some aspects of the process of economic development. In particular, we focus on the role of human capital and of wages. According to the above described mechanisms we can expect that both human capital and wages had to increase to allow the economic system to generate the higher quality goods and services and the income required to purchase them. Thus, we chose to modify some parameters which affect these two variables. First, we hypothesized that at least in some types of economic activities there could be a barrier in human capital. In these activities only human capital above this barrier could be employed. Second, we hypothesized that the weight of human capital in the production function could affect economic development processes. Third, we expected wages to affect economic development processes. In TEVECON wages are proportional to labor productivity according to a parameter k_w, henceforth called the wage parameter. Accordingly, in our experiments we vary the barrier in human capital, the weight of human capital in the production function and the wage parameter. We start by varying one parameter at a time and then we combined variations of two or more parameters (Table 1, Appendix).

The starting point of our experiments here was the comparison of the LQ and HQ scenarios described in Figures 6 a, b. These results show, that (i) the rate of employment growth (REG) is systematically higher in the LQ scenario, and that (ii) the rate of income growth (RIG) is initially higher for the LQ scenario but becomes higher for the HQ scenario at later times. In the following experiments we investigate the impact of the three above parameters on (i) the time required for HQ income to catch up and overtake LQ income, which we called ICUT, (ii) the relative REG for the two scenarios, and on (iii) the variance of income determined by the change from conservative (CP) to progressive (PP) preferences. ICUT was measured as the time at which the HQ income crossed the LQ income curve (see Figure 6a). ICUT is plotted as a function of the weight of human capital in the production function (Figure 7) and of the wage parameter k_w (Figure 8).

The most general trend observed is a fall in ICUT when both k_{Hi} or k_w increase. This means that the (LQ → HQ) transition would have occurred earlier if a higher intensity of human capital and a higher wage rate had been used in the economic system. However, the behavior of ICUT becomes more complex when the increases in the above two parameters are combined with increasing values of B_{hi}. In this case ICUT alternately rises or falls for different ranges of values of either B_{hi} or k_w. These more complex types of behavior could be understood by bearing in mind that the introduction of a human capital barrier excludes some workers from the labor force.

Figure 7. Effect of changing the weight k_{Hi} of Hi in the production function for different values of barrier in human capital B_{hi}

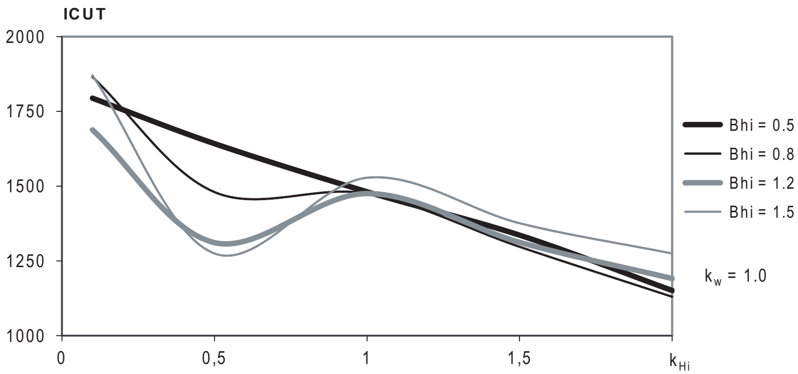
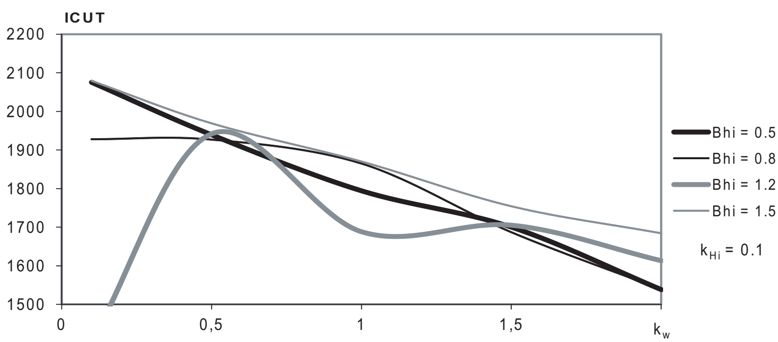


Figure 8. Effect of changing the wage parameter k_w for different values of barrier in human capital

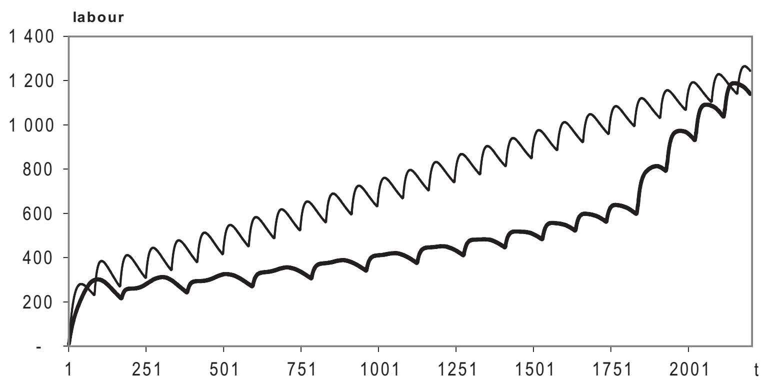


The resulting outcome would be due to the balance between the higher wages of the employed workers and the absent wages of the unemployed ones. The general point to be made here is that wages are both a source of costs and of revenues. The effect of rising wage rates and of rising levels and intensity of human capital depends on the balance of their effects on revenues and on costs. Also, we have to bear in mind that the introduction of an h_i barrier in the present state of TEVECON is equivalent to an internal differentiation of the labor force. Thus, introduction of a low h_i barrier into an economic system which has low wages and low human capital can have a very different effect than the introduction of a higher h_i barrier into an economic system which has high wages and high levels and intensity of human capital. The effect of the human capital barrier on the ICUT falls for higher values of both k_{Hi} and k_w. Thus, a system which already has high wages and high levels and intensity of human capital is less affected by the introduction of a human capital barrier than a system which has low wages and low levels and intensity of human capital.

Finally, we can observe that the LQ income curve is virtually unaffected by the changes in the three above parameters. This is the result of the fact that human capital and output quality are almost constant in the LQ case.

The same set of experiments described in table 1 was carried out for the relative REG of the LQ and HQ scenarios. The result described in Figure 6b showed that REG(LQ) was systematically higher than REG(HQ). In fact, the two curves diverged continuously. Furthermore, both curves were approximately linear in time. In the vast majority of the experiments we carried out REG(LQ) was greater than REG(HQ). However, for particular values of the parameters used, REG(HQ) increased considerably showing an inflection point in the employment curve (see for example Figure 9).

Figure 9. Employment for the LQ (light line) and HQ (heavy line) scenarios for B_{hi} = 0.5 and k_{Hi} = 2.0.



After the inflection pint the HQ employment curve can sometimes overtake the LQ one. The inflection point occurs at very long development times, which correspond to high levels of economic development. In other words, similarly to what happened for

Income, the evolution of employment shows a self-accelerating character in the HQ which is absent in the LQ case. In the HQ employment case this self-accelerating character seems to arise fairly suddenly while in the HQ income case it was continuous. However, even in the HQ employment case we can see premonitory signs of the inflection in the shortening of the ILCs which starts occurring form the beginning, a phenomenon which does not occur at all for the LQ scenario. Such shortening of the ILCs can be explained by (i) the increasing quality and internal differentiation of goods and services which lengthens the life cycles of the sectors producing them, as it can be seen by comparing the LQ and HQ cases (see also Saviotti, Pyka and Krafft, 2007); (ii) the increasing quality and internal differentiation of output can be become faster the more knowledge creating resources are present in the economic system. As in the income case the employment curve of the LQ scenario is almost unaffected by the changes in parameters used in our experiments. As for the income case this different sensitivity of the LQ and HQ cases to changes in parameters affecting human capital or wages can be explained by the much weaker feedback loop between demand and search activities existing in the LQ case.

The relative dynamics of income in the LQ and HQ scenarios is affected also by a change in preferences. Figures 10 and 11 compare the impact of preferences on the income curves for the LQ and HQ scenario with different parameter settings. Figure 10 corresponds to our standard scenario (Experiment 1 in Table 1) while Figure 11 corresponds to experiment 27. The results can be summarized as follows:

(i) The variance in income induced by a change of preferences from conservative (CP) to random (RP) and then to progressive (PP) for both the LQ and LQ cases increases in the course of time, that is the more highly developed an economic system is. In what follows we call this variance PIVI and we measure it as the difference between the income levels corresponding to PP and CP respectively at the maximum time at which we ran our model (the intercepts of the income curves with the vertical axis on the right of the diagram).

(ii) At equivalent times in our standard scenario (Exp 1, table 1) PIVI is larger for the LQ than for the HQ case.

(iii) When the barrier to human capital, the weight of human capital in the production function and the wage parameter are increased, either individually or in combination (Exps 2 – 44, Table 1) PIVI grows also for the HQ case and it can become comparable to that of the LQ case.

(iv) For very high values of k_{Hi} the income curve for the HQ case starts growing very rapidly at fairly long times and then abruptly stops. In these conditions the process of economic development becomes so unbalanced that it cannot proceed any further.

Figure 10. Income curves for the LQ (green curves) and HQ (blue curves) cases showing the impact of different preferences on income generation. The parameter settings correspond to the standard scenario

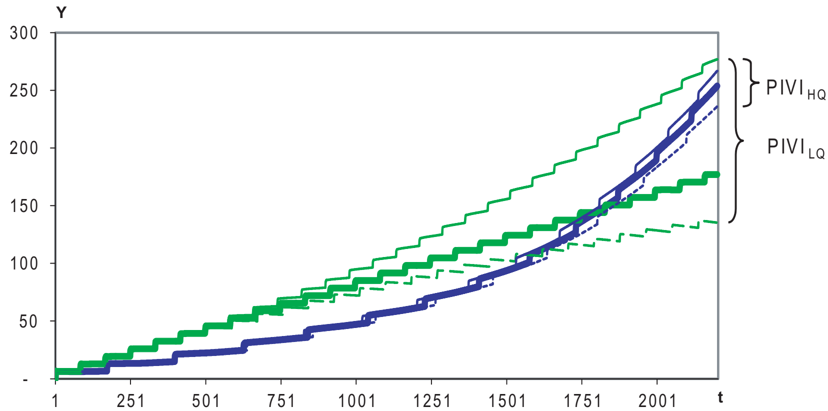
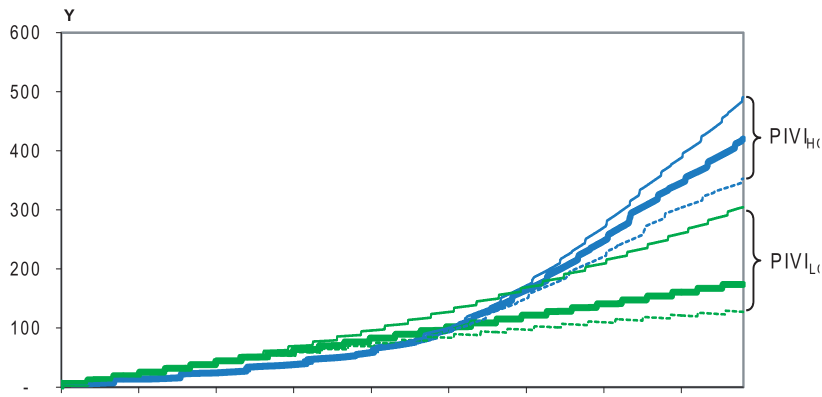


Figure 11. Income curves for the LQ (green curves) and HQ (blue curves) cases showing the impact of different preferences on income generation. The parameter settings correspond to higher values of the B_{hi} barrier in human capital and in the weight of human capital in the production function (Exp 27 in table 1)



The previous results can be interpreted as implying that the impact of changing preferences is likely to increase as economic development proceeds. In other words, differences in consumer preferences are likely to have a greater effect on the growth of income on those which are already rich than on relatively poor ones. In an economic system in which most people can just afford basic necessities the disposable income required to buy new goods and services would be absent or very scarce. In these conditions preferences could hardly exert any impact on income generation. Preferences can be expected to start exerting an impact when there is a disposable income with which consumers could choose to purchase different goods and services in addition to necessities.

While the previous conclusion makes sense in general it is not immediately clear why different preferences should have a greater impact on income formation in the LQ than on the HQ case. If we remember that in the LQ case the only choice consumers could have is that amongst different types of goods and services, but that within each type quality remains constant. As a consequence, ILCs would be shorter and the rate of growth of disposable income would initially be faster. Yet the effect of preferences on PIVI would still be lower even after HQ income had overtaken LQ income. PIVI for the HQ case can start growing and become comparable to that of the LQ case only after barriers to human capital, a higher weight of human capital in the production function and a higher wage rate had been introduced. Thus, although for both the LQ and the HQ cases different preferences start exerting an effect on the rate of growth of income, the time at which preferences start affecting income varies depending on the case and on the parameter setting used in the experiments. In particular, barriers to human capital, a higher weight of human capital in the production function and a higher wage rate seem to have a much higher impact on the HQ than on the LQ case. This in understandable because both levels of human capital and wages remain relatively flat in the LQ case, while they increase in the HQ case.

Let us conclude this section by pointing out that the term policy needs to be interpreted carefully in this context. Usually policies have a relatively short term orientation with respect to the time horizon we are envisaging in this paper. The parameters the influence of which we explored are related to human capital and to wages, two variables the importance of which in modern economic systems is still, and is likely to remain, very high. We have seen that rising wages and rising levels and intensity of human capital played a fundamental role in capitalist economic development by allowing to create both the competencies required to produce goods and services of higher quality and internal differentiation and the disposable income required to purchase them. These results cannot be interpreted as implying that economic development will always be positively affected by raising wages and levels and intensity of human capital. There are many examples in which a reduction in wages can positively contribute to economic performance. What matters is not wages per se but the combination of wages, human capital and other variables. Thus, even if rising wages and levels and intensity of human capital are required to sustain the long term development of the economic system, short term adjustments in their *combination* can be required to compensate for temporary slowdowns or bottlenecks. What matters is not wages or human capital per se but the way in which their co-evolution can create in a coordinated way new demand and the required purchasing power and preferences.

As for preferences, it is quite clear from our results that their impact on growth and development becomes increasingly important as the economic system becomes richer. As a consequence, the scope for activities which help consumers to form preferences for emerging goods and services increases with the level of economic development. This is particularly true for high levels and intensities of human capital and for high wage rates. However, we must remember that if facilitating the formation of preferences for emerging goods and services can positively affect economic development, a balance must be maintained in the economic system between speeding up the introduction of new goods and services and reduce the weight of pre-existing ones.

3. Conclusions

In this paper we study the co-evolution of innovation and demand and try to understand how it could have contributed to the long run development of the capitalist economic system by means of our TEVECON model. We show that the economic system can create the disposable income required for consumers to be able to purchase the new, higher quality and more differentiated goods and services created by innovation. The creation of such disposable income is due to the combination of growing productive efficiency (trajectory 1), growing variety (trajectory 2) and growing output quality and internal differentiation (trajectory 3). Furthermore, we show that consumer preferences can affect observed macroeconomic development paths. In particular, we show that consumers with progressive preferences led to higher rates of growth of output and of income than consumers with conservative preferences, where progressive preferences imply a strong relative propensity to purchase new goods and services at the expense of older ones. Thus, our results confirm that demand matters and that observed patterns of economic development can be explained by the co-evolution of innovation and demand.

After having established this point we explore the economic development paths that could be generated by different combinations of growing variety (trajectory 2) and growing output quality and internal differentiation (trajectory 3). This is done by choosing two rather extreme scenarios, one including only growing variety, which we called low quality (LQ), and one including both growing variety and growing output quality and internal differentiation, which we called high quality (HQ). The HQ scenario gives rise to a slower but richer growth path. The LQ scenario has a higher rate of creation of new sectors and consequently a higher rate of growth of employment but at the expense of having lower wages, lower sectoral demand and lower levels of human capital.

An important result of our comparison was that the HQ income was initially lower than the LQ one, but that at later times the situation was reversed with HQ income becoming dominant. We called this phenomenon the (LQ → HQ) transition. This is important because it seems to map some observed paths of economic development, in particular what we call the transition from necessities to imaginary worlds. Admittedly the two transitions are not identical but they both include the emergence of higher quality and more internally differentiated goods and services at a later stage of economic development. We then explore further the (LQ → HQ) transition to better understand long run mechanisms of economic development. To do this we vary some TEVECON parameters affecting human capital and wages. We find that growing wages and growing levels and intensity of human capital favour long run economic development. We then hypothesize that the (LQ → HQ) transition could have been the outcome of a virtuous circle in which growing human capital and growing wages provide both the competencies needed to produce higher quality and more internally differentiated goods and services as well as the disposable income required to purchase them. Our TEVECON model proves that this virtuous circle is possible but that it is not necessary. As in all co-evolutionary processes the necessary ingredients are required with the appropriate coordination.

Furthermore, we show that the LQ and HQ cases are both affected, although differently, by changing consumer preferences. In both cases, the variance in income produced by progressive (PP) and conservative (CP) consumer preferences tend to grow as economic systems become progressively richer. This points towards an important scope for policy, especially for those activities which help consumers to learn about new goods and services, a necessary condition to for them to have clear preferences.

We conclude this paper by pointing out that the policy implications we can derive here are long term. Thus, we have seen that growing wages and growing levels and intensity of human capital favour long run economic development. This conclusion cannot be translated into the short term prescription to keep raising wages and levels and intensity of human capital under any circumstances. What matters are not the individual values or trends of wages and of human capital but their combinations. Many adaptations can be required to overcome short term bottlenecks and to restore long run trends.

References

Acemoglu D., and F. Zilibotti, 1997. "Was Prometheus unbound by chance? Risk, diversification and growth." *Journal of Political Economy*. 105 (4) 709–751

Aghion P., and P. Howitt, 1992. "A model of growth through creative destruction." *Econometrica*, 60: 323–35

Amendola M., and J. L. Gaffard, 1998. *Out of Equilibrium*. Oxford, Oxford University Press.

Aoki M., and H. Yoshikawa, 2002. "Demand saturation-creation and growth." *Journal of Economic Behavior and Organization*, Vol 48, 127–154.

Andersen E.S., 2007. "Innovation and demand." in Hanusch H. Pyka A. (Eds) *Elgar Companion to Neo-Schumpeterian Economics*, Cheltenham, Edward Elgar.

Andersen E.S., 2001. "Satiation in an evolutionary model of structural economic dynamics." in Witt U. (Ed) *Escaping Satiation, the Demand Side of Economic Growth*, Berlin, Springer, pp. 165–186.

Aversi R., G. Dosi, G. Fagiolo, M. Meacci, and C. Olivetti, 1999. "Demand dynamics with socially evolving preferences." *Industrial and Corporate Change*, Vol. 8, 353–399

Bianchi M., (Ed) 1998. *The Active Consumer*, Routledge, London.

Ciarli T., A. Lorentz, M. Savona, and M. Valente, 2010. "The Effect Of Consumption And Production Structure On Growth And Distribution. A Micro To Macro Model." *Metroeconomica* 61:1, 180–218.

De Benedictis L., M. Gallegati, and M. Tamberi, 2009. "Overall trade specialization and economic development: countries diversify." *Review of World Economics*.

Diamond J., 1997. *Guns, Germs and Steel: the Fates of Human Societies*. New York, Norton.

Dosi G., 1982. "Technological Paradigms and Technological Trajectories: a Suggested Interpretation of the Determinants and Directions of Technical Change." *Research Policy*, Vol. 11, 147–162

Foellmi R., and J. Zweimuller, 2006. "Income distribution and demand induced innovations." *Review of Economic Studies* 82, 95−112

Hobsbawm E., 1968. *Industry and empire*, Harmondsworth, Penguin Books.

Imbs J., and R. Wacziarg, 2003. Stages of Diversification. *The American Economic Review*, 93 (1), 63–86.

Kaldor N., 1957. "A model of economic growth." *Economic Journal*, Vol 67, pp 591–624.

Kirman, 1998. "Self-organization and evolution." in Schweitzer F., Silverberg G. (Eds) Evolution und Selbstorganisation in der Ökonomie, Jahrbuch für complexität in den Natur, Sozial und Geissteswissenschaften, Vol 9. Duncker & Humboldt, Berlin.

Maddison A., 2007. *Contours of the World Economy*. 1-2030 AD, Oxford, Oxford University Press.

Matsuyama K., 2002. "The Rise of Mass Consumption Societies." *Journal of Political Economy* 110, 1035–70.

Metcalfe J.S., 2001. "Consumption, preferences and the evolutionary agenda." *Journal of Evolutionary Economics*, Vol 11, 37–58

Murmann J.P., 2003. *Knowledge and Competitive Advantage: the Co-evolution of Firms Technology and National Institutions*. Cambridge, Cambridge University Press.

Murphy K., A. Schleifer, and R. Vishny, 1989. "Industrialization and the Big Push." *Journal of Political Economy* 97,1003–26.

Nelson R. R., 1994. "The Co-evolution of Technology, Industrial Structure, and Supporting institutions." *Industrial and Corporate Change*, Vol 3, N° 1, pp. 47–63.

Nelson R. R., 2008. "What enables rapid economic progress: What are the needed institutions?." *Research Policy* 37, 1–11.

Nelson R. R., and D. Consoli, 2010. "An evolutionary theory of household consumption behaviour." *Journal of Evolutionary Economics*, Vol 20, 665–687.

Nelson R., and Winter S., 1982. *An Evolutionary Theory of Economic Change*. Cambridge, Mass, Harvard University Press.

Nicolis G., and I. Prigogine, 1989. *Exploring Complexity: an introduction*. Freeman New York.

Pasinetti L. L., 1981. *Structural Change and Economic Growth*. Cambridge, Cambridge University Press.

Perez C., 1983. "Structural change and the assimilation of new technologies in the economic system." *Futures*, Vol. 15, 357–375.

Pyka A., and G. Fagiolo, 2007. "Agent based modeling: a methodology for neo-Schumpeterian – Economics in Hanusch." H. Pyka A. (Eds) *Elgar Companion to Neo-Schumpeterian Economics*, Cheltenham, Edward Elgar.

Pyka A., and P. P. Saviotti, 2011. "Economic Growth through the Emergence of New Sectors." in: Mann, S. (ed.), *Sectors Matter-Exploring Mesoeconomics*, Springer, Heidelberg, Dordrecht, London, New York, 55–102.

Pyka A., and P. P. Saviotti, 2012. *Economic Development – More Creation than Destruction, in: Krämer, H., Kurz, H. and Trautwein, H.-M. (eds.), Macroeconomics and the History of Economic Thought – Festschrift in Honour of Harald Hagemann*, Routledge, London and New York, 351–367.

Rosenstein-Rodan, and N. Paul, 1943. "Problems of Industrialisation of Eastern and South-eastern Europe." *Econ. J.* 53 June-September, 202–11.

Saviotti P. P., 2001. "Variety, growth and demand." *Journal of Evolutionary economics*, Vol.11, 119–142.

Saviotti P. P., and K. Frenken, 2008. "Export variety and the economic performance of countries." *Journal of Evolutionary Economics*, 18, 201–218.

Saviotti P.P., J. S. Metcalfe, 1984. "A Theoretical Approach to the Construction of Technological Output Indicators." *Research Policy*, 13, 141–151.

Saviotti P.P., and A. Pyka, 2004a. "Economic development by the creation of new sectors." *Journal of Evolutionary Economics*, 14, (1), 1–35.

Saviotti P.P., and A. Pyka, 2004b. "Economic development, qualitative change and employment creation." *Structural Change and Economic Dynamics* Vol 15, 265–287.

Saviotti P.P., A. Pyka, and J. Krafft, 2007. *On the determinants and dynamics of the industry life cycle, presented at the EAEPE conference 'Economic growth, development, and institutions – lessons for policy and the need for an evolutionary framework of analysis'*, held in Porto, on November 1–3.

Saviotti P.P., and A. Pyka, 2008. "Micro and macro dynamics: Industry life cycles, inter-sector coordination and aggregate growth. *Journal of Evolutionary Economics*." Vol. 18, pp. 167–182.

Saviotti P.P., and A. Pyka, 2010. *The Co-evolution of Innovation, demand and growth, paper presented at conference 'Technical Change, History, Economics and Policy' to be held in SPRU*, Freeman Centre, University of Sussex, on March 29–30.

Schumpeter J., 1934. original edition 1912. *The Theory of Economic Development*. Cambridge, Mass, Harvard University Press.

Veblen T., 1915. *Imperial Germany and the Industrial Revolution*. MacMillan, New York.

Witt U., 2001. "Learning to consume: a theory of wants and the growth of demand." *Journal of Evolutionary Economics*, Vol. 11.

APPENDIX

Table Appendix

	h_i entry Barrier (B_{hi}) a	Weight of H_i in Production Function(k_{Hi}) b	Wage Function Parameter (k_w) c	
1.	0.0	0.1	1	Standard
2.	0.5	0.1	1	Entry
3.	0.8	0.1	1	Barrier
4.	1.2	0.1	1	Experiments
5.	1.5	0.1	1	
6.	0.0	0.5	1	Production
7.	0.0	1.0	1	Function
8.	0.0	1.5	1	Experiments
9.	0.0	0.1	0.5	Wage
10.	0.0	0.1	1.5	Function
11.	0.0	0.1	2.0	Experiments
12.	0.5	0.5	1	a&b
13.	0.5	1.0	1	
14.	0.5	1.5	1	
15.	0.5	2.0	1	
16.	0.8	0.5	1	
17.	0.8	1.0	1	
18.	0.8	1.5	1	
19.	0.8	2.0	1	
20.	1.2	0.5	1	
21.	1.2	1.0	1	
22.	1.2	1.5	1	
23.	1.2	2.0	1	
24.	1.5	0.5	1	
25.	1.5	1.0	1	
26.	1.5	1.5	1	
27.	1.5	2.0	1	
28.	0.5	0.1	0.1	a&c
29.	0.5	0.1	0.5	
30.	0.5	0.1	1.5	
31.	0.5	0.1	2.0	
32.	0.8	0.1	0.1	
33.	0.8	0.1	0.5	
34.	0.8	0.1	1.5	
35.	0.8	0.1	2.0	
36.	1.2	0.1	0.1	
37.	1.2	0.1	0.5	
38.	1.2	0.1	1.5	
39.	1.2	0.1	2.0	
40.	1.5	0.2	0.1	
41.	1.5	0.1	0.1	
42.	1.5	0.1	0.5	
43.	1.5	0.1	1.5	
44.	1.5	0.1	2.0	
45.	0.5	0.5	0.1	a,b&c
46.	0.8	1.0	0.5	
47.	1.2	1.5	1.5	
48.	1.5	2.0	2.0	

ENVIRONMENTAL TAXES, INEQUALITY AND TECHNICAL CHANGE[1]

Fabrizio Patriarca
Sapienza University of Rome
Francesco Vona
OFCE

Environmental innovations heavily depend on government policies and consumers' behaviour. This paper addresses the issue of how these two factors interact in shaping the transition to a green technology. We extend models of technological selection with heterogeneous agents and learning by including a weak hierarchy between green and polluting goods. For general distributions of agents' income and the explicit inclusion of a carbon tax, the model is not analytically tractable so we derive our results using numerical simulations. Given the level of income, carbon taxes are more effective when technological improvements brought about by wealthy pioneer consumers suffice in inducing the remaining population to buy the green good. In this case, a negative relationship between income inequality and tax effectiveness emerges. Taxes on polluting production have a regressive effect since they are mainly paid by poorer people who consume less of the green good. For these people, a negative wealth effect strongly contrasts the standard substitution effect of the tax. Finally, both lower inequality and taxes have the expected effect for intermediate levels of the learning parameter.

Keywords: Environmental Innovation, Inequality, Demand, Simulation Models.

Concerns for quality of life, sustainability of growth and environmental issues occupy an increasingly important position in the set of citizens' values, especially in developed countries where basic needs have been met (Inglehart 1995). Technical change is at

1. We wish to thank an anonymous referee, Alessandro Sapio and Mauro Napoletano, one of the editors of this issue, for particularly useful comments and suggestions. Usual disclaimer applies.

the centre of the political discourse being the unique way of reconciling current consumption patterns with both natural resources preservation and environmental quality. In the case of environmental technologies, policy interventions and the spontaneous involvement of citizens-consumers are particularly important as market prices do not reward for the lower environmental impact of green goods. Furthermore, two channels are recognized to be the most important drivers of environmental innovations (Beise and Rennings 2005): a direct market demand for green products and an indirect political pressure for the approval of ambitious environmental policies.

On the political side, consumers, local communities and environmental activists play a key role in signalling harmful effects of certain economic activities, in giving political voice to ethical issues (i.e. the rights of future generations) and in reinforcing the effectiveness of government interventions (Esty 1998). For instance, cooperatives and diffused ownership characterize the industry of wind turbines in Denmark (Johnson and Jacobsson 2003), while German environmental activists played a key role in sustaining ambitious feed-in-tariff programs (Lauber and Mez, 2004). On the economic side, agents' consumption choices and willingness-to-pay (WTP, henceforth) higher prices for products with low environmental impacts allow the creation of niche markets for these products. Policies can be targeted to promote the creation of niche markets through labelling, public procurement or regulation, e.g. car sharing requiring low-emission cars (Kemp *et al.* 1998). Another example is the one of the private provision of a public good where certain consumers accept to contribute to the good independently on what the others do (Kotchen 2006). All these examples suggest that the effectiveness of environmental policies depends on the distribution of preferences for environmental quality across heterogeneous agents.

This paper focuses on a particular aspect of the complex relationship between agents' heterogeneity in preferences, technology diffusion and environmental policy. In particular, we analyse how the effect of the policies on green technologies is mediated by the distribution of agents' preferences and by the budget constraints preventing these preferences to be realized, i.e. poverty and financial distress. With this aim in mind, we develop a simple model of technological selection with heterogeneous agents and pioneer consumers generating positive spillovers on the remaining population. These ingredients are common to a wealth of models, both analytical (Matsuyama 2002, Bertola *et al.,* 2006) and computational (Frenken *et al.* 2006, Cantono and Silverberg 2010). We contribute to this literature in two ways. First, heterogeneous attitudes towards green or non-green goods, embodying technologies with different environmental impacts,[2] depend in our model on microfounded agents' behaviour. More precisely, following our previous paper (Vona and Patriarca 2011, PV henceforth), we capture the fact that wealthier households care relatively more about environmental quality by introducing a weak hierarchy between green and non-green goods. This allows us to analyse not only the decision of buying the green good or not, as in related studies (e.g. Cantono and Silverberg 2010), but also the intensity of that choice. As a result, consumers' decisions depend on two income thresholds: a low threshold when a consumer starts buying the green good, a high threshold when she shifts from a mix of goods to full green consumption. Second, we extend PV (2011) in two ways: 1. by examining technological selection for general distributions of income, 2. by looking at the effect of a tax on polluting goods under different levels of income inequality.

These two extensions require a substantial departure from the methodology used in our original paper as it is difficult to preserve analytical tractability with many heterogeneous consumers not uniformily distributed and two thresholds moving endogeneously with technological learning. Besides, numerical simulations help in quantifying the effect of the tax in scenarios characterized by different learning speeds and levels of income inequality. Results of computer simulations not only generalize our previous theoretical and empirical findings on the reversion of the effect of inequality on the diffusion of the green good, but also contribute to explain the heterogeneneity in the effect of environmental policies on technological development (e.g. Vona *et al.,* 2012). In particular, we show that, given the initial level of income per capita, environmental taxes are more effective when positive

2. We assume that each good is produced with only one technology. Green (resp. non-green) goods are produced using a technology with low-(resp. high-) environmental impact.

spillovers from pioneer consumers can effectively trigger a second wave of demand for the green good. Thus, a negative relationship between income inequality and tax effectiveness emerges. In extreme cases, the regressive impact of polices like carbon taxes may have the paradoxical effect of reducing the number of consumers of green products. Our final contribution is methodological as we develop a model that has an analytically tractable core and an extension solved numerically. With respect to closely related percolation models of technology adoption (e.g. Cantono and Silverberg, 2010) and to ABM in general, this extension seems particularly promising as it enables to anchor simulation results to analytical ones.

The paper is organized as follows. Next section reviews the literature on inequality, environmental technologies and consumers' behaviour. In section 2, we remind the model and extends it for general distribution of income, Section 3 presents the main results, while section 4 concludes.

1. Related literature

Understanding the determinants of environmentally friendly behaviour represents the point of departure to entrench theoretical analyses on well-established stylized facts. Two strands of literature analyse the formation of green preferences: at the micro level, the empirical evidence on the determinants of the WTP for green goods; at the macro level, the one on the determinants of environmental regulation.

The macro strand of literature is quite scant given the lack of reliable time-varying data on environmental policies and regulation. Among the few exceptions, Dasgupta *et al.* (2001) and Easty and Porter (2002) show that GDP per capita is positively correlated with two independently built indicators of environmental policies, even when adding proxies of government efficiency and of costs of bureaucracy.[3] Interestingly, too, the index built by Easty and

3. The composite indicator used in Dasgupta et al. (2001) includes both environmental policy and environmental awareness and it is based on a survey conducted by the United Nation. The indicator used in Easty and Porter (2002) attempts to measure environmental regulation and uses data sources from the ESI project and the Global Competitiveness Report. Data sources: http://www.yale.edu/esi/ and http://www.weforum.org/issues/global-competitiveness

Porter tends to be significantly higher in more equal Nordic and central European countries. Recent work of Nicolli and Vona (2012) develops time-varying aggregate indicators of renewable energy policy[4] that are positively correlated with GDP per capita. Moreover, they show that lowering inequality positively affects the policy support for renewable energy, especially in high income countries and using policy indicators built with factor analysis. Inequality has a strong negative impact on public expenditures in green R&D as found by Magnani (2000) and by PV (2011), where the effect of inequality appears even stronger in the longer time span considered. Both in PV (2011) and in Nicolli and Vona (2012) a reversion in the relative effect of income levels and inequality emerge, that is: whereas for rich countries inequality negatively affects public policies and demand for green technologies, per-capita income is paramount in poorer ones. This evidence supports our claim that environmental quality is a good relatively higher in the hierarchical ranking.

At the micro level, several studies have shown that wealthier and more educated households are generally more willing to pay higher prices for green goods (Roe *et al.,* 2001; Wiser, 2007; Diaz-Rainey and Ashtonn 2009),[5] to participate voluntarily to the provision of green public goods (Rose *et al.* 2002, Wiser 2007, Kotchen and Moore 2007, Kotchen 2010) or to effectively buy green goods (OECD 2008, Kahn 1998, Gilg *et al.* 2005). It is worth noticing that the overall impact of richer households on the environmental quality can be either positive or negative as long as richer households consume more. However, their impact on technology through the demand of green goods is certainly positive. Also micro-evidence is consistent with our claim that environmental quality is a good relatively higher in the hierarchical ranking. In particular, sociological studies using value and social surveys show that "[the] concern for quality-of-life issues, such as free of speech, liberty and environmental protection... arise only after individuals

4. Data source: http://www.iea.org/textbase/pm/?mode=re

5. Peer effects in consumption are also found to be very relevant in the Contingent Evaluation of the WTP for clean energy carried out by Wiser (2007). In particular, the expected contribution of the others is found to be significantly correlated with the individual willingness to contribute. This result becomes relevant for the relationship between inequality and environmental quality if, as well known in the literature on peer effect in education, peer effects enter nonlinearly in the utility function.

have met their more basic materialist needs for food, shelter, and safety" (Gelissen, 2007; p. 393, see Diekmann and Franzen 1999, Franzen 2003).

On the theoretical side, the standard way to examine the relationship between income inequality and environmental quality is to look at the political-economy determinants of environmental regulation. Using a median voter argument, Magnani (2000) claims that income inequality and expenditures for environmental R&D can be negatively correlated if richer households prefer more environmental quality than poorer ones. Eriksson and Persson (2003) also derive a partial negative relationship between inequality and pollution in a political economy model where heterogeneous agents decide the optimal level of pollution control under the assumption that wealthier individuals are less affected by pollution.[6] Kempf and Rossignol (2007) study a similar problem but allow for a dynamic trade-off between growth and environmental quality. There, the median voter jointly decides the taxes devoted to finance two public goods: environment and infrastructures, which are conductive to growth. In line with previous studies, if the weight assigned to the "environment" in the utility function is low enough with respect to the one assigned to "consumption," a more unequal society would privilege production rather than the environment. Among the channels that support a negative impact of inequality and social segregation on the environment, recent studies (e.g. Rothman 1998, Roca 2003) claim that rich people are often able to divert the monetary benefits out-of-pollution from the cost of it. For instance, in a model of spatial sorting of agents by skills, Gawande *et al.* (2001) show that hazardous waste sites tend to be located in neighbourhoods with a higher fraction of poorer workers willing to accept higher pollution in exchange of jobs in the polluting sector.

All these models examine settings where technology does not change and hence neglect the role of environmental innovations, especially of those innovations that imply a redesign of the whole production process rather than the mere adoption of end-of-pipe solutions. Environmental innovations can be conviniently distin-

6. As in Magnani (2000), the result hinges upon the fact that, given the average income, a richer median voter can afford both more pollution control and more consumption.

guished between end-of-pipe and cleaner (or integrated) technologies: "Cleaner production reduces resource use and/or pollution at the source by using cleaner products and production methods, whereas end-of-pipe technologies curb pollution emissions by implementing add-on measures" (Frondel *et al.,* 2004, pp.1). The former are true innovations, both from the perspective of reducing net energy and material flows and from the one of economic agents, who have to change their behaviour to adopt these technologies.[7] More in general, a transition from polluting to a cleaner technology is best understood as a complex phenomenon involving changes in many institutional layers and the building of new social constituencies (Kemp 1994). Particularly important is the process through which new technologies acquire social legitimacy and become cost-effective. Overall, these socio-cultural features of new technologies are particularly important for green products that involve radical changes in habits and convey an intrinsic ethical motivation.

Following this argument, it is convenient to think at the dynamic interaction between consumers' behaviour and technological development as a prototypical feature of green technological transitions. Standard growth models are not well-equipped to deal with the path dependency emerging from demand-supply interactions. This weakness of standard models is even more relevant when consumers are heterogeneous and hence the dynamics of demand results from the aggregation of different evolving behaviour (e.g. Faber and Frenken 2009 on this argument in relation to environmental issues).

With regards to the diffusion of green products, agent-based computational models (ABM) have been applied to capture the intrinsic socio-cultural aspects of green technologies by introducing a richer set of assumptions on consumers' behaviour. These models typically analyse diffusion patterns of green goods in complex environment characterized by rich supply-side dynamics (Bleda and Valente 2009, Windrum *et al.* 2009) and using calibrated data to build scenarios for technological transitions

7. Examples of significant behavioural changes are: production of energy from renewable sources involving greater decentralization and self-production; change in the ownership structure to enlenght the durability of certain goods, e.g. cars; recycling and reusing activities; creation of consumers' networks to ensure steady demand to local producers of biofood.

(Schwarz and Ernst 2010). Within the broad class of simulation and ABM models, percolation models (Antonelli 1996, Silverberg and Verspagen 2005, Frenken *et al.* 2008) represent the most parsimonious approach to study technology selection with economies of scale and network externalities (e.g. Geroski 2000) when consumers are heterogeneous. Cantono and Silverberg (2010) apply these models to the case of the diffusion of environmentally friendly goods and analyse the effectiveness of a green subsidy. Diffusion of green consumption is constrained by both the heterogeneity in individual WTPs and the consumers' network structure, which affects the spread of information diffusion across potential adopters. To capture learning, the price of the new good decreases with the number of adopters. In this simple setting, the subsidy is effective only within an internal range of the learning parameter: when learning is too slow consumption does not take off, when it is too fast diffusion takes place anyway.

The logic of our model is related to the one of Cantono and Silverberg (2010) as lead-users and consumers' heterogeneity are also essential. Our paper, however, provides a different microfoundation for agent's adoption decision that depends explicitly on the opportunity cost of giving up consumption of the polluting good,[8] through income, and on the initial price gap between the two goods. As will be clearer in the next section, our model also analyses the relationship between the shape of the distribution of WTP for environmental quality and the final outcome.

2. The model

To analyze the choice between polluting and non-polluting goods, we adopt the framework of PV (2011). The simplest setting to address this issue is to consider two goods and two wants. Both goods satisfy a basic need, like food or shelter. The green good,

8. It is interesting to note that, in sociological surveys measuring the values for environmental quality, result strongly depends on the way the demand is made. In terms of absolute preferences developed and developing countries do not differ much. In turn, the higher propensity to spend in environmental quality of developed countries clearly emerges when the opportunity cost of environmental protection in terms, for instance, of foregone income is explicitly mentioned in the survey questionnaire (Inglehart 1995, Diekmann and Franzen 2003). Therefore, this opportunity cost should be also considered in building theoretical models.

indexed by 2, is more expensive but it also satisfies a non basic want like environment preservation. The green good enables agents to enjoy the same direct utility of the old one plus an additional utility linked to an "eco-friendly" motif. This is a convenient way to model preferences for the environment as it encompasses both the case in which "eco-friendly" goods are of better quality, and the one where they are consumed for "altruistic" reasons or moral obligations (see Eriksson 2004, Conrad 2005, OECD 2008).

The weak hierarchy between the two wants, the second being not necessary in an Inada sense, is essential to derive the particular shape of the Engel curve that, in turn, is crucial to derive our main results. As discussed in previous section, this assumption is also empirically founded.

More in details, we adopt the simplified framework of a utility function $w(.)$ that is continuous and additively separable in the two wants. In particular, $w(.)$ is concave in the basic want and linear in the second one. Thus[9]:

$$w(x_1; \mathrm{x}_2) = u(x_1 + x_2) + x_2\,; \tag{1}$$

where x_i is the quantity of the good i and $u(.)$ is a strictly concave function. Note that each unit of the second good gives a greater utility than each unit of the first one, so the first good is consumed only if the price of x_2 is sufficiently higher than the price of x_1. Now let m be the total income to be allocated between the two goods, δ_p the relative price gap, *i.e.* $p_2 – p_1 / p_1$ that represents a proxy of technological expertise in the production of the two goods. Under the previous hypotheses on the utility function, the first order condition for the internal solutions of this simple constrained optimization problem gives:[10]

$$u'(x_1 + x_2) = \frac{1}{\delta_P}\,; \tag{2}$$

9. Further details are discussed in PV (2011) where we also show that the linearity of the utility function in the second want is not necessary to derive our results. Note that the more general form $w_i(x_1; x_2) = u(x_1 + x_2) + \upsilon_i x_2$ would allow to capture heterogeneity in individual preferences for the environment.

10. The first order condition in Equation (2) states that $\tilde{x}$ corresponds to the level of consumption at which the ratio between the marginal utilities of the basic want $u'(x_1 + x_2)$ and of the other want (1) equals the marginal cost of substitution between the two goods

Considering the properties of u, and defining the function ϕ as the inverse of u', this condition has the following solution:

$$x_1 + x_2 = \phi\left(\frac{1}{\delta_P}\right) = \tilde{x} .; \tag{3}$$

Equation (3) implies that a mixed bundle is chosen only within an income thresholds $m \in (m^-, m^+)$. If income is not enough to buy the quantity $\tilde{x}$ of the cheapest good ($m < m^- = p_1\tilde{x}$) the green good is not consumed. When agents are rich enough to buy a quantity of good 2 ($m > m^+$), they will consume only that good which satisfies also the other want.

Figure 1. Engel's curves for the two goods

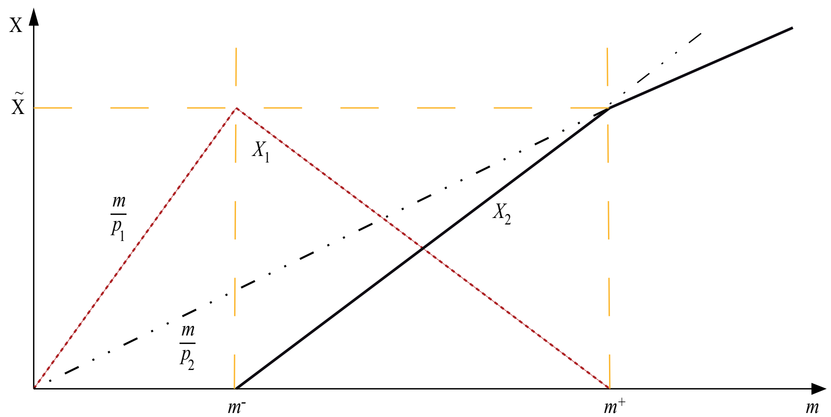


Figure 1 shows the Engel's curve derived by this analysis. The particular S-shape of the Engel's curve of x_2 is the main driver of all further developments. To give a preliminary intuition, it is worth to recall an important property of this curve (see PV 2011 for details). It is steeper in the region (m^-, m^+) than above m^+ Between m^- and m^+, the gradual substitution of the old good with the new one reinforces the positive effect on the consumption of due to the income expansion itself, while in the third region ($m > m^+$) substitution no longer occurs.

In what follows, it is useful to recall that the income thresholds (m^-, m^+) have a "dual" counterpart in the price domain. The "price gap" thresholds are important to analyse technological change in so far as, under standard competitive conditions in all markets, prices reflect costs and the inverse of the price gap reflects technological levels. Moreover, the two price gap thresholds, that correspond to a shift in consumers' behaviour, depend on m. The richer the consumer, the lower the two price gap thresholds required to start consuming the green good. Put it differently, rich consumers buy the green good even if the way of producing it is relatively inefficient.

2.1. The effect of a tax on the first good

In this section, we derive the basic analytical results on the effect of a tax that increases the price of good 1, *i.e.* a carbon tax. The non linearities in the income-demand curves derived in previous section imply that higher income people consume lower shares of the taxed good. Thus, the tax has a regressive impact. For consumers with incomes above m^+, the tax has clearly no effects since they do not consume the taxed good. Conversely, for consumers with incomes below m^+, the tax will have, as usual, the two contrasting substitution and income effects. The strength of each of these effects varies according to agents' shares of consumption, hence according to their income. To analyze the combined impact of these two effects, let us consider the effect of the tax on the two thresholds. A higher p_1 entails a lower δp. Since the marginal utility is a decreasing function (see Equation (2)), it is straightforward to show that this implies a lower $\tilde{x}$. That is, a tax on the old good always decreases the income threshold of consumption at which agents start consuming the new good. In formulas we have.[11]

$$\frac{\partial \tilde{x}}{\partial p_1} = \frac{1}{u''} \frac{p_2}{(p_2 - p_1)^2} < 0. \tag{4}$$

However, the effect of the tax on the income thresholds is not so trivial. Indeed, we have:

11. The continuity and differentiability of all functions considered give:

$$\frac{\partial \tilde{x}}{\partial p_1} = \frac{\partial \tilde{x}}{\partial \delta_P} \frac{\partial \delta_P}{\partial p_1} = \frac{\partial \phi(\frac{1}{\delta_P})}{\partial \delta_P} \frac{\partial \delta_P}{\partial p_1} = -\phi' \frac{1}{\delta_P^2} \frac{\partial \delta_P}{\partial p_1};$$

the definition of ϕ and the properties of the derivative of an inverse function gives Equation (4). .

$$\frac{\partial m^+}{\partial p_1} = p_2 \frac{\partial \tilde{x}}{\partial p_1} < 0. \tag{5}$$

$$\frac{\partial m^-}{\partial p_1} = p_1 \frac{\partial \tilde{x}}{\partial p_1} + \tilde{x} \gtrless 0. \tag{6}$$

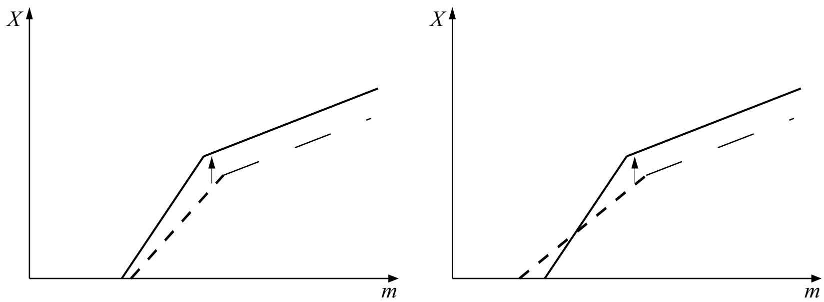
While the upper threshold will always diminish with the tax, thus favoring the demand shift to the green good, the effect on the lower threshold, *i.e.* the minimal income needed to consume the new good, deserves to be further discussed. The first term in Equation (6) represents the substitution effect. It is always negative and, according to Equation (4) is decreasing in the concavity of the utility function and increasing in the price gap. The second term of Equation (6) is the income effect, it is positive and, according to Equation (3), is increasing in the price gap. As a result, for not very concave utility functions and high price gap, the tax may increase the income threshold at which agents start consuming good 2. In this case, some households in the neighborhood of m^- are induced to consume less of the second good because the income effect associated with the higher cost of satisfying the basic need offset the substitution effect associated with a lower price gap. Although this can be seen as an extreme case, it is important to be aware of such possible reversing effects when designing incentive schemes to foster environmental preservation.[12] In the more general cases, the tax increases the demand of the new good of all agents but such an increase is much lower for poorer people that have a stronger negative income effect. Figure 2 summarizes the two possible effects of the tax on the Engel curves.

The case represented on the right panel of Figure 2 allows to visualize the range of middle-low incomes for which the tax has the paradoxical effect of reducing consumption of the green good. The left panel of Figure 2 displays instead the well-behaved case. It is also important to remark that, in both cases, the environmental tax has the standard regressive effect of benefiting wealthier households more than poorer ones (e.g. OECD 2004). In our model, this

12. For instance, this paradoxical result provides a different, simpler rationale to justify the joint adoption of a carbon tax and a subsidy for the green good The standard justification is that the tax is needed for the environmental externality, while the subsidy for the learning or knowledge externality (Jaffe *et al.* 2005). Mix of taxes and subsidies are also commonly observed in practice. The subsidy here can be used to offset the negative income effect of the tax.

result depends on a simple compositional effect. Poorer households consume a large fraction of the dirty good that becomes more expensive. Notice also that subsidizing the consumption of the green good cannot be enough to offset this regressive effect as long as poorer households keep buying a greater fraction of the dirty good. Income transfers are hence required to offset the redistributive effects of environmental policies.

Figure 2. The effect of a tax on the Engel curve



Grounded on these analytical results, the next section address the issue of the effectiveness of environmental taxes on the diffusion of the green good in a context of heterogeneous agents, drawn from a left-skewed distribution of income. Before this, the next sub-section briefly summarizes the relationship between aggregate demand of x_2 and the shape of the income distribution.

2.2. Income Distribution and aggregate demand

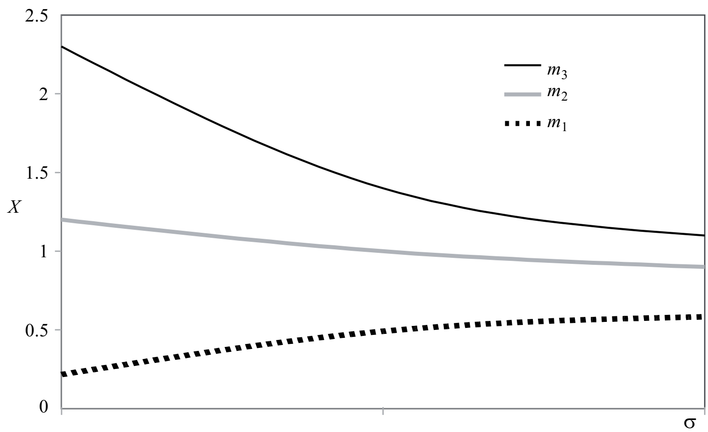
In an economy where agents are heterogeneous in their incomes, the non-linearities in the Engel curves imply that the diffusion of good 2 jointly depends on the average income and on the level of income inequality. With the purpose of giving preliminary insights on this process, let us consider numerical examples drawn from a log-normal distribution of income with a concave shaped utility for the basic want.[13] This is also the distribution of consumers' characteristics chosen by Cantono and Silverberg (2010), which, however, do not analyse the role played by the second moment of the distribution.[14]

13. In particular, we take: $u(x) = ln(x)$ and $\delta p = 2$.

14. Also the functional distribution of income matters on the diffusion pattern of a new good (see Patriarca and Vona 2009).

In Figure 3, we plot the aggregate demand X_2 as a function of the variance in income distribution for different mean income levels ($m_1 < m_2 < m_3$). If the mean income is relatively high, an unequal distribution implies an increase in the number of agents with income under the threshold m^-. For the characteristics of the Engel curve for x_2, that is steeper in ($m^- < m^+$), a redistribution would have in this case a positive impact on the diffusion of the green good. Conversely, in relatively poorer economies, higher income dispersion enables fewer rich people to consume the green good, which can at most emerge as niche consumption.

Figure 3. Variance of the income distribution and diffusion of the green good for different levels of the mean income



This reversal effect of inequality on the diffusion of x_2 is a consequence of the S-shaped feature of the Engel curve of the new good. In turn, the S-shaped relationship depends both on the assumption of a weak hierarchy between the two goods and on the fact that very poor consumers do not buy the green good. It is interesting to note that the S-shaped feature of our Engel curve does not allow to sketch a uniform relationship between inequality and the diffusion of x_2 as it would be for concave—or convex-shaped curves considered in the previous literature on the "aggregation argument" (e.g. Heerink *et al.* 2001). For a standard aggregation argument, if the rich buy relatively more of the green good, higher inequality would generate more consumption of green goods. In our model, instead, middle income households are the only ones for which an income expansion translates in a more than proportional increase of x_2. Hence, there is a reversal in the effect of inequality on the diffusion of the green good for sufficiently high levels of average income.

3. The effect of a tax on the diffusion of green technologies

3.1. The dynamic setting

To analyse the effectiveness of the carbon tax, we consider an environment where technology improves endogenously. As well-documented in the literature on demand-driven innovation (e.g. von Hippel 1988), initially pioneer consumers are willing to buy more expensive innovative products. Their consumption is a source of positive externalities as long as it triggers price reductions that may enable low-budget consumers to adopt these products (pioneer consumer effect, PC henceforth). However, an "excessive income distance" between pioneer consumers and the remaining population prevents the process of diffusion to fully trickle down (consumption polarization effect) to other consumers. The overall effect of the heterogeneity in consumers' characteristics, notably the variance, on the diffusion pattern depends on which of the two effects prevails.

The simplest way of including technological spillovers from pioneer consumers consists in introducing a positive relationship between the growth of demand for X_2 and technological change, *i.e.* a global externality. This assumption is a quite standard way to capture pioneer consumers' spillovers (e.g. Matsuyama 2002, Cantono and Silverberg 2010). Let us denote with γ_i the technological level in sector i, which is equal to the inverse of p_i *i.e.* $p_i = 1/\gamma_i$. We chose a particular linear form for the learning function:

$$\gamma_{2,t} = \gamma_{2,t-1} + c(X_{2,t-1} - X_{2,t-2}), \tag{7}$$

where c measures the effectiveness of technological change. We now analyze the process of diffusion of good 2, by considering the initial condition in which the green good appears at time 1, with a technological level $\gamma_{2,0}$ low enough as to induce a niche demand for this good by few pioneer consumers. Clearly, we also assume a positive technological gap at time 0 *i.e.* $\gamma_{1,0} - \gamma_{2,0} > 0$ being the green technology initially less developed.

Once the niche level of demand emerges, the process of technological progress involves a self-reinforcing process of decreases in p_2 and thus increases in the demand of this new good. For a given mean income, the dynamics of demand depends on the technological parameter c and on the mass of consumers that, in correspondence to each technological improvement, increase their consumption of x_2 the latter being a function of the income distribution. We consider a realistic and general distribution of incomes: the incomes of a population of 1000 agents are extracted from lognormal distributions. The higher complexity of this model with respect to the original paper requires the use of simulation methods.

We compare the diffusion patterns for the new good of two random samples with the same mean (set at a level that allows for a niche consumption of the green good at time 0) but with different variances associated with a Gini coefficient of respectively 0.22 and 0.44, which are the lower and upper bounds of Gini coefficients in OECD countries.[15] The results of this preliminary exercise are shown in Figure 4: the left panel considers a lower level of the learning parameter while the right one a middle level.[16] In the left panel, when technical change is too slow, the PC effect prevails so the unequal society guarantees greater diffusion of the green good. In the right panel, the level of c potentially allows a full diffusion of the new good and the more equal society outperforms the unequal one.

In both cases, technological progress is initially faster in less equal countries because of the stronger PC effect. When technology becomes more mature, however, the more equal population recovers and overcomes the less equal one given the larger number of potential followers, i.e. the middle class is larger. This result, jointly with the empirical evidence presented in the PV (2011), seems to confirm that pioneer consumers play an important role in explaining early stages of technological development, while the mass of potential adopters is more important in later phases. In the

15. See, e.g., OECD on-line statistics: http://stats.oecd.org
16. Respectively, c and $c = 0.2$.

right-hand panel of Figure 4, the crossing-point between the two diffusion curves corresponding to different levels of inequality highlights this leap-frogging effect that—it is worth to remark—occurs only for a sufficiently high learning potential. Finally, Figure 5 replicates the analysis of Figure 4 for a lower level of income per capita. In this case, the green good diffuses less in the equal society also in correspondence to higher levels of the learning parameter. This further exercise generalizes our analytical result on the reversion in the effect of inequality depending on the level of income per capita (see PV 2011).

Figure 4. Dynamics of consumption of the green good

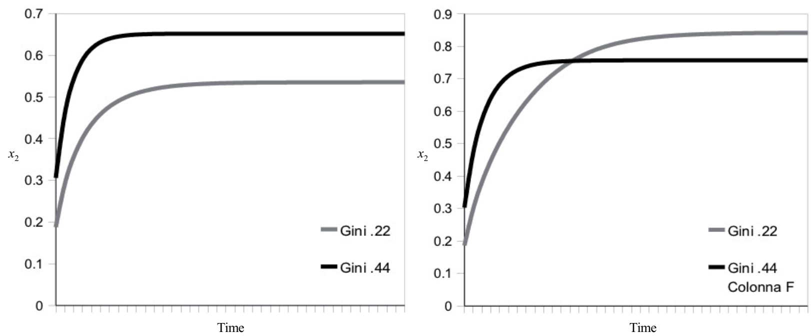
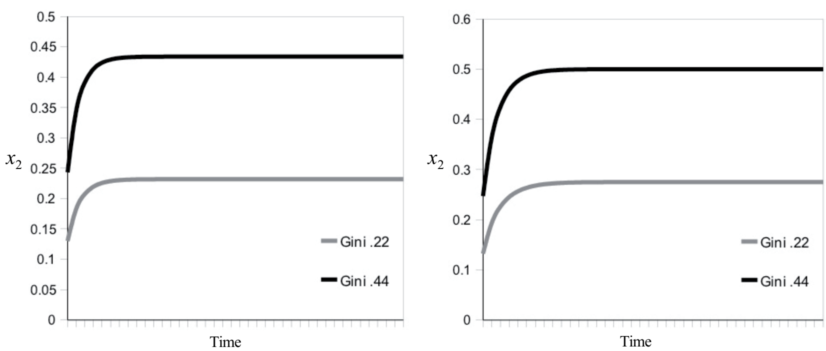


Figure 5. Dynamics of of consumption of the green good (lower mean income)



3.2. The effect of a carbon tax

To examine the effect of a carbon tax, we use a slightly different approach and extract randomly 100 couples of populations of 1000 consumers. Each couple of populations is a random sample from two lognormal distributions having the same mean but two different variances that correspond to the lower and upper bounds of the Gini coefficient in OECD countries, respectively .22 and .44. For every couple we analyze the effect of a 5% carbon tax on the diffusion of the new good. We run the model for four cases: with and without the tax for each of the two population. As we already discussed in the previous section, the dynamics of the system depends on the income distances between agents. In each population, these distances vary although they are all randomly extracted from the same stochastic process. The 100 replications allow to make the results independent from the income distances of the specific population. First, we consider a benchmark case setting the parameter c at an intermediate level (the same as in the left panel of Figure 4), then we will move to the more general case, by varying c in its relevant range.[17]

Table. Final Levels of X_2

		no tax	tax	% change
Population1	mean	844.8	993.4	17.75
(Gini=0.22)	std. dev.	[1.9]	[1.5]	[1.6]
Population2	mean	760.4	839.1	10.31
(Gini=0.44)	std. dev.	[1.8]	[1.5]	[1.8]

Source: Simulated values for 100 couples of populations of 1000 consumers.

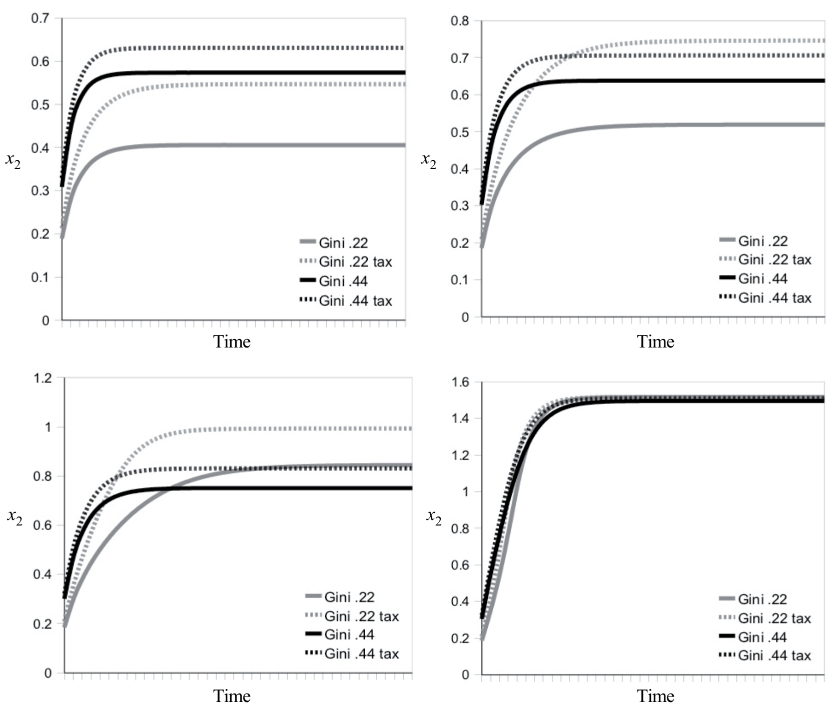
The first two columns in Figure 1 show the average final demand levels of of the new good in the four cases. The third column shows the average relative increase of the final demand level involved by taxation. The result states the higher effectiveness of the tax for the more equal population. In particular, the average improvement is 17.75 for the equal population with respect to an average of improvement of 10.31 for the unequal one. The difference in the effect of the tax is highly statistically

17. The parameters for the benchmark case are: c = 0.2, $p^1 = 1$, the initial level of the price gap is $\delta p = 2$ and the mean income is $\mu = 2$.

significant since standard deviations are very low. Comparing column 1 and 2, it is evident that the tax amplifies the positive effect of lowering inequality on the diffusion of the green good.

The final robustness exercise consists in exploring the effect of the tax for different learning parameters. For the sake of simplicity, we consider two populations, characterized respectively by a Gini equal to .22 and .44 respectively, for which we plot the dynamics of the demand of the new good for different levels of the parameter c ($c_1 < c_2 < c_3 < c_4$[18]). This simple graphical analysis of specific cases allows to draw some insights on the joint role of the parameters of the model. Results are shown in Figure 6, where the third panel is the benchmark in Figure 4.

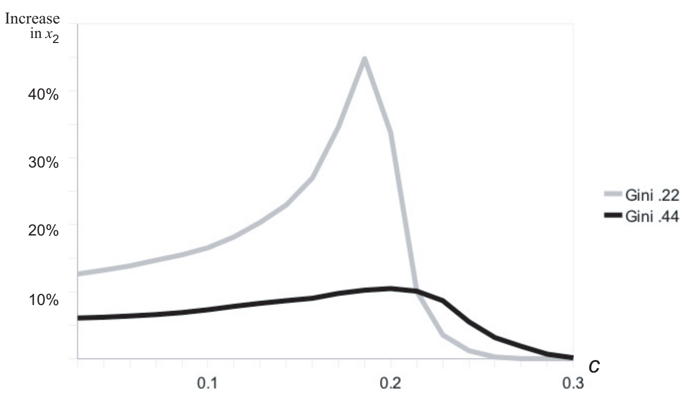
Figure 6. The effect of a tax for different learning parameters



18. We set: $c_1 = 0.16$; $c_2 = 0.18$, $c_3 = 0.20$ and $c_4 = 0.28$.

The positive effect of lowering inequality documented in the previous section is confirmed, except for high levels of the learning parameter (south-east panel of Figure 6). Similarly to Cantono and Silverberg (2010), a very effective technological learning renders useless both the carbon tax and an income redistribution. Panel north-west in the same figure confirms that the unequal society outperforms the equal one for low levels of the learning parameter. However, the tax is more effective in the equal society also in this scenario. The north-east and south-west panels stress the continuity of our main argument: the tax can be very effective when a lower inequality favours the formation of a second wave of demand for the green good, while it is relatively ineffective when inequality is too high. Interestingly, too, the tax allows the more equal system to outperform the less equal one also for relatively low learning parameters (see north-east panel in Figure 7). This important result reinforces our result (see also PV 2011) that environmental policy turns out being significantly more effective in more equal societies.

Figure 7. Tax effectiveness (average relative increase in the diffusion of the new good due to taxation) and learning parameters



To conclude, as a robustness analysis, we take the previous 100 couples of random populations and run the dynamics of the system for 20 consecutive values of c in its relevant range (0, 0.3). As in the table in Figure 1, for each value of c we obtain the average relative increase in the final level of the X_2 due to the tax for the two inequality cases. Figure 7 shows the result of this exercise.

Previous conclusions are strongly confirmed: tax effectiveness is almost everywhere higher for the equal case; the relation is reverted only for higher values of c, that is, when the effectiveness of the tax tends to zero as technological change allows by itself for a wide diffusion of the new good. Furthermore, the tax has a stronger impact for intermediate values of the learning parameter, especially in the equal case.

4. Conclusion

This paper generalizes to the case of realistic distributions of heterogeneous agents our previous theoretical and empirical findings on the reversion of the effect of inequality on the diffusion of green goods (PV 2011). In correspondence to low levels of income per capita, high inequality maximizes the positive effect of early adopters and positively affects diffusion. In turn, the reverse occurs for high levels of income per capita as pioneer consumers can effectively trigger middle class consumption, provided the income distance is sufficiently low.

The second and main result of the paper is to provide a rationale for the heterogeneity in the effect of environmental policies on green technologies (e.g. Vona *et al.* 2012). First, notice that the policy is regressive as it increases the income required to buy the basic good and hence reduces residual income that poor households can devote to the green good. For middle—and low—income households, the tax can bring a negative wealth effect that may overcome the standard substitution effect and, under certain conditions, it leads to the paradoxical result of a reduction in the overall diffusion of the green good.

In the dynamic setting, environmental taxes are generally more effective when, given the level of income, technological externalities from pioneer consumers suffice in inducing less wealthier households to buy the green good, and thus when inequality is lower. The tax can benefit also the middle and low classes only if this negative income effect in the short-term is more than offset by an effective increase in the consumption of the green good in the long-term. Another interesting result, similar to the one of Cantono and Silverberg (2010), is that the tax appears to be significantly more effective when learning is neither too slow nor too fast. When learning is too slow, green consumption remains anyway in a niche. When learning is too fast, the transition does not depend on the level of inequality. Instead, relatively unequal societies with a larger pioneer consumer effect transit fast to the new steady state characterized by fully green consumption.

Finally, as a methodological contribution, our analytically tractable model can represent a useful benchmark for numerical simulations and extensions accounting also for local interactions. Two extensions appear particularly promising. First, in the spirit of ABMs, one could explicitly set local consumers' network to examine how the distribution of income across space shapes consumers' habits and technological development. Second, one could model the green good as an impure public good to investigate how agents' implicit cooperation in the provision of that good is affected by different level of income inequality and different expected gains from new technology.

References

Antonelli, C., 1996. "Localized knowledge, percolation processes and information networks." *Journal of Evolutionary Economics* 6, 281-295.

Beise M., Rennings K., 2005. "Lead markets and regulation: a framework for analyzing the international diffusion of environmental innovations." *Ecological Economics* 52, 5-17.

Bertola, G., Foellmi, R., Zweimüller J., 2006. *Income Distribution in Macroeconomic Models*. Princeton University Press, Princeton and Oxford.

Bleda, M., Valente, M., 2009. "Graded eco-labels: a demand-oriented approach to reduce pollution." *Technological Forecasting and Social Change 76*, 512-524.

Cantono, S., Silverberg, G., 2009. "A percolation model of eco-innovation diffusion: the relationship between diffusion, learning economies and subsidies." *Technological Forecasting and Social Change* 76, 487-496.

Conrad, K., 2005. Price "Competition and Product Differentiation When Consumers Care for the Environment." *Environmental and Resource Economics* 31, 1-19.

Dasgupta, S., Mody, A., Roy, S., Wheeler, D., 2001. "Environmental regulation and development: a cross country empirical analysis." *Oxford Development Studies* 29, 173-87.Diaz-Rainey, I., Ashton, J., 2009. Characteristics of UK consumers' willingness to pay for green energy. Norwich Business School, University of East Anglia, *mimeo*.

Diekmann, A., Franzen, A., 1999. "The wealth of nations and environmental concern." *Environment and Behavior* 31, 540-549.

Easty, D., Porter, M., 2001. *Ranking National Environmental Regulation and Performance: A Leading Indicator of Future Competitiveness?*. In: The Global Competitiveness Report 2001 Porter, M., Sachs, J., (eds.), New York, Oxford University Press.

Easty, D., 1998. "NGOs at the World Trade Organization: Cooperation, Competition, or Exclusion." *Journal of International Economic Law* 1, 123–147.

Eriksson, C., 2004. "Can green consumerism replace environmental regulation? a differentiated-products example." *Resource and Energy economics* 26, 281-293.

Faber, A., Frenken, K., 2009. "Models in evolutionary economics and environmental policy. Towards an evolutionary environmental economics." *Technological Forecasting and Social Change* 76, 462-470.

Frenken, K., Silverberg, G., Valente, M., 2008. "A percolation model of the product lifecycle." *UNU-MERIT Working Paper Series* 073, United Nations University, Maastricht.

Frondel, M., Horbach, J., Rennings, K., 2004. "End-of-Pipe or Cleaner Production? An Empirical Comparison of Environmental Innovation Decisions Across OECD Countries." *ZEW discussion paper* 04-82.

Gawande, K., Berrens, R., Bohara, A.K., 2001. "A consumption-based theory of the environmental Kuznets curve." *Ecological Economics* 37, 101–112.

Gelissen, J., 2007. "Explaining popular support for environmental protection." *Environment and Behavior* 39, 392-415.

Geroski, P. , 2000. "Models of technology diffusion." *Research Policy* 29, 603-625.

Gilg, A., Barr, S., Ford, N., 2005. "Green Consumption or Sustainable Lyfestiles? Identifying the Sustainable Consumer." *Futures* 37, 481-504.

Heerink, N., Mulatu, A., Bulte, E., 2001. "Income Inequality and the Environment: Aggregation Bias in Environmental Kuznets Curves." *Ecological Economics* 38, 359-367.

Inglehart, R., 1995. "Public support for environmental protection: Objective problems and subjective values in 43 societies." PS: *Political Science and Politics* 15, 57-72.

Jaffe, A., Newell, R., Stavins, R., 2005. "A tale of two market failures: Technology and environmental policy." *Ecological Economics* 54, 164-174.

Johnson, A., Jacobsson, S., 2003. "The Emergence of a Growth Industry: A Comparative Analysis of the German, Dutch and Swedish Wind Turbine Industries. in: Metcalfe, S., Cantner, U., (Eds): Change, Transformation and Development." *Physica/Springer*, Heidelberg, 197-228.

Kahn, M., 1998. "A Household Level Environmental Kuznets Curve." *Economics Letters* 59, 269-273.

Kemp, R. , 1994. "Technology and the transition to environmental sustainability: The problem of technological regime shifts." *Futures* 26, 1023-1046.

Kemp, R., Schot, J., Hoogma, R., 1998. "Regime shifts to sustainability through processes of niche formation: the approach of strategic niche management." *Technology Analysis and Strategic Management* 10, 175–195.

Kempf, H., Rossignol, S., 2007. "Is Inequality Harmful For The Environment In A Growing Economy?" *Economics and Politics* 19, 53-71.

Kotchen, M., (2006). "Green Markets and Private Provision of Public Goods." *Journal of Political Economy* 114, 816-845.

Kotchen, M., 2010. "Climate Policy and Voluntary Initiatives: An Evaluation of the Connecticut Clean Energy Communities Program." *NBER Working Papers* 16117.

Kotchen, M., Moore, M., 2007. "Private provision of environmental public goods: Household participation in green-electricity programs." *Journal of Environmental Economics and Management* 53, 1-16.

Lauber V., Mez, L., 2004. "Three Decades of Renewable Electricity Policies in Germany." *Energy & Environment* 15, 599- 623.

Magnani E., 2000. "The Environmental Kuznets Curve, Environmental Protection Policy and Income Distribution." *Ecological Economics* 32, 431-443.

Matsuyama K., 2002. "The Rise of Mass Consumption Societies." *Journal of Political Economy* 110, 1035-70.

Nicolli F., Vona, F., 2012. "The evolution of renewable energy policy in Oecd countries: aggregate indicators and determinants." *Documents de Travail de l'OFCE* 2012-13.

OECD, 1994. *The Distributive Effects of Economic Instruments for Environmental Policy*. Paris, OECD.

OECD, 2008. *Household Behaviour and the Environment: Reviewing the Evidence*, Paris.

Patriarca F., Vona, F., 2009. "Structural Change and the Income Distribution: a post-Keynesian disequilibrium model." *Working paper 5 department of Economics*, University La Sapienza.

Pfaff, A., Chauduri, S., Nye, H., 2004. "Endowments, preferences, technologies and abatement: growth-environment microfoundations. *International Journal of Global Environmental Issues* 4, 209 - 228.

Roca J., 2003. "Do individual preferences explain the Environmental Kuznets Curve?" *Ecological Economics*, 45: 3-10.

Roe B., Teisl, M.F., Levy, A., and Russell, M., 2001. "US consumers' willingness to pay for green electricity." *Energy Policy* 29, 917-925.

Rothman D., 1998. Environmental Kuznets curves real progress or passing the buck?: a case for consumption-based approaches." *Ecological Economics* 25, 177–194.

Schwarz N., Ernst, A., 2009. "Agent-based modeling of the diffusion of environmental innovations: An empirical approach." *Technological Forecasting and Social Change* 76, 497- 511

Silverberg G., Verspagen, B., 2005. "A percolation model of innovation in complex technology spaces." *Journal of Economic Dynamics and Control* 29, 225-244.

von Hippel E., 1988. *The Sources of Innovation*. New York: Oxford University Press.

Vona F., F. Patriarca, 2011. "Income Inequality and the Development of Environmental Technologies." *Ecological Economics* 70, 2201-13.

Vona F., F. Nicolli, L. Nesta, 2012. "Determinants of Renewable Energy Innovation: environmental policies vs. market regulation." *Documents de Travail de l'OFCE* 2012-05.

Windrum, P., T. Ciarli, C. Birchenhall, 2009. "Consumer heterogeneity and the development of environmentally friendly technologies." *Technological Forecasting and Social Change* 76, 533-551.

Wiser, R., 2007. "Using contingent valuation to explore willingness to pay for renewable energy: A comparison of collective and Voluntary Payment Vehicles." *Ecological Economics* 62, 419–432.

HIGH WIND PENETRATION IN AN AGENT-BASED MODEL OF THE ELECTRICITY MARKET

THE CASE OF ITALY[1]

Eric Guerci

Department of Quantitative Methods for Economics, University of Genoa

Alessandro Sapio

Department of Economic Studies, Parthenope University of Naples

In this paper, we build a realistic large-scale agent-based model of the Italian electricity market and run simulations to investigate how a significant increase in wind capacity can affect electricity prices at the national level when the wind resource is geographically concentrated, as in the case of Italy. The simulator implements both cost-based and oligopoly models in which electricity companies learn to bid strategically. We compare a scenario based on the 2010 wind supply and a scenario based on the maximum potential wind capacity as estimated in technical reports. Results confirm the beneficial effect of low-cost renewable energy in reducing average market prices, but simulated power flows in the grid suggest that congestion in the electricity network induced by high wind penetration creates market power opportunities that can offset the price reduction effects.

Keywords: Electricity Market, Wind Power, Agent-Based Modeling.

A number of environmental and security issues in recent years have pushed energy economists and policy makers to analyze the prospects of renewable energy sources. Government programs for the abatement of greenhouse gases (GHG) emissions caused by, among other things, combustion of fossil fuels, have been adopted

1. The authors are grateful to the special issue editors, Jean-Luc Gaffard and Mauro Napoletano, and to an anonymous referee for useful comments and suggestions. REF-E (Economics, Engineering, Energy, Environment) is thankfully acknowledged for sharing information about the electricity industry and for providing their dataset on the Italian electricity production pool.

to face growing worries about anthropogenic climate change. Foreign dependence of most countries from fossil fuels, which are nearing depletion and cause diplomatic clashes, call for the intensified exploitation of locally available energy sources. The high human and economic losses due to nuclear disasters (Three Miles Island, Chernobyl, Fukushima), the fear of nuclear proliferation, and the signs of diminished financial attractiveness of nuclear investments (Bradford, 2012) trigger the quest for alternative, cheap energy sources. Hydroelectricity, wind, solar radiation, biomass, tidal waves, and geothermal heat pumps are renewable as they are replenished over time through natural processes; they are locally available; and their GHG emissions and impact on climate change are negligible.

While the environmental benefits of renewables are not under discussion, the debate is open concerning their economic efficiency (see Joskow (2011) and references therein). Power plants using progressive renewable energy technologies, such as wind power and photovoltaics, are characterized by small efficient scale and low variable costs, but upfront costs are high. Market entry therefore needs to be subsidized. The ensuing burden on public budgets is more acceptable for tax-payers if, besides bringing environmental benefits, renewables push electricity prices down.

The impact on electricity prices of increasing power production from renewables has been examined in a number of papers with a focus on wind power, a highly dynamic renewable source in terms of growth rates of installed capacity.[2] Sensfuβ *et al.* (2008), Sáenz de Miera *et al.* (2008), Twomey and Neuhoff (2010), Green and Vasilakos (2010), Banal-Estañol and Rupérez-Micola (2011b) and Sioshansi (2011) are among the main references. Price reduction effects are commonly found, since greater availability of a power source with negligible marginal costs causes the displacement of high-cost, fossil-fueled power plants. However, faced with the prospects of lower prices, oligopolistic power generating companies (gencos henceforth) may look for market power opportunities. One such opportunity can be given by congestion in the transmission grid caused by increased wind outputs. This is likely to occur

2. Worldwide wind capacity has grown from 74 GW in 2006 to 197 GW in 2010 (World Wind Energy Association report 2010).

when wind production is geographically concentrated and there are significant transmission bottlenecks. High electricity demand and high wind production may jointly cause congestion, thereby magnifying local market power and partly or wholly offsetting price reduction effects.

Italy is an interesting test bed for hypotheses concerning the price effects of increasing wind penetration.[3] Thanks to relatively high wind speeds in its southern regions and islands, Italy has a strong potential for wind power production, and a rather fast growing wind power market thanks to incisive support policies. Wind production is concentrated in Italian regions where gencos possess relatively high market power. In particular, limited competition ensues when transmission lines between different zones of the electricity network are congested. This is far from being a rare event due to significant bottlenecks in the Southern Appennine and between Sicily and Calabria, that emerge because the grid was not designed to accommodate such levels of power generation. Because of such bottlenecks, forced outages of wind turbines have been ordered by the transmission system operator, in the face of threats for reliability and security of supply. This caused the loss of about 10.7% of wind energy in 2009, and 5.6% in 2010 (APER, 2011; Lo Schiavo *et al.*, 2011). Further reasons why wind power is a hot topic for Italy include the potential to stimulate development opportunities for regions whose per-capita incomes are below the national average, and concern that public resources in support to renewables may be appropriated by organized crime, which is particularly strong in those regions.

We are interested in verifying two research hypotheses. First, electricity prices decline as wind penetration increases. Second, growing wind power penetration causes an increase in congestion frequencies. These research questions are investigated by means of an agent-based model depicting an electricity market in which heterogeneous, boundedly rational, and capacity-constrained oligopolistic gencos serve a time-varying price-inelastic demand. Gencos engage in price competition in a uniform price (non discriminatory) auction. Their portfolios include both thermal power

3. Wind penetration is the fraction of electricity demand satisfied by means of wind power production.

plants and wind plants, but wind power is supplied unstrategically. Strategy learning on thermal capacity is modeled by means of genetic algorithms. The market-clearing price is set by the market operator using supply offers, taking into account transmission constraints between zones. A model along these lines is used by Guerci and Fontini (2011) for an assessment of the potential impact of nuclear power in Italy, and by Guerci and Sapio (2011) for a comparison between agent-based and optimizing models of the Italian electricity market. Previous applications of agent-based modeling to the electricity market include Bower and Bunn (2001), Bunn and Oliveira (2003), Sun and Tesfatsion (2007), Rastegar *et al.* (2009), while Weidlich and Veit (2008) and Guerci *et al.* (2010) provide critical surveys. To our knowledge, the impact of wind power on the Italian electricity market has never been studied before; most papers focus on Spain, Germany, and the UK.

In the simulations of the model, we set the parameters on wind power supply, electricity demand levels, and cost coefficients as equal to real-world data from the January 2010 Italian electricity market sessions, collected from various sources (REF-E, GME, GSE, Terna).[4] The simulated market outcomes in the scenario based on January 2010 data are then compared with a "wind potential scenario" in which we assess what would happen if, all else being given, the Italian wind power supply reached its potential, *i.e.* the maximum amount of wind power that could be produced, given the Italian orography and the geographical distribution of wind speeds, pressures, temperatures, and available land. Cost-based scenarios are also simulated for the sake of assessing the extent of markups charged by power generating companies.

Our findings show that electricity prices drop as wind supply reaches its potential, but prices remain well above marginal costs. Sharper drops are observed when demand is low than at times of peaking demand, thereby magnifying volatility. The sensitivity of electricity prices to wind power fluctuations, detected through regressions controlling for power demand, is larger in the wind potential scenario. Looking at the patterns of network congestion,

4. REF-E is a Milan-based research and consulting company specialized in energy and environmental economics. Gestore dei Mercati Energetici (GME) is the Italian electricity market operator. Gestore dei Servizi Energetici (GSE) is an Italian State-owned company in the field of renewables. Terna is the Italian electricity transmission system operator.

we find that high wind penetration comes at the cost of more frequent separation between the southern regions, rich in wind, and the 'windless' northern regions. The ensuing market power opportunities partly offset the price reduction effects of high wind penetration. This is a novel result in the literature.

The value added of agent-based modeling in the analysis of the effects of renewables on market outcomes lies in greater realism of the assumptions on behaviors and market structure than the commonly used alternatives (Cournot, Supply function equilibrium, auction-theoretic models). Beyond that, the gencos' environment is a large scale economic system with complex interactions between competing gencos and between possibly congested zones. In such circumstances, full optimization is impracticable, in the sense that a global optimum, if it exists, may not be found in a reasonable amount of time (Simon, 1978). This would force gencos to engage in search for satisficing solutions, which we model by means of genetic algorithms. Most previous works on the price effects of wind power relied on simple models depicting a small number of symmetric, profit-maximizing companies (e.g. duopolies) and ignored issues of grid congestion. Agent-based modeling allows to deal with a very detailed and realistic model of an electricity system—including the real-world structure of the Italian transmission grid and the true spatial distribution of power generating facilities—in an oligopolistic setting. This comes at the cost of giving up the assumption of perfect rationality, but for the reasons given above, bounded rationality provides a better approximation of individual behaviors even when economic agents are specialized in sophisticated activities, such as bidding in the electricity market.

The paper is structured as follows. Section 1 gives a brief overview of the electricity sector and of wind power production in general and in Italy. In Section 2, we summarize the existing literature on the impact of wind generation on electricity market outcomes. We then outline an agent-based model of the electricity market in Section 3, which also describes the implemented learning algorithm. Section 4 illustrates the simulated scenarios, whereas the results are in Section 5. Conclusions and suggestions for further research are provided in Section 6.

1. Basics of the electricity industry

This section offers basic information on the structure and functioning of the electricity industry. The interested reader can refer to the books by Stoft (2002), Kirschen and Strbac (2004), and Harris (2006) for thorough expositions of electricity economics.

Electricity is a property of certain subatomic particles (e.g. electrons, protons) which couples to electromagnetic fields and causes attractive and repulsive forces between them. Trading electricity amounts to trading the availability to supply electrical energy at given times. As such, electricity is not storable: one can only store the means to generate it—e.g. one can keep reserve capacity, or store water behind a dam. Electricity is produced by gencos through turbines activated by several alternative means, such as: combustion of fossil fuels (oil, coal, natural gas), biomass, or biofuels; the potential energy of water stored behind dams or in reservoirs (hydroelectricity); the kinetic energy of wind; geothermal energy; heating of fuels using sunlight (concentrated solar power); conversion of sunlight using the photoelectric effect (photovoltaics). Once generated, power is injected in sources (linked to the power plants) and withdrawn in sinks (loads), which, together with the transmission line system, constitute an interconnected transmission network or grid, through which electricity flows as an alternating current (AC) and is transported over long distances. Because AC repeatedly changes direction, it is impossible to link specific suppliers to specific users: all power is pooled in the network. This is due to Kirchhoff's law, stating that the sum of the currents entering any node (*i.e.*, any junction of wires) equals the sum of the currents leaving that node. Therefore, equality between demand and supply is a technical necessity. But, due to physical transmission constraints (some interconnecting branches may have small capacity values), there can be congestion in the transmission grid. From an economic viewpoint, this can be stated as if a supply and demand matching mechanism based on a purely economic merit order criterion cannot be implemented because of the implicit rationing determined by the Kirchoff's law. Hence an appropriate mechanism based on Kirchoff's law needs to be implemented in order to correctly account for the electrical power flows among the different areas of the grid.

Power from the grid is withdrawn by distributing companies and large industrial consumers. The former, in turn, supply small commercial users and households by means of low-voltage distribution networks and often are also integrated in the retail segment (e.g. billing and metering services). Final electricity consumers are in most countries allowed to choose among competing retailers, or stick to a regulated contract with a public utility. In the short run, final demand is price-inelastic (Considine, 1999; Halseth, 1999; Earle, 2000). The use of regulated tariffs is widely cited as shielding end users from hourly and daily price fluctuations, hence causing limited short-run sensitivity of electricity demand to prices. Demand responsiveness programs, involving real-time metering of electricity consumption and time-of-use pricing, are being experimented in many countries, with results that are so far below expectations (see Kim and Shcherbakova (2011) and references therein).

In countries where the electricity industry has been liberalized, wholesale trading of electricity takes place in organized markets and over-the-counter. Trading concerns the physical delivery of electricity as well as derivatives (forwards, futures, options) on various horizons. The day-ahead market draws much attention in research on electricity economics. Participation to that market involves the submission of production and consumption plans for the day after. Periodic (uniform or discriminatory) double auctions and bilateral continuous time trading are among the adopted trading setups. Further market sessions (adjustment market, market for reserves, real-time market) are held between the day-ahead session and delivery time. Such sessions allow buyers and sellers to adjust their forward and/or day-ahead positions in light of updated information on demand and plants availability, so that the transmission system operator can insure balancing between withdrawals from and injections into the grid.

1.1. The Italian wholesale electricity market

In Italy, day-ahead wholesale trading of electricity takes place in the Italian Power Exchange (Ipex), run by Gestore dei Mercati Energetici (GME), a State-owned company. The Ipex day-ahead market is a closed, uniform-price (non discriminatory), double auction. Each day, market participants can submit bids concerning each hour of the next day. Demand bids are submitted by large industrial consumers, by independent power providers (who serve final users) and by Acquirente Unico, a State-owned company that takes care of final customers who have not switched to competitive retailers. Supply bids are presented by gencos.

The market operator, GME, has the responsibility of clearing the day-ahead market, for each hour of the following day, based on supply and demand bids received from market participants. Market-clearing occurs through a constrained optimization algorithm with the objective to minimize the total expenditure for electricity. The constraints are given by minimum and maximum capacity constraints for each plant, and by transmission limits between zones. A zone is a subset of the transmission network that groups local unconstrained connections. Zones are defined and updated by the transmission system operator, or TSO (Terna in the case of Italy) based on the structure of the transmission power-flow constraints.[5] Choice variables for GME in such optimization problem are the dispatch quantities of electricity to be generated by each power plant, and electricity prices that remunerate electricity production. If transmission constraints are not binding, day-ahead supply offers are remunerated by the same price, the System Marginal Price (SMP). However, when lines are congested the optimal solution involves the calculation of zonal prices.[6] In all cases, electricity buyers pay a weighted average of zonal prices, called PUN (Prezzo Unico Nazionale, or single national price), with weights equal to zonal demand shares. At the optimum, GME calls into operation power plants in merit order, *i.e.* giving priority to offers for the lowest prices. In the merit order typically renewables come first, followed by coal-fired and gas-fired plants. By the same token, demand bids are ordered in decreasing price order. The day-ahead electricity demand curve, however, is typically very steep, consistent with low short-run demand elasticity to price.[7]

---

5. Zones in the Italian grid are: North, Central North, Central South, South, Sardinia, Sicily, plus some limited production poles, namely Brindisi, Foggia, Monfalcone, Priolo, and Rossano. In limited production poles, transmission capacity is lower than the installed power.
6. Holders of long-term contracts receive the contract price; subsidized plants receive regulated tariffs.

1.2. Wind power in general and in Italy

Italy's rich endowment of renewable energy sources—such as hydroelectricity (from the Alps), sunlight, wind (Southern Appennine, Sicily, Sardinia), geothermal energy (in Larderello, Tuscany)—puts it at the forefront of the battle against global warming. Recent years have witnessed the fast growth in wind power capacity in Italy—from 2123.4 MW in 2006 to 5797 MW in 2010. As a result, Italy in 2010 ranked 6^{th} in the world for installed wind capacity penetration, behind China, USA, Germany, Spain, and India, and 9^{th} in terms of wind capacity per land area (19.2 kW/sqkm), the first being Denmark with 86.6 kW/sqkm (source: World Wind Energy Association report 2010).

Wind turbines produce electricity by exploiting airflows. Because work done by a moving mass is proportional to the square of speed, power generated by a wind turbine goes like the cube of wind speed. The relationship between wind speed and power is tuned by the characteristics of the wind turbine, by air density, and by temperature. Minimum and maximum capacities of wind plants are also defined in terms of wind speed: they start operating when wind speed is about 4 m/s; above 25 m/s the turbine is automatically shut down, and a brake is applied to prevent mechanical damage (Bartolazzi, 2006).

Wind speeds are highly variable across space, altitude, and time. As to spatial heterogeneity, average wind speeds and nameplate capacities in Italian regions are given in Table 1. As it can be noticed, the top Italian regions for wind capacity and output fall into just four zones: Center-South, South, Sicily, and Sardinia.

---

7. We have computed the arc elasticity of national day-ahead electricity demand with respect to price in a neighborhood of the single national price (PUN) for all of the 744 hourly market sessions held in January 2010 (our period of interest in the subsequent analysis). We have chosen a pretty large neighborhood of the PUN, namely [0.5 PUN; 1.5 PUN], not to be too conservative, taking also into account that the real-world demand curves are discrete. The median arc price-elasticity of demand is 0.1080, the mean is 0.1051, and the 95% percentile is 0.1689.
That the price-elasticity of electricity demand in Italy is low may sound surprising in light of the wide program of real-time electricity consumption metering implemented by Enel, the former monopolist, as early as 2005. About 90% of final customers were equipped with smart meters by the end of 2009. This is a key step towards stimulating demand response, together with a time-of-use pricing scheme, that has been in place in Italy even before the creation of the day-ahead market in 2004 (Torriti *et al.*, 2010). However, until December 2011 the difference between the peak and off-peak retail prices was fixed by the energy regulator at 10%, hence it did not adequately reflect intra-day wholesale price fluctuations (Lo Schiavo *et al.*, 2011).

Such spatial concentration reflects differences in average wind speeds.[8]

Table 1. Wind power in Italy in 2010

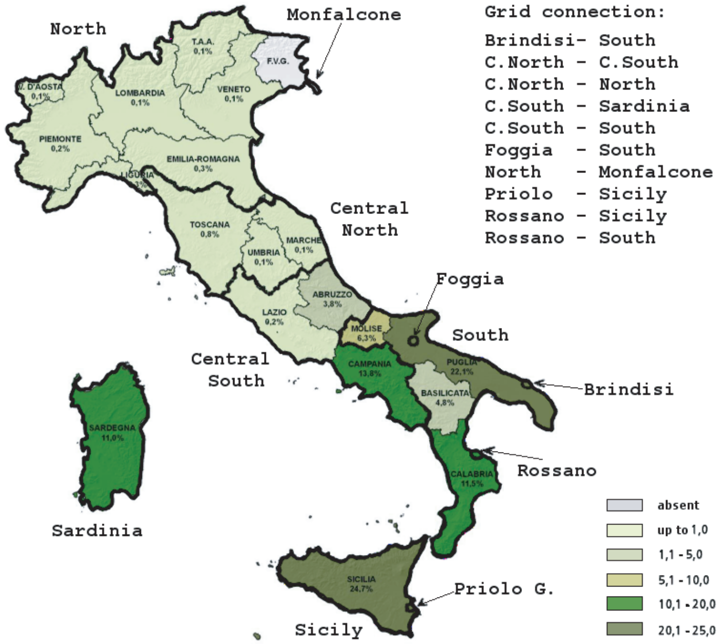
Regions	Zones	n. plants	Inst. capacity	Output
			(MW)	(GWh)
Piedmont	North	7	14.4	21.4
Aosta Valley	North	1	0.0	0.0
Lombardy	North	1	0.0	—
Trentino Alto Adige	North	5	3.1	2.2
Veneto	North	5	1.4	1.7
Friuli Venezia Giulia	North	—	—	—
Liguria	North	15	19.0	34.8
Emilia Romagna	North	15	17.9	24.7
Tuscany	Central-North	17	45.4	76.1
Umbria	Central-North	1	1.5	2.3
Marche	Central-North	3	0.0	0.0
Lazio	Central-South	7	9.0	15.1
Abruzzo	Central-South	25	218.4	329.3
Molise	Central-South	23	367.2	532.3
Campania	Central-South	76	803.3	1333.2
Apulia	South	134	1287.6	2103.2
Basilicata	South	28	279.9	458.3
Calabria	South	31	671.5	952.3
Sicily	Sicily	62	1435.6	2203.0
Sardinia	Sardinia	31	638.9	1036.1

Source : GSE Rapporto Statistico 2010 - Impianti a fonti rinnovabili.

The map of Italy in Figure 1 shows wind penetration rates in Italian regions, in varying degrees of green, as well as the zonal market subdivision adopted by the transmission system operator (bold lines). The transmission lines connecting zones are listed in a grid connection table (top-right). The Italian transmission grid has the shape of a chain that connects the northern, almost windless zones, to the southern zones, rich in wind.

8. See the Italian Wind Atlas: http://atlanteeolico.rse-web.it/viewer.htm.

Figure 1. Map of Italy showing the regional distribution of installed wind capacity (in varying degrees of green) and the borders between zones in the transmission grid (thick black lines)*



*** Zonal connections are reported in the top right corner.**

High variability of wind speeds across time—that is, intermittency in the availability of the wind resource—implies that wind outputs are less controllable than outputs from fossil-fueled power plants, hence they cannot be set strategically in the short run. Large discrepancies can arise between wind supply and wind output if wind power must be offered in advance of delivery.[9] Gencos that cause such imbalances incur costs related to correcting their positions in the adjustment and real-time markets.

Although power plants using renewables are characterized by negligible marginal costs, high upfront costs—the cost of the turbine, foundation, electrical equipment, grid connection—tend to discourage their adoption. Various support schemes have there-

9. In a study on the Nordic countries, Holttinen (2005) mentions 30%-50% wind outputs were forecasted wrong over a time horizon of 7-38 hours ahead of delivery.

fore been enacted around the world. Support to renewables in Italy is channeled by means of feed-in tariffs[10] and green certificates.[11]

2. Literature review: the effects of wind power supply on electricity prices

A growing literature examines the price effects of high wind penetration in electricity markets under different assumptions on market structure. The main result of this literature is that, since wind plants have lower variable costs, fossil fueled power plants are displaced in the merit order, yielding a downward pressure on wholesale electricity prices.

Price reduction effects due to wind power are found by Saenz de Miera *et al.* (2008), who perform an empirical analysis focusing on the Spanish wholesale electricity prices. In their model, perfectly competitive gencos submit linear supply functions in order to maximize expected profits. Expected profits depend on a stable probability distribution of wind generation, and the opportunity costs of CO_2 allowances are internalized. The marginal cost of wind is assumed below that of fossil-fuel generation. A first exercise consists in comparing the market prices in three consecutive days with similar levels of electricity demand in order to isolate the impact of wind generation from the other factors affecting the market price. In a second exercise, the authors simulate the market solution in the absence of wind generation using data for 2005, 2006, and the first 5 months of 2007, and compute the difference between the prices simulated with and without wind generation. The findings indicate a reduction in wholesale electricity prices

10. Since 1992, energy generated through renewable sources is sold to the transmission system operator at a tariff set by the energy market regulator and revised annually. Legislative Decree 78/99 (Art. 12) obliges the transmission system operator to re-sell the subsidized power at prices determined through a merit order. Small plants (below 200 kW for wind plants) that started producing after December 2007 can choose to take advantage of an alternative feed-in tariff (*tariffa onnicomprensiva*). In this case, it is GME who withdraws renewable energy from producers and sells it on the market.

11. Legislative Decree 79/99 (art. 11) stipulates that producers and importers of energy from conventional power plants supply, from 2002 on, a minimal required amount of renewable energy. Such percentage was 3.8%, and has been increased by 0.75% every year, reaching 7.55% for 2012. Obliged parties can meet this requirement either by injecting electricity from renewables in their portfolios, or by purchasing green certificates for an equivalent amount. The green certificates market is managed by GME. Plants using renewables that started operations after April 1999 are assigned a number of green certificates that is proportional to their outputs.

due to wind power. Moreover, such price reduction is found to be greater than the increase in the costs for the consumers arising from feed-in tariffs.

More recent works assuming oligopolistic market structures have demonstrated that price-reduction effects of wind power penetration are exacerbated if thermal power plants are run by gencos with market power. Indeed, when wind is low, gencos with market power face a higher residual demand, therefore they find it optimal to withhold capacity, pushing up the wholesale electricity price. In a supply function equilibrium (SFE) model of the UK electricity market, Green and Vasilakos (2010) show that market power with high wind penetration magnifies price volatility: because strategic gencos view wind sales as a shift in their residual demand functions, increased wind penetration adds to the uncertainty in the residual demand faced by them.

Banal-Estañol and Rupérez-Micola (2011b) analyze an auction-theoretic model in which electricity demand, known with certainty, is served by two symmetric high-cost plants and a low-cost plant. Two cases are analyzed: the low-cost plant can generate intermittent energy (a wind plant) or dispatchable energy (a nuclear plant). In both cases, the model solution features multiple equilibria (with lower prices when high-cost plants are not pivotal, *i.e.* when their capacity is not essential for market-clearing) and mixed-strategy equilibria. The introduction of low-cost capacity—whether intermittent or dispatchable—depresses prices, but wind pushes price down more than nuclear when high-cost gencos are pivotal, and less so when they are not. Moreover, prices remain well above marginal cost even after substantial increases in wind capacity, and wind causes greater volatility. The reason why electricity prices stay above marginal costs is that wind intermittency makes it more difficult for gencos running high-cost plants to coordinate on low-price equilibria.[12] The main insight raised by the authors is that, in the presence of multiple equilibria, intermittency alters the process of equilibrium coordination.

12. The results are robust to several extensions of the analysis (larger strategy spaces, experience-weighted learning, high-cost capacity replacement through low-cost capacity, joint-ownership of low-cost and high-cost plants, risk aversion).

As a downside to price reduction effects, downward pressure on wholesale prices can discourage wind investments, since it amounts to a negative correlation between availability of wind power and electricity prices. Hence, wind plants receive, on average, lower prices than conventional plants—which can be dispatched at any time. Such a negative correlation can be mitigated if wind output is higher in months of high demand—but it is stronger when gencos running thermal power plants possess market power. Twomey and Neuhoff (2010) illustrate this effect in a model with a non-strategic wind generator and conventional Cournot gencos. According to their simulations, the average difference between energy prices of wind and thermal generation can be more than £20/MWh for some parameter settings. The results of the SFE oligopoly simulations of Green and Vasilakos (2010), based on the UK electricity market, confirm this. Sioshansi (2011) builds a Stackelberg oligopoly model wherein a wind generator acts as the leader, and fossil-fueled gencos—the followers—compete in SFE fashion. Using data concerning the ERCOT (Texas) market in 2005, a scenario including additional 10 GW of wind capacity shows that as more wind enters the market, the discrepancy between the average value of overall electricity sales and the average value of wind energy sales grows, depressing the profitability of wind generators.

The models that have been used in the wind power and electricity markets literature are open to critical observations concerning their general validity as well as their usefulness for the Italian case. One advantage of cost-based models is that, in assuming away strategic interaction and learning, they allow the modeler to give a rich description of the transmission grid. One may also argue that opportunities for market power exercise will vanish after a substantial increase in wind power penetration, thereby making marginal cost bidding a realistic behavioral assumption. Yet, in absence of grid investments that will relieve the existing transmission bottlenecks, market power in Italy should hardly disappear, given the spatial concentration of Italian wind. Modeling strategic interaction is therefore needed. Among models with strategic interaction, the Cournot model, as in Twomey and Neuhoff (2010), assumes quantity-setting gencos and a downward-sloping demand function, thereby violating the evidence of price-based competition and inelastic demand. SFE models (Green and Vasilakos, 2010) confine their analyses to symmetric oligopolies and company-wide supply functions. This is useful to avoid convergence problems of the kind highlighted by Baldick *et al.* (2004). Convergence problems are even more severe if, in order to consider transmission constraints, gencos are assumed to optimally set a supply function for each plant, with non-linear dependencies between marginal profits from different plants, and if plant-level marginal cost slopes are small (see Hobbs *et al.*, 2000). In such a case, even slight changes in parameters (e.g. in fuel prices) can push the model out of the convergence region, as mentioned by Sapio *et al.* (2009). Auction theory provides a more realistic framework for our analysis, as in Banal-Estañol and Rupérez-Micola (2011b). However, the authors' game-theoretical results crucially rely on mixed strategy equilibria and on equilibrium coordination between multiple equilibria. A mixed strategy equilibrium can be interpreted as the distribution of pure strategies in a large population of agents, or as plans of action that depend on exogenous, payoff-irrelevant factors (Rubinstein, 1991). Neither interpretation is appealing as a description of supply behaviors in electricity markets. Consider also that the practical application of mixed strategies is further hindered if their implementation is costly (Abreu and Rubinstein, 1988)—and when transmission constraints are involved, such computational costs can be high.

More generally, since the gencos' environment is a large scale system with complex interactions between competing gencos and between possibly congested zones, full optimization is likely impracticable, in the sense that a global optimum, if it exists, may not be found in a reasonable amount of time (Simon, 1978).[13] Following Simon (1972), this is even more true for a large-size combinatorial problem, such as competition in prices between gencos endowed with diversified portfolios. Temporal specificities in electricity market operations (Glachant and Finon, 2000) act as further constraints on computationally-intensive decision processes. Thus, gencos are very likely to rely on bounded rationality. Agent-based models are well suited to represent gencos that engage in search for 'satisficing' solutions (to use Simon's jargon) in

13. The definition of complexity that is relevant for our purposes was given by Simon (1962): a complex system is a system made up of a large number of parts that interact in non-simple ways.

an oligopolistic market with price competition and a highly detailed representation of the transmission grid. Search in agent-based models occurs through inductive strategy selection methods based on learning dynamics, that are shown to possess high predictive power with respect to agents' behaviors (Roth and Erev, 1995; Camerer and Ho, 1999).[14]

3. An agent-based model of the Italian day-ahead market

3.1. The model

Consider a day-ahead electricity market populated by G gencos and a large number of electricity consumers. Company g ($g = 1,\ldots,G$) owns a portfolio of N_g thermal power plants and W_g wind plants. Thermal power plants can use either of F fuels, indexed by $f = 1,\ldots,F$. Power plants inject power in a transmission network connecting Z zones indexed by $z = 1,\ldots,Z$. Demand in each zone is price-inelastic.

Wind plants owned by genco g are placed in different zones, so that $\sum_{z=1}^{Z} W_g^z = W_g$. Wind plant j in zone z can produce at null marginal costs within the feasible production interval $[\underline{Q}_{g,j}^{w}, \overline{Q}_{g,j}^{w}]$. Wind power is offered at zero price and the offered volume is set equal to day-ahead forecasts. Sold quantities are denoted by $\hat{Q}_{g,j}^{w}$. The unitary remuneration consists of feed-in tariff TP.

Each thermal generation unit $i \in \mathcal{I}_g = \{1,\cdots,N_g\}$ is characterized by a feasible production interval defined by lower $\underline{Q}_{g,i}$ and upper $\overline{Q}_{g,i}$ production limits, so that dispatched power $\hat{Q}_{g,i}$ must satisfy $\underline{Q}_{g,i} \le \hat{Q}_{g,i} \le \overline{Q}_{g,i}$ [MWh]. The cost function of the i^{th} thermal power generating unit run by g and using fuel f is given by

$$TC_{g,i}\,(Q_{g,i}) = FP_f\,(a_{g,i} \cdot Q_{g,i} + b_{g,i}) \tag{1}$$

where FP_f [Euro/GJ] is the price of the fuel f and the term within round parentheses corresponds exactly to the amount of fuel in GJ required to generate an energy of $Q_{g,i}$ MWh. The dimensionless coefficient $a_{g,i}$ and the coefficient $b_{g,i}$ [GJ/MWh] refer to the technology-specific efficiency of the power plant and are assumed

14. To be fair, Banal-Estañol and Rupérez-Micola (2011b) perform simulations in which gencos select among equilibria through an Experience-Weighted Attraction mechanism.

time-invariant. In particular, $a_{g,i}$ specifies the relationship between the energy input and output, whereas $b_{g,i}$ denotes the real value of no-load costs that are born only if the plant is dispatched.[15]

The marginal cost $MC_{g,i}$ for g's thermal plant i is constant across the feasible production interval:

$$MC_{g,i}\,(Q_{g,i}) = FP_f \cdot a_{g,i} \tag{2}$$

Let $\mathcal{I}^r_g = \{1, \cdots, N^r_g\} \subset \mathcal{I}_g$ denote the set of all thermal power plants belonging to genco g in zone z using technology f, where $r = (z, f)$. Thus $\cup_r \mathcal{I}^r_g = \mathcal{I}_g$. For each hour of the following day, genco g sets for all plants in r the same markup level $m^r_g \in \mathcal{A}^r_g$.[16] Thus the action space of genco g is $\mathcal{A}_g = \times_r \mathcal{A}^r_g$, that is, the Cartesian product of the markup spaces $\mathcal{A}^r_g$ for the representative plant of g in r.[17]

Each genco g bids to the day-ahead market session of hour h a pair of price and quantity values for each generating unit $i \in \mathcal{I}_g$. A strategy for genco g is defined as

$$s_g(m^1_g, \ldots m^{R_g}_g) = \{(Q_{g,1}, P_{g,1}), \ldots, (Q_{g,N_g}, P_{g,N_g})\}$$

where R_g is the number of genco g 's representative plants. Each pair is defined by a limit price $P_{g,i} = m^r_g \cdot MC_{g,i}$ ([Euro/MWh]) with $i \in \mathcal{I}^r_g \subset \mathcal{I}_g$ and a quantity of power $Q_{g,i} = \overline{Q_{g,i}}$ [MWh], that is, gencos are assumed to bid the maximum capacity of their thermal power plants.

Upon receiving all generators' bids, the market operator clears the market by performing a total welfare maximization subject to equality constraints posed by zonal energy balance (Kirchhoff's law) and inequality constraints, *i.e.* the maximum and minimum

15. The cost structure of a thermal power-plant includes several terms which can be grouped in two distinct components, that is, fixed costs (such as debt and equity obligations associated with plant investments) and operating costs. The latter occur only if production takes place (*i.e.*, if fuel combustion takes place) and are commonly broken down into variable costs, no-load costs, startup and shutdown costs (see Kirschen and Strbac, 2004). In our model, only variable and no-load costs are considered and are both introduced in Equation (1). No-load costs in power engineering refer to quasi-fixed costs. They correspond to the hypothetical cost incurred by a generator if it could be kept running at nearly zero output.

16. This allows to reduce the size of the strategy space. See Müsgens and Neuhoff (2006) for a similar assumption.

17. For instance, the cardinality of the action space of a generation company owning 4 representative plants is $|\mathcal{A}_g| = |\times_r \mathcal{A}^r_g| \approx 10^4$.

capacity of each power plant and inter-zonal directional transmission limits (Kirschen and Strbac, 2004).[18] The dual of this welfare maximization, given perfectly inelastic demand, is the total production cost minimization. The optimal solution consists of a set of zonal prices ZP_z, for $z = 1,..., Z$, and dispatched quantities of electricity $\hat{Q}_{g,i}$ for $g = 1,..., G$ and $i = 1,..., N_g$.[19]

The profit per hour for genco g, Π_g, is obtained as the sum of the profits from representative thermal generating units and the profits from wind power generating units:

$$\Pi_g = \sum_{r=(z,f)} \left[ZP_z \sum_{i=1}^{N_g^r} \hat{Q}_{g,i} - \sum_{i=1}^{N_g^r} TC_{g,i}(\hat{Q}_{g,i}) \right] + TP \cdot \sum_{z=1}^{Z} \sum_{j=1}^{W_g^z} \hat{Q}_{g,j}^{w} \tag{3}$$

3.2. The learning algorithm

How do gencos decide their offers to the day-ahead market? The boundedly rational behavior of gencos is formalized here by assuming that gencos search for 'good enough' or 'satisficing' markup levels by means of a genetic algorithm. The genetic algorithm goes through T runs, indexed by $t = 1,..., T$. In run t, a population of P markup vectors evolves across K_t generations by means of selection, mutation, and crossover operators. Across runs, gencos compute the prices and profits associated to various points in the space of markups, treating the markups of their opponents as fixed. The conjectured markups of their opponents are updated after each run, allowing gencos more exploration.

The learning algorithm can be schematized as follows.

• *Initialization of the simulation*: at the beginning of run 1 each genco g draws uniformly a population of P–1 markup vectors, whose p-th element is ,

$$M_{g,p} = [m_{g,p}^{1},\dots,m_{g,p}^{r},\dots,m_{g,p}^{R_g}] \in \mathcal{A}_g$$

with $m_{g,p}^{r} \geq 1 \quad \forall r$ and p. Markup levels are drawn from the set $\{1.00, 1.04, 1.08, ..., 5.00\}$. Along with the P–1 randomly drawn

18. This optimization problem is known as DC optimal power flow (DCOPF).
19. Zonal prices are the shadow prices of the active power balance constraints in each zone in the minimization problem.

markup vectors, we include the markup vector $M_g = [1.00,...,1.00]$ (*i.e.* all plants bidding at marginal costs).

• *Initialization at each run t*: at the beginning of run t each genco g draws one markup vector M_g with probability

$$\pi_{M_{g,p},t} = \frac{e^{F_{M_{g,p}}/\lambda}}{\sum_{M_{g,p}} e^{F_{M_{g,p}}/\lambda}}$$

(a logit probabilistic choice model) from the population of P markup vectors. $F_{M_{g,p}}$ is the relative frequency of $M_{g,p}$ in the population of P markup vectors, and λ is a parameter that affects the probability of choosing a markup vector. As $\lambda \to 0$, the probability of choosing the markup vector with the highest frequency goes to 1. Only at run 1 we impose that the markup vector being chosen is $M_g = [1.00,...,1.00]$.

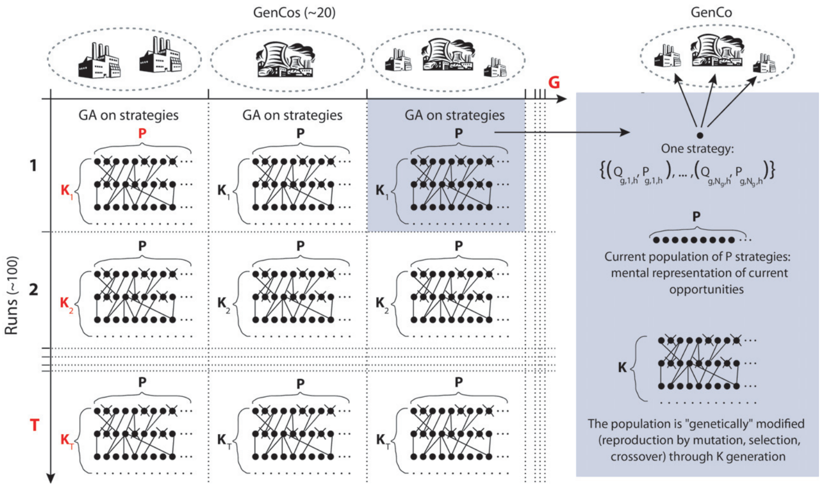
• *At each run t*: at generation $k \in \{1,...,K_t\}$ of run t, for each markup vector of the current population of size P genco g computes the zonal prices ZP_z, $z = 1,...,Z$, and its own profits Π_g under the conjecture that all other gencos play the markups selected at the beginning of run t. Given the current profits/fitnesses for each candidate strategy, a genetic procedure based on selection, mutation, and crossover establishes the next population to be used at $k + 1$ if $k + 1 \leq K_t$ or at run $t + 1$ otherwise.

The learning algorithm is depicted in Figure 2. The left part of the figure shows the evolution of the algorithm through T runs for all G gencos, whereas the right part zooms into the behavior of one genco in one generic run.

In the simulations, we adopt the following parameter settings: $P = 200, T = 50$, $K_1 = 3,..., K_{50} = 20$. Our simulation results are therefore based on the markups selected from the K_{50}th generation. Notice that the number of generations K_t changes across runs. The idea is to favor exploration in initial rounds (smaller K_t) and then to let agents exploit their experience (larger K_t).[20]

20. We assume that gencos learn independently to bid strategically on each hourly market, *i.e.*, no interrelationships are considered among such hourly auctions. The reason is that gencos bid simultaneously on all 24 hourly auctions scheduled for a day.

Figure 2. A schematic representation of the learning algorithm



Notice that search as described by the above learning algorithm need not end up on either a local or a global optimum, even if they exist. The idea here is that gencos set a stopping rule for their search that is independent of the optimality properties of the markups finally selected. The amount of search performed through the genetic algorithm can be thought of being viewed by a genco as satisfactory.

4. Simulation scenarios

By simulating the agent-based model described above, we aim to provide answers to two research questions. First, increasing wind penetration yields price reduction effects. Second, congestion frequencies increase with the rate of wind penetration. These research questions are investigated by building simulation scenarios that differ only for the amount of wind supply: a 'real' scenario, based on January 2010 wind data, and a 'wind potential' scenario, in which we set wind power production to its maximum potential, estimated for Italy in some technical reports. For the sake of comparison, further simulations are run in which all gencos bid their marginal costs. Detailed descriptions of each scenario are provided below.

4.1. Real strategic scenario

In the real scenario, plant-level and demand parameters are measured using January 2010 data from the Italian day-ahead electricity market.[21] We use data from a number of sources. A first source is a database covering most Italian thermal power plants, supplied to us by REF-E, including information on the capacity range (minimum and maximum capacity), technical coefficients of cost functions based on engineering estimates, and transmission constraints. In order to compute costs we also use REF-supplied data on fuel prices and CO_2 prices. We draw hourly data on zonal demands, imports, and the amount of power from renewable sources supplied into the Italian Power Exchange (Ipex) from the website of the market operator GME. Hereby we take into account day-ahead electricity demand after subtracting net imports. A database of Italian wind farms (featuring denomination, technical characteristics, and localization of each plant) is supplied by Terna, the transmission system operator. In the reference period, the available set of power plants consisted of 156 wind power plants and 223 thermal generating units (coal-fired, oil-fired, combined cycle, turbo-gas). Those power plants were independently or jointly owned by 19 different gencos.

We distinguish between core gencos, which behave as oligopolists, and fringe gencos that behave competitively. The competitive fringe includes seven price-taking companies (AA API, AEM, ATEL, Elettrogorizia, EnPlus, Italgen, Set) that own only one power-plant each, and a company, Sorgenia, which owns four generating units all located in the north zone where almost one third of all thermal units are installed. The remaining eleven companies (A2A, AceaElectrabel, EDIPOWER, EDISON, EGL, ENEL Produzione, ENIPOWER, EOn, ERG, IRENMERCATO, TIRRENO Power) behave as oligopolists.

One important issue in implementing our agent-based model concerns the measurement of wind supply. Ideally, one would like to have information on the technical characteristics of individual

21. Focusing on a recent year is an advantage in view of the increasing trend in the Italian wind capacity. In January, the Italian power consumption is at its highest on average, and in 2010 January was the second highest month in terms of wind production, the first being December—but power consumption falls during Christmas festivities (sources: GME and GSE Annual Reports).

wind turbines, as well as wind speeds, pressures, and temperatures at their exact locations. Such information is usually not available. In most papers, it is assumed that all gencos use wind turbines of a standard type—e.g. of given size and height—and wind outputs from such standard wind turbines are computed by plugging into the wind output-wind speed relationship weather data recorded at nearby weather stations. This approach neglects heterogeneity among wind farms and discrepancies in meteorological conditions between the wind farm location and the weather station location.

Our approach is to use the available data on wind power offers. Such offers are ostensibly determined by gencos based on their day-ahead forecasts of wind outputs. Because of intermittency, wind outputs typically deviate from their predicted values, hence blurring the information conveyed by wind offers. Yet, since imbalances imply monetary penalties, players in the wind industry spend resources to refine their forecasts, as testified by Niglio and Scorsoni (2008) in their description of wind forecasting methods employed by GSE. Our wind offers have the advantage of reflecting the individual choices of actual gencos based on weather conditions at the precise locations of wind plants.

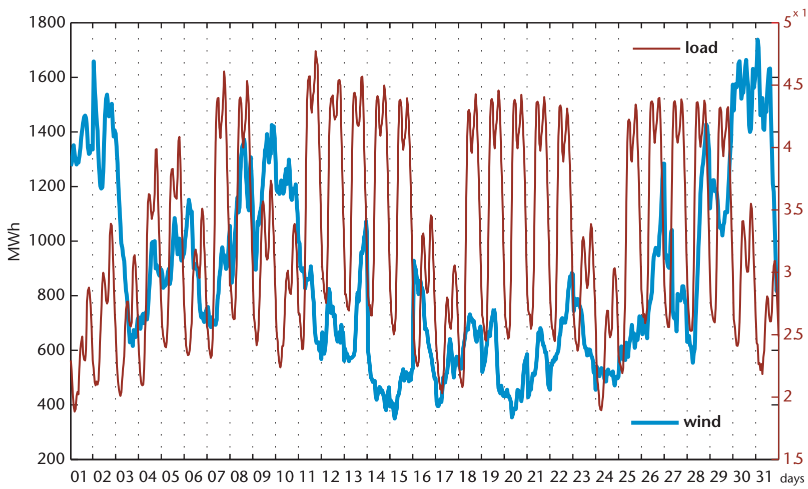
A simple statistical look at our data reveals a number of empirical facts. First, demand bids in the day-ahead market are inelastic in the relevant price range (see also footnote 7). Second, plant-level marginal cost functions are linear, with extremely low slopes. This justifies ex post the assumption of constant plant-level marginal costs of Equation (2). Third, gencos' portfolios are diversified across technologies/fuel types and include plants localized in different Italian regions. Fourth, about 75% of the sell offers consist of a single price-quantity point, even though up to 4 points can be submitted. Finally, intra-day patterns of wind offers are heterogeneous across gencos, due to different locations of the wind plants and different forecasting accuracies.[22]

The time profile of the Italian wind supply and electricity demand in January 2010 is represented in Figure 3, with hourly frequency. This plot shows that, while electricity demand follows quite regular intra-day and weekly patterns—only slightly blurred

22. Wind offer profiles for some gencos are flat across consecutive hours, despite wind intermittency.

by holidays in the first days of the year, wind supply is quite erratic, with fluctuations that suggest stronger winds at the beginning and at the end of January 2010.[23] Interestingly, these plots highlight the presence, in the same month, of market sessions in which the balance between wind supply and electricity demand was very different: low wind with high demand (the Jan 11-15 and Jan 18-22 weekdays), high wind with high demand (Jan 8, Jan 28), high wind with low demand (first and last weekend). This should allow to have a rather complete assessment of the potential effects of wind on electricity prices.

Figure 3. Time series of electricity demand and wind supply, Italy, January 2010



4.2. Wind potential strategic scenario

As a way to detect the price effects of high wind penetration, we build a scenario in which the wind outputs are scaled up, and all the other variables and parameters (the number of gencos, the composition of their thermal portfolios, zonal demands, fuel prices, plant efficiencies, thermal generation capacities, transmission constraints) are kept at the January 2010 level. In particular, we aim to scale up wind outputs until reaching a wind penetration rate similar to that considered by previous papers. Twomey and

23. No entry of new wind plants occurred during January 2010.

Neuhoff (2010) analyzed scenarios with wind covering 15% and 30% of the UK electricity demand. Other studies on the UK (Sinden, 2007; Oswald *et al.*, 2008) considered 20% and 16% scenarios, respectively.

It turns out that similar wind penetration rates can be attained if the January 2010 wind outputs are scaled up to the Italian wind potential, *i.e.* the maximum amount of on-shore wind energy that could in principle be produced, given the Italian orography and the geographical distribution of wind speeds, pressures, temperatures, and available land. Using data supplied by CESI, SPS Italia estimated the Italian wind potential to be about 60 TWh (31 GW), corresponding to about 20% of total electricity consumption in Italy.[24] A study performed by the University of Utrecht gave figures of 69 TWh (34.5 GW). For comparison, notice that wind power production covered about 2.6 % of demand as of January 2010; and that the 1999 Italian White Book targeted to install 12 GW of wind power capacity by 2020. See Ronchi *et al.* (2005) for further details. We therefore simulate our agent-based model as if the wind potential estimated by CESI-SPS was already available in January 2010.[25]

4.3. Cost-based scenarios

In addition to the above strategic scenarios we simulate cost-based scenarios—that is, we run the agent-based model as if all gencos bid their marginal costs. We perform this exercise with both wind supply at its January 2010 levels and at its potential. The outcomes of such cost-based scenarios will be compared with those of the corresponding strategic scenarios, in order to highlight the extent to which markups are eroded by the entry of additional wind power capacity. Notice also that some downward pressure on electricity prices should be expected even if gencos asked their

24. The CESI-SPS study focuses on only locations able to guarantee at least 1750 hour/year of wind power production, and assumes 25 MW of wind capacity per squared Km, and that only 2% of the available land would be occupied by wind farms.

25. Neither the CESI-SPS study, nor the present paper consider the potential for off-shore wind power production in Italy. According to RSE's Wind Atlas, annual average wind speeds of 7-8 m/s are recorded south-west of Sardinia and south of Sicily; 6-7 m/s south-west of Sicily, around Sardinia, and offshore Apulia. RSE (Ricerca sul Sistema Energetico) is a State-owned company performing research on the electricity industry.

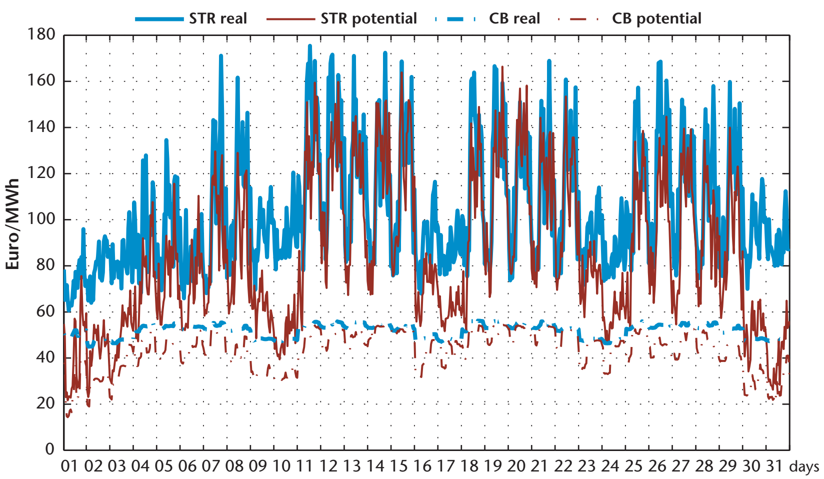
marginal costs, simply because expensive thermal plants are displaced in the merit order.

5. Results

After running the agent-based model, we obtain four simulated hourly-frequency time series of the single national price (PUN)—one for each scenario (real strategic, potential strategic, real cost-based, potential cost-based). In addition, our model yields, as outcomes, the quantity sold by each plant. Together with zonal demands and transmission limits, this allows to determine which lines are congested at each hour, and therefore what zonal configurations appear in the Italian transmission network.

The time dynamics of the single national price is compared across scenarios in Figure 4. The horizontal axis reports each hour of the days indicated. As the picture shows, the day-ahead electricity price fluctuates in a periodic fashion, following the daily and weekly cycles of economic activity. Prices are lower during nights and weekends, and during the first days of January due to holidays—but reach high peaks in the central hours of working days, when electricity demand is at its highest. Prices in strategic scenarios are much higher than cost-based simulation outcomes,

Figure 4. Time dynamics of Italian day-ahead electricity prices, January 2010: baseline cost-based, baseline strategic, wind potential cost-based, and wind potential cost-based scenarios



testifying to market power exercise. The impact of increased wind power capacity can be grasped by comparing the price series for the strategic real and wind potential scenarios. Electricity prices in the wind potential strategic scenario are nearly always below those obtained in the baseline strategic scenario. Such a price reduction effect is particularly strong during periods when demand is lower: weekends, nights, and holidays. A price reduction effect is observed even when gencos bid their marginal costs, as expected. Electricity prices are well above marginal costs even when wind capacity is at its potential. This result is consistent with findings by Banal-Estañol and Rupérez-Micola (2011b).

Wind power affects electricity prices not only across scenarios, but also over time. For each scenario, using hourly-frequency data, we estimate linear regressions of the simulated PUN on wind supply, the real-world national electricity demand, supply from hydropower and other non-wind renewables, and a dummy equal to 1 for all hourly sessions between Jan 1, 2010 and Jan 6, 2010 included and 0 otherwise. Such dummy accounts for exogenous shifts in demand caused by Christmas holidays, that in Italy last until January 6 included. All non-binary variables are in natural logarithms, hence regression coefficients can be interpreted as elasticities. Wind supply is at the January 2010 value in the real scenario, and at the potential value in the wind potential scenario. Estimates are obtained using ordinary least squares as well as a robust estimator, to control for heteroskedasticity and outliers.[26] Estimated coefficients and 95% confidence intervals are reported in Table 2. As indicated by the table, estimated coefficients for demand are positive and statistically significant, with larger magnitudes in strategic scenarios than in cost-based ones, and when wind is at its potential. Closer to our focus, the coefficient estimates for wind supply are all negative and statistically significant—showing that wind production can yield a downward pressure on electricity prices. More in detail, doubling wind supply yields nearly a 4% drop in electricity prices in the real, strategic scenario, a drop that becomes larger (36%) under the wind potential, strategic scenario.

26. Regressions have been performed in Matlab, using the 'regress' and 'robustfit' commands.

Table 2. Estimated coefficients from log-linear regressions of PUN on electricity demand, wind supply, supply from hydropower and other renewables, and holiday dummy: hourly frequency, various scenarios and estimation methods. 95% confidence intervals are reported in square brackets

Variables	Conf .Int.	Real, strat. OLS	Real, strat. Robust	Pot., strat. OLS	Pot., strat. Robust	Real, cost-b. OLS	Real, cost-b. Robust	Pot., cost-b OLS	Pot., cost-b Robust
Constant	95%	-3.397 [-3.840 -2.954]	-3.552 [-4.016 -3.087]	-5.835 [-6.398 -5.271]	-5.968 [-6.542 -5.393]	2.381 [2.269 2.492]	2.210 [2.099 2.321]	0.822 [0.428 1.216]	1.138 [0.813 1.464]
Demand	95%	0.663 [0.533 0.793]	0.664 [0.527 0.801]	2.116 [1.958 2.275]	2.091 [1.930 2.253]	0.501 [0.468 0.534]	0.543 [0.510 0.576]	1.117 [1.007 1.228]	1.012 [0.921 1.104]
Wind supply	95%	-0.039 [-0.065 -0.013]	-0.037 [-0.064 -0.010]	-0.367 [-0.398 -0.335]	-0.361 [-0.393 -0.329]	-0.021 [-0.028 -0.015]	-0.017 [-0.023 -0.011]	-0.345 [-0.367 -0.324]	-0.299 [-0.317 -0.280]
Hydropower & other renewables	95%	0.151 [0.009 0.293]	0.165 [0.016 0.314]	-0.905 [-1.079 -0.732]	-0.870 [-1.046 -0.693]	-0.372 [-0.408 -0.337]	-0.404 [-0.439 -0.368]	-0.599 [-0.720 -0.478]	-0.559 [-0.659 -0.459]
Holidays	95%	-0.097 [-0.125 -0.069]	-0.102 [-0.132 -0.073]	0.024 [-0.010 0.059]	0.019 [-0.0157 0.054]	0.059 [0.052 0.066]	0.066 [0.059 0.073]	0.089 [0.065 0.113]	0.106 [0.087 0.126]

Elasticities of electricity prices to wind supply are around 2% in the real, cost-based scenario, and equal to -0.345 (OLS) or -0.299 (robust fit) in the potential, cost-based scenario. Regression results thus suggest that price reduction effects are much more seizable when wind reaches its potential. Moreover, comparing the coefficients of demand across scenarios suggests that increasing wind supply brings additional volatility to the market.

In Table 3 we report the congestion frequencies for each transmission line, *i.e.* the fraction of hours in which each line was congested. Congestion frequencies are computed for each scenario, and for different groups of days (workweek and weekend) and hours: off-peak (10 pm-6 am) and peak (7 am-9 pm). We observe that, although congestion becomes slightly more rare in the Rossano-Sicily line under the wind potential scenarios, congestion becomes more frequent precisely in lines that connect the zones hosting the bulk of wind capacity with each other or with the other zones: Central North-Central South, Central South-Sardinia, Central South-South, and Priolo-Sicily (except for peak hours in weekdays).[27] Notice that most wind capacity installed in Apulia and Campania is connected through the Central South-South line, which is affected by significant bottlenecks. The Central South-Sardinia line also connects zones with above average wind capacity, while the Central North-Central South line is likely congested because of exports of wind power from the Central South zone to the Central North zone. Congestion frequencies in other lines change only slightly. What we observe, thus, is that increasing wind penetration comes at the cost of increased congestion episodes that effectively separate the 'wind-intensive' southern regions from the 'windless' north, and that cause fragmentation even among southern zones. One may conjecture that such zonal separation creates more opportunities for market power exercise in the southern zones. If so, it might as well be that price reduction effects of wind penetration are partly offset by such stronger market power. The fact that increased wind power production yields more congestion may also be the reason why electricity prices lie above marginal costs even in the wind potential strategic scenario. Recall that Banal-Estañol and Rupérez-Micola (2011b) attributed this to equilibrium coordination attempts by agents. Our simulations provide a different explanation.

Table 3. Market splitting under strategic scenarios

Periods	Scenarios	BR S	CN CS	CN N	CS SA	CS S	FG S	MF N	PR SI	RS SI	RS S
Off-peak (10pm-6am)	**Real, workweek**	0.01	0.24	0.03	0.18	0.12	0.00	0.00	0.04	0.91	0.04
	Potential,workweek	0.00	0.47	0.06	0.52	0.19	0.00	0.00	0.21	0.82	0.01
	Real, weekend	0.01	0.21	0.05	0.07	0.07	0.00	0.00	0.01	0.93	0.04
	Potential, Weekend	0.00	0.62	0.04	0.52	0.20	0.00	0.00	0.47	0.66	0.00
Peak (7am-9pm)	**Real, workweek**	0.03	0.15	0.09	0.15	0.40	0.12	0.00	0.22	0.94	0.10
	Potential,workweek	0.00	0.25	0.11	0.58	0.54	0.04	0.00	0.15	0.79	0.04
	Real,weekend	0.01	0.12	0.06	0.07	0.19	0.01	0.00	0.06	0.92	0.06
	Potential, weekend	0.00	0.45	0.05	0.61	0.50	0.01	0.00	0.21	0.77	0.01

Legend: BR = Brindisi, CN = Central North, CS = Central South, MF = Monfalcone, N = North, PR = Priolo, RS = Rossano, SA = Sardinia, SI = Sicily, S = South.

27. The reader can refer to Fig.1 for a graphical representation of zones in the Italian grid.

6. Conclusion

In this paper, we have built an agent-based model with the purpose of assessing the impact of high wind power penetration on electricity prices in Italy. Our findings show that, as wind supply reaches its potential, electricity prices decrease, although they remain above marginal costs. Wind power fluctuations bring more volatility to the market, as testified by the fact that price falls more sharply when demand is low, thereby magnifying volatility. The elasticity of electricity prices to wind power fluctuations, detected by means of regression methods, is larger in the wind potential scenario. The patterns of network congestions show that high wind penetration implies more frequent separation between the southern regions, rich in wind, and the 'windless' northern regions. Our conjecture is that zonal separation induced by high wind penetration creates market power opportunities that, if exercised, offset the price reduction effects of wind.

The main policy implication of our results is that transmission investments in the southern zones would we worthwhile, since they would bring further electricity price reductions, to the benefit of consumers. Additional investments in the grid would of course put pressure on the public budget. Whether price reductions would be enough to compensate citizens for the additional tax burden is an interesting research question for the future. In particular, one could explore a further scenario relaxing the transmission constraints that separate the northern and the southern zones. This would confirm or falsify our conjecture that, in the case of Italy, price reduction effects of wind are partly offset by increased market power driven by congestion. Our policy implications are in any case in line with the idea that, since the existing power transmission grids had been conceived to support power generation by large centralized plants, large-scale use of renewable energy, fed into the grid by a myriad atomistic producers, will require the transition to smart grids able to 'communicate' with its users.

The foregoing analysis is by no means a complete assessment of costs and benefits from wind power. In performing comparisons between scenarios, a number of variables have been held constant, such as the efficiency and the vintages of power plants. Technological progress may cause obsolescence of the currently available projections on wind potential. Moreover, since intermittency of wind power production gives rise to large discrepancies between programmed and actually produced energy, supplemental energy reserves for balancing the system are required. Such reserve capacities will most likely be supplied by new thermal plants that will have low utilization rates. Generators will therefore be encouraged to install low-cost and low-efficiency plants with greater GHG emissions (Oswald *et al.*, 2008; Strbac *et al.*, 2007).

Future research will also have to take account of endogenous responses of gencos and energy users to the actual and expected impact of high wind power penetration. Following Twomey and Neuhoff (2010), one could build scenarios in which gencos engage in strategic forward trading. Further scenario analyses could be motivated by the likely increase in demand response induced by the diffusion of distributed generation facilities, relaxing the assumption of inelastic demand along the lines indicated by Banal-Estañol and Rupérez-Micola (2011a).

References

Abreu D., A. Rubinstein, 1988. "The structure of Nash equilibrium in repeated games with finite automata." *Econometrica* 56 (6). 1259–1281.

APER, 2011. Rete e vento. *Lo sviluppo della rete elettrica italiana per la connessione e l'integrazione della fonte eolica*. Tech. rep., Italian Association of Wind Producers.

Baldick R., R. Grant and E. Kahn, 2004. "Theory and application of linear supply function equilibrium in electricity markets." *Journal of Regulatory Economics* 25 (2), 143–167.

Banal-Estañol A. and A. Rupérez-Micola, 2011a. "Behavioural simulations in spot electricity markets." *European Journal of Operational Research* 214 (1), 147–159.

Banal-Estañol A. and A. Rupérez-Micola, 2011b. "Production intermittence in spot markets." *Economics working papers*, Department of Economics and Business, Universitat Pompeu Fabra.

Bartolazzi A., 2006. *Le energie rinnovabili*. Hoepli.

Bower J. and D. W. Bunn, 2001. "Experimental analysis of the efficiency of uniform-price versus discriminatory auctions in the England and Wales electricity market." *Journal of Economic Dynamics and Control* 25 (3-4), 561–592.

Bradford P., 2012. "Energy policy: The nuclear landscape." *Nature* 483, 151–152.

Bunn D. W. and F. S. Oliveira, 2003. “Evaluating individual market power in electricity markets via agent-based simulation.” *Annals of Operations Research*, 57–77.

Camerer C., T.-h. Ho, 1999. “Experience-weighted attraction learning in normal form games.” *Econometrica* 67, 827–874.

Considine T. J., 1999. “Economies of scale and asset values in power production.” *The Electricity Journal* 12 (10), 37–42.

Earle R. L., 2000. “Demand elasticity in the California power exchange day-ahead market.” *The Electricity Journal* 13 (8), 59–65.

Glachant J.M. and D. Finon, 2000. “Why do the European Union's electricity industries continue to differ?”. In C. Menard, ed., 2000. *Institutions, Contracts, and Organizations*. London: Edward Elgar, 432–456.

Green R. J. and N. V. Vasilakos, 2010. “Market behaviour with large amounts of intermittent generation.” *Energy Policy* 38 (7), 3211–3220.

GSE, 2010. *Rapporto statistico 2010 - impianti a fonti rinnovabili*. Tech. rep., Gestore dei Servizi Energetici - GSE.

Guerci E. and F. Fontini, 2011. “Phased in nuclear power and electricity prices: the case of Italy.” *Marco Fanno Working Papers* 0130, Dipartimento di Scienze Economiche "Marco Fanno".

Guerci E., M. A. Rastegar and S. Cincotti, 2010. “Agent-based modeling and simulation of competitive wholesale electricity markets.” In *Handbook of Power Systems II*, pages 241–286. Springer-Verlag Berlin Heidelberg, 2010.

Guerci E. and S. Sapio, 2011. “Comparison and empirical validation of optimizing and agent-based models of the Italian electricity market.” In: *Proceedings of EEM11—8th International Conference On The European Energy Market.*

Halseth A., 1999. “Market power in the Nordic electricity market.” *Utilities Policy* 7 (4), 259–268.

Harris C., 2006. “Electricity markets: pricing, structures and economics.” *Wiley finance series*. John Wiley and Sons Inc.

Hobbs B. F., C. B. M. Metzler and J.-S. Pang, 2000. “Strategic gaming analysis for electric power systems: An MPEC approach.” *IEEE Transactions on Power Systems* 15, 638–645.

Holttinen H., 2005. “Optimal electricity market for wind power.” *Energy Policy* 33 (16), 2052–2063.

Joskow P. L., 2011. Comparing the costs of intermittent and dispatchable electricity generating technologies. *American Economic Review* 101 (3), 238–244.

Kim J.-H. and A. Shcherbakova, 2011. “Common failures of demand response”. *Energy* 36 (2), 873–880.

Kirschen D. and G. Strbac (Eds.), 2004. *Fundamentals of Power System Economics*. John Wiley and Sons.

Lo Schiavo L., M. Delfanti, E. Fumagalli and V. Olivieri, 2011. "Changing the Regulation for Regulating the Change: Innovation Driven Regulatory Developments in Italy: Smart Grids, Smart Metering and E-Mobility". *IEFE Working* Paper No. 46

Müsgens F. and K. Neuhoff, 2006. "Modelling dynamic constraints in electricity markets and the costs of uncertain wind output". *Cambridge Working Papers* in Economics 0610, Faculty of Economics, University of Cambridge.

Niglio G. and G. Scorsoni, 2008. "Short-term forecasting of wind power in Italy—the Gestore Servizi Elettrici (GSE) experience." *Proceedings of IFAC Modeling and Control of Economics Systems 2001 Meeting*, Klagenfurt, Austria.

Oswald J., R. Raine and H. Ashraf-Ball, 2008. "Will British weather provide reliable electricity?" *Energy Policy* 36 (8), 3202–3215

Rastegar M. A., E. Guerci and S. Cincotti, 2009. "Agent-based model of the italian wholesale electricity market." In: *Proceedings of EEM09* - 6th International Conference On The European Electricity Market.

Ronchi E., P. degli Espinosa, N. M. Caminiti and G. Onufrio, 2005. "Lo sviluppo dell'eolico in Italia." *Dipartimento Politiche della sostenibilitá* - DS Roma. http://digilander.libero.it/lucixk/dossiereolico.pdf

Roth A. E. and I. Erev, 1995. "Learning in extensive-form games: Experimental data and simple dynamic models in the intermediate term." *Games and Economic Behavior* 8 (1), 164–212.

Rubinstein A., 1991. "Comments on the interpretation of game theory." *Econometrica* 59 (4), 909–924.

Sáenz de Miera G., P. del Rio González and I. Vizcaino, 2008. "Analysing the impact of renewable electricity support schemes on power prices: The case of wind electricity in spain." *Energy Policy* 36 (9), 3345–3359.

Sapio S., V. Canazza, A. Rivoiro, N. Barabaschi and L. Pastore, 2009. "Implementazione in elfo++ del modello supply function equilibrium multi-gruppo." *Quaderno* Ref. n. 55/2009 33, 897–913.

Sensfuß F., M. Ragwitz and M. Genoese, 2008. "The merit-order effect: A detailed analysis of the price effect of renewable electricity generation on spot market prices in germany." *Energy Policy* 36 (8), 3076–3084, August.

Simon H., 1962. "The architecture of complexity." *Proceedings of the American Philosophical Society* 106 (6), 467–482.

Simon H., 1972. "Theories of bounded rationality." In McGuire, C.B., and Radner, R., eds. *Decision and Organization*. North-Holland.

Simon H., 1978. "Rationality as process and as product of thought." *American Economic Review* 68 (2), 1–16.

Sinden G., 2007. "Characteristics of the UK wind resource: Long-term patterns and relationship to electricity demand." *Energy Policy* 35 (1), 112–127.

Sioshansi R., 2011. "Increasing the value of wind with energy storage." *The Energy Journal* 32 (2), 1–30.

Stoft S., 2002. *Power system economics: designing markets for electricity*. IEEE Press. Wiley-Interscience.

Strbac G., A. Shakoor, M. Black, D. Pudjianto and T. Bopp, 2007. "Impact of wind generation on the operation and development of the UK electricity systems." *Electric Power Systems Research* 77 (9), 1214–1227.

Sun J., and L. Tesfatsion, 2007. "Dynamic testing of wholesale power market designs: An open-source agent-based framework." *Computational Economics* 30, 291–327, October.

Torriti J., M. G. Hassan and M. Leach, 2010. "Demand response experience in Europe: Policies, programmes and implementation." *Energy* 35 (4), 1575–1583.

Twomey P. and K. Neuhoff, 2010. "Wind power and market power in competitive markets." *Energy Policy* 38 (7), 3198–3210, July.

Weidlich A. and D. Veit, 2008. "A critical survey of agent-based wholesale electricity market models." *Energy Economics* 30 (4), 1728–1759.

W. W. E. A., 2011. "World wind energy report 2010". *Tech. rep., World Wind Energy Association* – WWEA, April.

Achevé de rédiger en France
Dépôt légal : octobre 2012
Directeur de la Publication : Philippe Weil
Publié par les Éditions du Net SAS 92800 Puteaux

Réalisation, composition : Najette Moummi