

Les modèles de microsimulation dynamique dans l'analyse des réformes des systèmes de retraites : une tentative de bilan

Gaël Dupont*, Cyrille Hagneré† et Vincent Touzé‡

N°2003-10

Décembre 2003

Résumé

Initiés à la fin des années 60 aux Etats-Unis, les modèles de microsimulation dynamique appliquée à l'économie connaissent un essor important depuis une quinzaine d'année, lié à l'accroissement des capacités de calculs et à la disponibilité de données longitudinales. En simulant des trajectoires individuelles, la microsimulation dynamique permet d'analyser l'évolution des systèmes de retraites en se basant sur une hétérogénéité réelle des agents et sur l'ensemble des non-linéarités intrinsèques aux systèmes. Potentiellement, elle peut donc améliorer grandement la compréhension des réformes effectuées ou potentielles, et fournir une mesure précise du montant des droits à la retraite ainsi que des redistributions intra et intergénérationnelles. Elle a en outre la capacité d'intégrer des approches ou des résultats en provenance d'autres types d'analyse, telles que la microéconométrie comportementale.

Cet article propose de dresser un bilan des enseignements des modèles de microsimulation dynamique dans l'analyse des réformes des systèmes de retraite. Dans une première partie, nous dressons une typologie des différentes méthodes possibles dans le domaine de l'analyse des retraites. Dans un second temps, la nature de l'exercice de microsimulation dynamique est précisée ; les différentes sources d'erreurs dans le calcul des trajectoires individuelles et dans l'élaboration de résultats agrégés sont alors étudiées. Cette discussion milite en faveur d'une meilleure connaissance théorique des propriétés statistiques de la microsimulation dynamique, et incite à la création de bases de données répondant directement aux besoins propres de la microsimulation dynamique. Deux exemples élémentaires illustrent certaines propriétés liées à l'emploi de simulations stochastiques. Enfin, nous recensons les principaux modèles existants dans les pays développés, et nous résumons les enseignements issus de l'expérience française à travers le modèle DESTINIE de l'INSEE.

Mots clés : Microsimulation dynamique, études prospectives de long terme, modélisation des système de retraites par répartition.

Code JEL : C15, H55, J11, J26

* Département analyse et prévision – OFCE.

† Département des études – OFCE et THEMA (CNRS).

‡ Département des études – OFCE, THEMA (CNRS) et Institut d'études politiques de Lille.

Cet article est la version révisée d'un texte issu d'un rapport réalisé pour la Commission des finances de l'Assemblée nationale. Nous tenons à la remercier pour son soutien financier. Nous exprimons également toute notre gratitude à Emmanuelle Walraet, pour des discussions très stimulantes qui nous ont permis de profiter de sa connaissance pratique de la microsimulation dynamique, à Alain Jacquot et Ronan Mahieu, nos deux discutants lors de la journée sur les modèles de microsimulation organisée par la CNAF, et aux deux re-lecteurs anonymes de la Revue pour leurs très nombreuses propositions d'amélioration et de correction du texte initial. Bien entendu, les opinions exprimées ou les erreurs et omissions n'engagent que leurs auteurs.

Correspondance à : vincent.touze@ofce.sciences-po.fr.

**Dynamic Microsimulation Models Using to Analyze
Retirement Systems Reforms:
An Essay of Synthesis**

Dynamic microsimulation models applied to economics were first introduced in the United States at the end of the 1960's. Since the 1980's, they have been rapidly developing due to the increase in computational capabilities and to the availability of longitudinal data.

Dynamic microsimulation provides a tool for analysing nonlinear pension systems by simulating individual trajectories of heterogeneous economic agents. Hence, this method highly improves the understanding of pension reforms, the evaluation of pensioners' rights and intra- as well as inter-generational redistributions. Moreover, it takes into account methods and results from other kinds of analyses, such as behavioural micro-econometrics.

This article sets out a synthesis of the achievements of dynamic microsimulation studies of pension systems. The first section draws up a typology of useful methods in the field of pension systems studies. In the second section, a technical assessment of the method is carried out. Various sources of errors in the calculation of individual trajectories and in the determination of aggregated results are examined. Two simple numerical examples illustrate some properties of stochastic simulations. Our analysis recommends better knowledge of the statistical properties of dynamic microsimulation. It favours setting up databases which directly answer the needs of dynamic microsimulation. Finally, the third section reviews an extensive range of dynamic microsimulation models from across OECD countries, and sums up the achievements of the French INSEE model, "DESTINIE".

Keywords: Dynamic Microsimulation, Long Run Prospective Studies, Unfunded Retirement Systems Modelling.

JEL code: C15, H55, J11, J26

Introduction

Les évaluations de l'avenir des retraites sont souvent résumées à des projections du taux de dépendance, dont le principal mérite est de révéler, à législation sociale inchangée, les déséquilibres financiers majeurs potentiellement induits par l'allongement de la durée de la vie et l'accélération du mouvement liée à l'entrée massive en retraite des générations du baby-boom. Le débat sur les retraites est ainsi ramené à une simple règle d'équilibre budgétaire où la réflexion consiste principalement à examiner deux positions extrêmes – le maintien du taux de cotisation retraite ou le maintien du taux de remplacement sur le salaire brut – pour finalement recommander une position intermédiaire qui est celle de la parité entre les revenus des retraités et ceux des salariés¹ ; ce qui signifie que le taux de remplacement sur les salaires nets doit être maintenu constant (*cf.* rapport Cottave, 1991 ; Dupont et Sterdyniak, 2000).

Malgré des possibilités d'amélioration, cette approche atteint vite ses limites si on souhaite mesurer avec une certaine précision le montant des droits à la retraite et si on s'intéresse aux conséquences redistributives des réformes. En effet, pour des raisons évidentes d'hétérogénéité sociale et du caractère complexe² des formules de calcul du montant des retraites, il apparaît que les incidences redistributives des réformes, au sein de chaque cohorte, peuvent être importantes et que leurs impacts sur le montant total des droits à la retraite peuvent être difficiles à mesurer. Une approche basée sur des évaluations à un niveau individuel s'avère ainsi indispensable pour répondre à toute une série de questions pour lesquelles le législateur aimerait disposer de sérieux éclaircissements :

- i) Quelle solidarité entre les générations le système de retraite exprime-t-il ? Quelle règle d'équité retenir pour la mesurer (*cf.* Fleurbaey et Michel, 1992) ? Par exemple, en matière d'inégalités entre les générations, l'histoire française récente révèle des évolutions disparates (*cf.* Legris et Lollivier, 1995 ; Chauvel, 1998).
- ii) Quelle équité au sein d'une même génération ? Existe-t-il des inégalités de traitement selon les secteurs professionnels d'appartenance (durée d'activité, accès à la capitalisation, différences objectives d'espérances de vie, etc) ? Le caractère peu universel des retraites en France se traduit, *a priori*, par de profondes différences selon les secteurs d'activités ou selon le sexe (*cf.* Lagarde *et alii*, 1982).
- iii) Dans quelles mesures les réformes sont-elles efficaces puisqu'elles influencent également les comportements, notamment les choix de cessation d'activité ?
- iv) Quel est le soutien politique ? Dans la mesure où les réformes sont susceptibles de traiter inégalement les ménages, il peut être important de disposer d'une évaluation du crédit électoral des réformes proposées.

Plus généralement, au regard du caractère complexe des règles de calcul des pensions retraite, la réalisation d'une mesure pertinente de l'incidence redistributive et de l'évolution

¹ Cette évolution repose sur un choix particulier d'élasticité d'ajustement des paramètres institutionnels au taux de dépendance (*Cf.* Le Cacheux et Touzé, 2003).

² Présence de non-linéarités, besoin de connaître les carrières complètes, prises en compte d'informations qualitatives, etc.

des agrégats de finances publiques, suite à un changement de la législation, nécessite de disposer d'un échantillon représentatif de la population. Si l'information relative aux individus de l'échantillon est assez complète, il est alors possible de simuler pour chacun d'eux des variables économiques pertinentes telles que le montant des retraites. Profitant de l'accroissement des capacités de calcul et du développement des outils statistiques, cette technique, dite de *microsimulation*, fournit aux décideurs publics des réponses de plus en plus précises, notamment en introduisant certains comportements économiques. Toutefois, dans le domaine des retraites, la tâche est compliquée par le caractère *dynamique* de la simulation qui découle de l'aspect temporel de la problématique.

Cet article tente d'établir un bilan de l'approche par microsimulation dans l'analyse des réformes des systèmes de retraite. Les motivations de ce travail sont multiples. Il y a d'abord des interrogations sur les enseignements, que l'on peut tirer des modèles de microsimulation, dans le domaine du financement des retraites. En quoi sont-ils utiles pour comprendre les réformes sociales ? De quelle manière se distinguent-ils des autres travaux de modélisation (équilibre général, macro-projection, ...) ? Il y a ensuite des interrogations d'ordre méthodologique. Comment fonctionnent-ils ? Quelles sont leurs propriétés ? Quel est leur degré de fiabilité ? Nous souhaitons ainsi comprendre en quoi les expériences réalisées à l'aide de modèles de microsimulation dynamique sont susceptibles d'apporter des enseignements nouveaux dans le cadre de la réforme des systèmes de retraite.

L'article s'organise autour de trois parties. La première partie tente de situer la microsimulation par rapport aux autres types de modélisation. Il s'agit de dessiner une typologie au regard des deux axes d'analyse complémentaire que constituent la dimension microéconomique et le bouclage macroéconomique. Dans la seconde partie, la microsimulation dynamique est présentée. Cette étape permet d'identifier les informations utilisées et de préciser la nature technique de l'exercice. Il s'agit également d'étudier la fiabilité des résultats. Dans la dernière partie, nous recensons les principaux modèles et nous nous attardons plus spécifiquement sur l'expérience française. En conclusion, nous fournissons quelques éléments de mise en perspective. En annexe de l'article, deux exemples illustratifs sont proposés pour saisir plus particulièrement les difficultés que peut poser la simulation stochastique pour construire des agrégats.

1. La modélisation dans l'analyse des réformes des systèmes de retraite : une typologie

Il existe six catégories d'outils d'analyse permettant d'appréhender les problématiques de retraite. Ces approches peuvent être classées selon deux axes fondamentaux que sont la dimension microéconomique et la dimension macroéconomique :

- *L'existence d'une approche microéconomique* signifie que l'unité de référence de modélisation est l'individu et que l'hétérogénéité³ des ménages peut être appréhendée à différents degrés (âge, revenus sur le cycle de vie, structure familiale, durée de vie, régimes sociaux d'appartenance, etc). Ces individus peuvent réellement exister (par exemple, ils sont observés dans le cadre d'une enquête ou d'une banque de données administratives) ou être introduits de manière fictive ; leur « invention » doit alors répondre à un impératif particulier, comme celui d'identifier des structures de ménages (des cas-types) qui soient représentatives de groupes sociaux réellement existants, ou un besoin de créer de façon « intelligente » un ensemble de ménages dont les propriétés en termes de distributions des caractéristiques socio-économiques sont compatibles avec des distributions réellement observées par ailleurs (par exemple, création de cohortes ou de populations fictives). La modélisation de l'histoire des cycles de vie de chacun de ces individus peut reposer sur un fondement microéconomique explicite des comportements qui traduit des arbitrages économiques de type coût - bénéfice dans une dimension globalement cohérente –en principe, les comportements de travail, de consommation et d'épargne devraient résulter du même cadre global d'arbitrage sur l'ensemble du cycle de vie comme dans une formulation théorique d'économie pure– ou résulter d'une approche plus *ad hoc* (modélisation qui insiste d'abord sur la pertinence statistique sans véritable identification des arbitrages) ou sur une prédétermination complète des destins (par exemple, *cf. infra*, dans les modèles de cas-types dynamiques : un individu qui resterait au SMIC toute sa vie).
- *L'existence d'une approche macroéconomique* traduit le fait que l'analyse incorpore un bouclage macroéconomique complet, c'est-à-dire qu'au niveau de l'économie générale, on observe des forces de rappel vers un équilibre entre les différents agrégats que sont les offres et les demandes de travail, de capital ou de biens et services par l'intermédiaire d'ajustement du système de prix. Le bouclage signifie aussi le fait que l'Etat garantit un certain équilibre intertemporel en matière de finances publiques, et notamment de financement des retraites. Des règles d'ajustement entre les paramètres qui fixent le niveau des pensions et le taux de cotisation retraite de façon à préserver l'équilibre budgétaire, sont alors explicitement prises en considération. Dans notre typologie, on considère un modèle comme véritablement

³ Le degré d'hétérogénéité dépend de la variété d'individus considérés. Elle se situe à un niveau transversal (hétérogénéité instantanée entre individus) et à un niveau longitudinal (différences de destins). Elle se situe également à un niveau *ex ante* (les individus naissent différents) ou *ex post* (les individus naissent égaux et sont soumis au même processus aléatoire)

macroéconomique lorsque l'ensemble des bouclages économiques est pris en considération.

Une modélisation idéale se devrait d'incorporer les approches microéconomiques et macroéconomiques dans leurs versions les plus détaillées et les plus pertinentes du point de vue des besoins de l'analyse économique (*cf.* questions énoncées dans l'introduction). Mais, pour des raisons évidentes de limites de l'observation et du caractère complexe de la psychologie des ménages et du système productif, il n'existe pas à ce jour d'outil de modélisation sans défaut⁴. En fait, les différents instruments recensés méritent tous d'être utilisés⁵ car chacun est en mesure d'apporter des réponses à des questions différentes ou d'offrir des visions complémentaires à des questions communes. Outre les méthodes de microsimulation, il existe les cinq méthodes suivantes :

- Les projections de tendances permettent de prendre en compte simplement les implications des évolutions démographiques sur le niveau de vie relatif des actifs et des retraités et sur les cotisations de retraite dans un cadre macroéconomique exogène.
- L'approche par cas-types consiste à évaluer les droits à retraites des individus selon leur profil de carrière. Elle permet de comprendre l'effet des règles de calculs des pensions sur un plan purement individuel.
- La modélisation microéconométrique des comportements cherche à valider sur le plan empirique des modèles de comportement en supposant les individus rationnels. Elle ne permet pas d'effectuer une analyse globale des systèmes de retraite car l'analyse est en général trop partielle, mais elle appréhende les comportements (par exemple le choix du départ en retraite) et leurs évolutions induites par des modifications de la législation.
- Les modèles d'équilibre général calculable à générations imbriquées (MEGCGI) sont des maquettes bouclées d'équilibre général des marchés avec ajustement des prix et fondements microéconomiques explicites tant en matière d'offre que de demande. Ils ont une cohérence théorique interne forte.
- Les modèles macroéconométriques sont des modèles néo-keynésiens où les prix ne s'ajustent qu'avec retard. Ils sont en général sollicités pour les analyses de court et moyen termes. Leur potentiel d'application aux retraites est *a priori* plutôt réduit.

Les modèles macro-économétriques et les MEGCGI prennent en compte un bouclage macroéconomique. Les grandeurs économiques agrégées ont une influence sur d'autres grandeurs avec des effets de second ordre. Dans le cas des MEGCGI, ils reproduisent en général des équilibres de plein emploi et ont une portée de long terme, tandis que les modèles macro-économétriques ont essentiellement une portée à court ou moyen terme et permettent, par exemple, d'analyser les phases de transition démographique en termes de déséquilibres

⁴ La grande difficulté pour calculer le montant total des droits à la retraite plaide, d'une certaine manière, en faveur d'une évaluation en grandeur nature, et donc continue, à partir d'une comptabilité individuelle.

⁵ Les projections annuelles à 75 ans et les différents autres travaux de la *Social Security Administration* aux Etats-Unis sont un exemple d'utilisation combinée de différents outils.

sur les marchés et notamment celui du travail. Ces modèles permettent d'étudier l'impact des évolutions démographiques et des réformes des systèmes de retraites sur l'ensemble de l'économie (taux de chômage, taux d'épargne...), ce qui n'est pas possible avec les approches sans bouclage macroéconomique.

Parmi les approches sans bouclage, on distingue les modèles où seule la dimension individuelle est étudiée (cas-types, analyse de comportements microéconomiques) des modèles dans lesquels les grandeurs macroéconomiques sont exogènes. En ce qui concerne les projections de tendance, les variables démographiques et économiques telles que la natalité, la mortalité, l'immigration, le taux de chômage, le taux de productivité sont exogènes. Cette toile de fond étant posée, on étudie les conséquences du vieillissement sur la répartition d'un revenu national exogène. Un type de comportement macroéconomique est parfois pris en compte dans ces projections : l'offre de travail *via* l'effet de flexion (Cf. Chauvin et Plane, 2001).

Sur le plan microéconomique, on distingue les approches qui modélisent les comportements comme le résultat d'une maximisation sous contrainte de celles dans lesquelles les choix ne sont pas explicitement fondés (formulation *ad hoc*). L'approche par cas-types et les projections de tendance se bornent à postuler des profils de revenus, sans étudier, théoriquement ou empiriquement, si ces revenus influencent les choix des agents économiques. Quant aux modèles avec bouclage macroéconomique, ils prennent nécessairement en compte des comportements puisque le bouclage repose sur ces derniers. Dans les MEGCGI, la cohérence théorique prédomine puisque les comportements sont fondés dans un même cadre d'analyse microéconomique, et la principale unité de modélisation est un individu représentatif (et donc fictif) d'une cohorte. En revanche, le contenu empirique est faible (méthode de calibrage sur une période de référence). En comparaison, dans les modèles macroéconométriques, la conformité des équations de comportement aux données observées est privilégiée mais les comportements s'observent seulement à un niveau agrégé, et ces derniers ne sont pas nécessairement fondés d'un point de vue microéconomique.. Quant aux modèles microéconométriques, ils s'intéressent aux comportements.

Parmi ces cinq types de modélisation, la microsimulation dynamique est parente aux cas-types et aux modèles microéconométriques, la maîtrise de ces deux outils étant souvent une phase préliminaire pour construire le modèle de microsimulation dynamique⁶. Ce dernier est un outil de simulation de trajectoires individuelles. En ce sens, il permet de constituer des histoires économiques complètes d'individus représentatifs de la population. Par essence, le niveau de modélisation est microéconomique, mais les équations qui expliquent les destins individuels sont en général *ad hoc*, et donc souvent sans véritable fondement théorique : il s'agit, en effet, la plupart du temps de capter des relations d'inférence économiquement justifiables et avec de bonnes propriétés statistiques. En outre, il n'y a pas de bouclage macroéconomique⁷, notamment pour la simple raison que la sphère productive n'est pas modélisée.

Le tableau ci-dessous récapitule les différentes configurations possibles selon qu'il existe ou non un bouclage macroéconomique et une base d'analyse microéconomique, avec ou sans une « véritable » modélisation des comportements microéconomiques, et selon la

⁶ Sur ce plan, l'élaboration d'un modèle de microsimulation statique repose sur une organisation similaire.

⁷ L'ajustement est supposé automatique : on connaît l'évolution du taux de chômage et la croissance de la productivité ; cette évolution est indépendante de la réforme du système de retraite car il n'y a pas d'effets retour via la consommation des ménages ainsi que la durée d'activité qui joue sur la taille de la population active.

nature des unités individuelles modélisées (individus réels ou fictifs). Les six modèles considérés ont été numérotés de (1) à (6).

Avec cette typologie, il existe dix configurations et quatre d'entre elles ne se voient associées aucun modèle. Elles ont été numérotées de (a) à (d). Ce résultat provient du fait que dans notre choix de typologie, nous nous restreignons à des formes de modélisation dans leurs configurations élémentaires. Des formes plus élaborées de modèles peuvent conduire à des mutations de statut des modèles (versions hybrides), ce qui pourrait légitimer la présence de certains modèles dans plusieurs cases. En particulier, une telle discussion pourrait devenir pertinente avec les différents développements possibles de la microsimulation (introduction de bouclages macroéconomiques, modèles véritablement comportementaux, etc). Ce tableau conduit, probablement, à des imprécisions, mais il a le mérite de résumer les principales possibilités offertes comme choix de modélisation et de proposer une localisation de chacun des six outils existants de façon relativement simple. Chacun des six modèles est présenté plus en détail ci-après.

TYPOLOGIE DES METHODES D'ANALYSE DES REFORMES DES SYSTEMES DE RETRAITE

			Bouclage macro : Non	Bouclage macro : Oui
Analyse micro : Non			<i>(1) projections de tendances</i>	<i>(5) modèles macro- économétriques</i>
Analyse micro : Oui	Pas de modélisation « véritable » des comportements	Individus fictifs	<i>(2) cas-types dynamiques</i>	<i>(b)</i>
		Individus réels	<i>(3) microsimulation dynamique simple</i>	<i>(c)</i>
	Modélisation micro-fondée des comportements	Individus fictifs	<i>(a)</i>	<i>(6) MEGCGI</i>
		Individus réels	<i>(4) modélisation microéconométrique des comportements</i>	<i>(d)</i>

(1) Les projections de tendances

Les projections de tendances⁸ sont les façons les plus simples de prendre en compte les implications des évolutions démographiques sans se préoccuper de l'hétérogénéité des situations individuelles ni du bouclage macroéconomique, notamment l'effet de l'évolution des taux de cotisation, des ratios de dépendance, du niveau des prestations sur la dynamique économique. Il s'agit d'une approche approximative qui a une visée principalement illustrative. On fait une hypothèse de croissance dépendant de l'hypothèse de croissance de la productivité et d'évolution du chômage et de la population active⁹. Un cadre économique général est ainsi dressé à partir d'hypothèses sur l'évolution de l'âge moyen effectif de départ à la retraite. On déduit alors l'évolution du ratio de dépendance (rapport entre le nombre de retraités et le nombre de travailleurs). Compte tenu d'une hypothèse d'évolution du niveau de vie relatif entre retraités et actifs occupés (et entre chômeurs et actifs occupés), on en déduit la

⁸ Cf. Cornilleau et Sterdyniak (1995, 2003), Rapport Charpin, les travaux de la DREES repris par le COR.

⁹ La projection de population peut néanmoins reposer sur des matrices de transition assez élaborées et inclure un certain détail, et donc une hétérogénéité, notamment en termes de catégories sociales.

part des pensions retraite (et des prestations versées aux chômeurs) dans le PIB. Cela fournit une évaluation simple de l'évolution du poids des retraites pesant sur les actifs occupés due aux seuls changements démographiques.

Les projections de tendance permettent néanmoins d'aborder une formulation simplifiée de l'équilibre des finances des régimes de retraites. Ce dernier s'écrit :

$$(\text{pension moyenne}) \times (\text{ratio de dépendance}) = (\text{taux de cotisation}) \times (\text{salairé moyen}).$$

Le salaire moyen considéré est le coût du travail (salaire brut et cotisations patronales). Le taux de cotisation défini ici est un taux apparent. Il mesure le ratio entre les cotisations sociales versées et le coût du travail.

Appelons tr , le taux de cotisation retraite, r , le ratio de dépendance, et ar , le taux de pension net¹⁰, rapport entre la pension moyenne nette et le salaire moyen net. A chaque période, on a :

$$(\text{pension moyenne nette}) = ar \times (\text{salaire net moyen}).$$

L'équilibre du système de retraites s'écrit alors :

$$ar \times (1 - tr) \times r = tr ;$$

autrement dit :

$$tr = (ar \times r) / (1 + ar \times r).$$

Si $(1 - \alpha)$ est la part des salaires dans la valeur ajoutée, la part des retraites dans le PIB est :

$$(1 - \alpha) \times tr .$$

Avec cette relation simple (que l'on peut bien sûr généraliser un peu, notamment, en prenant en compte d'autres prestations sociales), on obtient la relation qui décrit la contrainte comptable à chaque période. Elle montre les conséquences de l'augmentation du ratio de dépendance (r) du fait du vieillissement : soit le taux de cotisations augmente, soit l'âge de départ augmente (r diminue), soit la pension moyenne relative (ar) diminue.

Cette approche ne décrit pas la réalité institutionnelle. Elle ne permet pas de prendre en compte la diversité des situations présentes et futures. Cependant, elle a deux grands avantages. Le premier est sa simplicité. Le second est qu'elle permet d'aborder la question des retraites de façon globale et de poser la problématique essentielle de la répercussion du vieillissement démographique sur le niveau de vie des actifs et des retraités. Plus précisément, elle permet de s'extraire d'un débat purement technique et institutionnel sur les réformes des régimes et d'évaluer le « triangle des contraintes » du système de retraite par répartition : le vieillissement se traduira nécessairement par une augmentation des cotisations sociales et/ou une diminution du niveau de vie relatif des retraités et/ou un recul de l'âge de départ à la retraite. La question du choix social devient effective lorsque l'ampleur des évolutions est chiffrée. C'est ce que permet cette approche. En France, le Conseil d'Orientation des Retraites (COR) estime par exemple que le maintien du rapport entre la pension moyenne nette et le salaire moyen net à son niveau actuel (78 %) suppose une augmentation du taux de cotisation de 15 points (par rapport au salaire brut)¹¹. Pour maintenir ce ratio sans augmenter les

¹⁰ Ce taux de remplacement net mesure le niveau de parité entre les pensions et les salaires nets. Ce ratio joue un rôle crucial dès lors qu'on souhaite comparer le niveau de vie des retraités à celui des salariés.

¹¹ En supposant une augmentation d'un an de l'âge de départ en retraite moyen, conséquence de la réforme du Régime général et du recul de l'âge de début des carrières.

cotisations, l'âge de départ à la retraite doit augmenter de 9 ans. Le débat est posé. Le COR décrit alors les solutions mixtes possibles.

Cette approche permet principalement de soumettre les débats techniques à un choix de société beaucoup plus large, mais elle est insuffisante notamment pour s'interroger sur les inégalités entre régimes ou à l'intérieur des régimes (par exemple, les pénalités fortes en cas de départ anticipé ou retardé...). Elle ne permet pas une approche descriptive de l'évolution des régimes existants (équilibres financiers, niveau des pensions versées). On peut cependant, de façon *ad hoc*, faire une estimation hors modèle de l'évolution du niveau relatif moyen des pensions nettes dans les régimes existants à partir de cas-types, et l'injecter dans la maquette macroéconomique. C'est ainsi que le COR prévoit, en l'état actuel des règles des régimes de retraite, une réduction de 78 à 64 % du taux de pension net. Cela supposerait d'augmenter le taux de cotisation de 9 points ou de reculer l'âge de 6 ans.

Le principal inconvénient de cette méthode est lié à la difficulté d'évaluer les réformes d'un point de vue pratique, c'est-à-dire, proposer leur mise en place et étudier leurs conséquences à un niveau individuel (identifier les gagnants et les perdants). A cet égard, l'approche par les cas-types¹² et la microsimulation peuvent être d'une grande utilité.

(2) *Les cas-types dynamiques*

L'approche par cas-types consiste à évaluer le niveau des pensions dont bénéficient et bénéficieront les retraités selon une carrière professionnelle donnée et des caractéristiques sociodémographiques données (nombre d'enfants, statut du foyer, ...). On mesure en général le taux de remplacement plutôt que le niveau de pension : le taux de remplacement (rapport entre la pension liquidée et le dernier salaire) permet d'évaluer dans quelle mesure le jeune retraité conserve son niveau de vie quand il prend sa retraite, mais aussi d'évacuer de la mesure la dérive liée à l'augmentation générale du niveau de vie dans l'économie. On peut aussi calculer le taux de rendement interne.

Autant d'un point de vue statique, les cas-types sont faciles à définir (les salariés au SMIC, les cadres moyens, etc.), autant d'un point de vue dynamique, la notion de cas-types est difficile à cerner¹³, notamment du simple fait qu'il existe de la mobilité sociale (un salarié ne reste pas forcément au SMIC pendant toute sa vie). En général, les cas-types adoptés sont représentatifs d'une certaine réalité¹⁴, mais il paraît difficile de se donner une distribution des cas-types représentative des différences observées dans la société. La pertinence d'un cas-type peut aussi différer selon les cohortes concernées : ainsi, pour la cohorte née en 1945, une entrée dans l'activité à l'âge de 16 ans est une hypothèse réaliste et pertinente ; elle ne l'est plus si on s'intéresse à une cohorte née en 1970. Par ailleurs, le choix d'un cas-type peut reposer uniquement sur sa capacité à illustrer une propriété particulière (avantage, anomalie, ...) du système de retraite.

¹² Par exemple, Dupont et Sterdyniak (2000) complètent leur analyse basée sur des projections de tendance par une série de cas-types censée illustrer des histoires salariales pertinentes.

¹³ Le principal mérite de la microsimulation est à ce stade de générer des cas-types qui ont leur pertinence, dans le sens où leur choix repose sur des propriétés de mutation sociale réellement observées, et d'être capable d'en évaluer les poids dans la population sociale, même si à propos de ces poids estimés, des questions de validité statistique restent probablement posées.

¹⁴ Des travaux récents tentent de développer des méthodes basées sur l'analyse des données pour identifier des profils types représentatifs (*cf.* Briard, 2003).

L'étude de cas-types nécessite une connaissance fine de la législation en matière de calcul des retraites (*cf.* le modèle OSCARI de la DREES – Ministère des Affaires sociales, voir Raynaud, 2003). Les approches par cas-type permettent d'avoir une vision simple du niveau de vie assuré par le système malgré sa complexité (systèmes de plafonds, cumul de plusieurs pensions...) et donnent une idée claire du niveau de vie des retraites futures et des propriétés du système dans toutes ses particularités, notamment la présence de non-linéarités ou discontinuités dans le calcul des pensions (voir par exemple, Dupuis et El Moudden, 2002).

Les cas-types peuvent être utilisés pour mesurer l'effet des réformes sur l'évolution des niveaux de vie ou les implications du système en termes de redistribution intra-génération (par exemple, savoir si les carrières ascendantes sont défavorisées ...). Ils permettent de comparer la générosité des différents régimes en évaluant les taux de remplacement assurés par chaque système à des profils de salaires identiques. Elle est aussi utilisée pour comparer la générosité des systèmes de retraites dans différents pays, notamment lorsque la logique des systèmes est proche (pays d'Europe continentale, pour une comparaison entre l'Allemagne, la France et l'Italie, voir Chagny *et alii*, 2001).

L'approche par cas-types est un point de départ nécessaire pour débiter l'analyse d'une réforme, mais inopérante pour appréhender les propriétés globales. Cette approche pourrait connaître un développement intéressant s'il était possible d'attribuer à chaque cas-type une pondération susceptible de reproduire des propriétés distributionnelles globalement observées. En effet, de telles pondérations permettraient d'amorcer une analyse plus globale et donc de constituer des agrégats de dépenses (pensions versées) et de recettes (cotisations sur les salaires). La limite du cas-type repose, pour l'instant, sur la difficulté de trouver ces pondérations. Une façon naturelle de dépasser cette limite est le recours à un modèle de microsimulation dynamique.

(3) La microsimulation dynamique

La microsimulation dynamique est un outil de simulation de trajectoires individuelles qui reflètent les parcours socio-économiques d'un échantillon d'individus représentatifs de la population. La nature représentative de l'échantillon permet de reproduire des ordres de grandeurs agrégées (moyenne des revenus, différents moments de la distributions des revenus, taux de chômage, etc) tout en conservant la richesse d'évaluation basée sur un niveau individuel. Ces modèles reproduisent alors une réalité agrégée de la démographie et des droits à la retraite. Ils permettent donc de déduire un équilibre des finances publiques à partir d'unités d'observation microéconomiques¹⁵.

La technique de microsimulation dynamique, dans sa version la plus simple, n'introduit pas véritablement des comportements. Cependant, les modèles de microsimulation peuvent être améliorés dès lors qu'on est en mesure d'introduire des règles de choix individuels qui reposent sur un comportement de maximisation d'un objectif sous contrainte, d'où une classification également possible avec les travaux de microéconométrie comportementale (*cf.* Klevmarken et Olovsson, 1996). Par ailleurs, si un certain bouclage macroéconomique est également pris en compte grâce à une utilisation associée avec un modèle macroéconométrique, un classement, dans le tableau, en position (c) ou (d) est envisageable. Enfin, certains modèles (*cf.* le recensement fourni dans la troisième partie)

¹⁵ Dans une certaine mesure, les évaluations sont basées sur ces cas-types réels (partie historique) et probables (partie simulée).

reposent sur des populations ou cohortes fictives, ce qui signifie que potentiellement, ils peuvent également se classer en (a), ou même avec les cas-types ou avec les modèles d'équilibre général calculable, en (b), si on a en plus un bouclage macroéconomique. Dans l'état actuel de la modélisation, la réalité est plutôt proche de la configuration de la case (3), car aucune tentative ne peut présenter un véritable fondement microéconomique global des comportements ni un bouclage réellement pertinent.

(4) Modélisation micro-économétrique des comportements

De nombreuses études économétriques essaient de tester et d'estimer des modèles comportementaux fondés sur des hypothèses de rationalité. Ces études permettent de mesurer des modèles comportementaux. Ces applications portent notamment sur la démographie (choix du nombre d'enfants), sur les choix de participation au marché du travail et aussi sur la modélisation des choix de départs en retraite (*cf.* le *survey* de Bommier *et alii*). Il n'y a pas de véritables mises en perspective car il faudrait pour cela disposer de prévisions.

Les modèles microéconométriques ont un « intérêt réduit » pour analyser les réformes dans la mesure où l'étroitesse de la portée de leur résultat dépend du particularisme du comportement étudié (choix de durée d'activité, sensibilité de l'offre de travail pendant l'activité, etc). Cependant, ces modèles peuvent être réemployés dans le cadre d'une microsimulation dynamique (sur ce point, voir Klevmarken, 1997). Dans ce cadre là, il peut se poser des problèmes de cohérence globale entre les différents modèles comportementaux car les microfondements ne sont pas nécessairement compatibles.

(5) Les modèles macro-économétriques

Ces modèles reposent sur un schéma d'analyse néokeynésien où la structure causale développée est du type : à court terme, la demande agrégée (la consommation, l'investissement, les dépenses des administrations publiques et les exportations) détermine l'offre agrégée (la production et les importations) qui, en retour, détermine les revenus et le taux de chômage ; les conditions de l'offre ont, à court terme, un effet sur le commerce extérieur, via la compétitivité et les tensions sur les capacités de production, d'une part, et sur la consommation via l'inflation, d'autre part. Les équations qui expriment les comportements macroéconomiques sont estimées par des procédures économétriques classiques employées pour les séries temporelles.

Ces modèles ne sont pas conçus pour s'appliquer à la réforme des systèmes de retraite ; pourtant leur potentiel d'utilisation se situe à plusieurs niveaux :

- A l'issue de la proposition d'équivalence ricardienne (Barro, 1974), de la théorie du cycle de vie (Ando *et alii*, 1957) ou du revenu permanent (Friedman, 1957), des auteurs comme Feldstein (1974) ont proposé d'introduire les droits à la retraite dans la fonction de consommation macroéconomique et ont trouvé une corrélation positive¹⁶.

¹⁶ Sur ce point, les études sont parfois contradictoires, et depuis la critique de Lucas (1976), ces outils de modélisation sont sujet à une utilisation prudente car les coefficients des équations traduisent une certaine anticipation de la politique économique ; si celle-ci change, ils se trouvent modifiés. La critique de Lucas doit pouvoir s'étendre aux modèles comportementaux sur données microéconomiques si ces derniers n'intègrent pas de façon pertinente les anticipations. Les MEGCGI échappent par construction à la critique de Lucas.

- Ils peuvent fournir une analyse de l'impact du vieillissement en termes de demande de bien et service (conséquences des transferts entre générations sur les consommations des jeunes et des vieux ; mutation de la structure de la demande et conséquences sectorielles).
- Les modifications de la structure démographique ont selon toutes vraisemblances des effets importants sur le marché du travail. Par exemple, l'entrée massive en retraite des générations du *baby-boom* devraient libérer de nombreux postes¹⁷. Dans le domaine de l'évaluation des retraites à long terme, une réflexion sur le taux de chômage est donc essentielle (*cf.* Rapport du Conseil économique et social). Leur attrait potentiel réside donc dans l'incorporation d'un taux de chômage endogène résultant de déséquilibres sur l'ensemble des marchés.

On peut noter en France le travail réalisé par Beffi *et alii* (2003) à l'aide du modèle MESANGE développé par l'Insee. Ce modèle présente des propriétés keynésiennes (régime de demande) à court terme et néoclassiques (régime d'offre) à long terme. Ils examinent différentes voies d'ajustement des paramètres institutionnels face à la déformation de la pyramide des âges. L'horizon de simulation est relativement long : l'année 2020 a été retenue.

(6) Les modèles d'équilibre général calculable avec générations imbriquées (MEGCGI)

Les MEGCGI fournissent des éléments d'évaluation sur les comportements d'accumulation de richesse, de consommation et d'offre de travail. L'ensemble des secteurs de l'économie est pris en compte : cohortes de ménages, entreprises et secteur public. L'économie est bouclée, les richesses produites sont consommées ou stockées. Les finances publiques sont intertemporellement équilibrées. Des marchés en situation de concurrence pure et parfaite fournissent des prix qui égalisent les offres et les demandes optimales de l'ensemble des agents¹⁸. La notion d'individu dans un modèle d'équilibre général calculable à générations imbriquées est abstraite, il s'agit d'un individu représentatif de l'ensemble d'une cohorte, et parfois d'une classe sociale ; ce dernier maximise son bien-être sur l'ensemble de son cycle de vie compte tenu des revenus qu'il anticipe rationnellement (salaires nets de cotisation, pensions retraite). Les trajectoires individuelles sont connectées à la dynamique macroéconomique globale de l'économie par l'intermédiaire de l'évolution des prix (le salaire et les taux d'intérêt) et l'évolution des finances sociales. Ce type d'exercice de simulation a une cohérence interne forte d'autant plus que les comportements sont parfaitement micro-fondés (voir le *survey* de Le Cacheux et Touzé, 2002). Les points faibles de ces approches proviennent en grande partie d'un contenu empirique, quant à lui, plus réduit au regard des autres approches. Notamment, la valeur des paramètres n'est pas obtenue par des techniques économétriques mais par des méthodes de calibrage, pour lesquelles il n'existe pas véritablement de tests (au sens de la théorie statistique) de la validation empirique.

¹⁷ Voir par exemple l'étude de Chauvin et Plane (2001) sur l'économie française qui tient compte à la fois de l'évolution de la population active (ralentissement de la croissance jusqu'en 2005 puis réduction à partir de 2010) sur le chômage (pour un scénario donné d'évolution de la demande de travail) et du chômage sur l'offre de travail (effet de flexion : les chercheurs d'emploi se découragent lorsque le chômage est élevé et se retirent du marché du travail).

¹⁸ Les travaux d'équilibre général calculable de type Blanchet (1992) avec des comportements de consommation spécifiés de façons *ad hoc* pourraient se classer selon notre typologie en (b).

Dans ces modèles, l'hétérogénéité en dehors de la prise en compte de l'âge a été jusqu'à présent rare. Cela signifie que les propriétés redistributives des systèmes de retraite sont difficiles à saisir. Néanmoins, la nouvelle génération de modèle introduit l'hétérogénéité de deux façons : soit une hétérogénéité innée de productivité (chaque cohorte est subdivisée en sous-groupes de productivité, sortes de cas-types dynamiques où il n'y pas de mobilité sociale pendant le cycle de vie entre chacun des sous-groupes) ou existence de chocs idiosyncratiques sur le revenu (sorte de microsimulation stochastique¹⁹ avec des chocs de productivité, un risque chômage ou un risque de mortalité) qui conduisent à des histoires différentes sur le cycle de vie. Kotlikoff (1995) a été l'un des premiers à s'intéresser à l'hétérogénéité. Sur ce sujet, on peut trouver en France deux applications récentes : par exemple, Hénin et Weitzenblum (2002) s'intéressent à la première configuration (chaque cohorte est subdivisée en plusieurs groupes de revenu), tandis que Hairault et Langot (2002, 2003) traitent un modèle qui se rapproche de la seconde configuration.

2. La microsimulation dynamique : principes et propriétés

La microsimulation est, à la base, une démarche de physicien et elle a été développée pour étudier les propriétés de systèmes stochastiques complexes. Le préfixe « micro » signifie que la simulation porte directement sur des entités élémentaires (cellules, particules ou individus), considérées alors comme les plus petites, et non sur des agrégats –ce dernier se trouve seulement simulé indirectement–. Il s'agit donc d'effectuer des projections sur ces entités élémentaires et non seulement sur leur masse²⁰. Les capacités actuelles de calcul font que cette technique s'emploie dans de nombreuses disciplines : la physique, la biologie, l'épidémiologie, l'oncologie, la chimie, ..., dans des optiques opérationnelles ou à des fins purement scientifiques, mais aussi dans le domaine de la démographie (projection de population), de la logistique et la gestion de réseaux de transport. Il faut noter aussi des utilisations potentielles en matière d'organisation pour les entreprises, notamment en ce qui concerne la gestion du personnel (prévoir les départs et les promotions, mais aussi estimer le passif social) et la gestion des stocks (par exemple, quelle est la règle qui minimise le risque d'une rupture de stock ?).

La microsimulation peut être statique ou dynamique. Les modèles statiques reposent sur des observations en coupe, et chaque entité élémentaire est étudiée sans perspective dynamique. En revanche, les modèles dynamiques ont pour objet de fournir des projections de trajectoires pour chaque composante élémentaire et ils sont construits à partir d'observations longitudinales. Appliqué à un ensemble composé d'individus (la population d'une nation), le modèle dynamique va être utilisé pour faire vieillir²¹ (ou rajeunir s'il s'agit de reconstituer des trajectoires individuelles incomplètement observées) un échantillon initialement observé de la

¹⁹ Dans ce type de modélisation, l'hypothèse d'un *continuum* de ménages élimine, pour calculer les agrégats, les problèmes liés à l'insuffisance de la taille des échantillons, comme cela peut se rencontrer en microsimulation dynamique (Cf. *infra*).

²⁰ Cf. Castiglione (2001) : "The use of computers has brought great advantages in handling complex models in two different although complementary ways: (1) the mathematical formulation of a model can be solved numerically using sophisticated algorithms to find a good "numerical" solution; (2) a system can be analysed in term of its constituents, i.e. the overall dynamics can be simulated by its very microscopic elements and global quantities can be compared with experimental data".

²¹ Il existe des modèles « faussement » dynamiques. Le vieillissement de la population consiste seulement à modifier les poids de chaque âge dans la population totale de façon à reproduire la déformation de la pyramide des âges. Cette technique a un sens lorsqu'il s'agit d'adapter des informations individuelles à une période différente de celle qui a servi à construire l'échantillon.

population, mais aussi pour faire « naître » de nouveaux individus puis pour « raconter » leur histoire, de façon à reproduire une évolution ressemblante, dans le sens où les propriétés observées en termes de corrélations croisées et de distribution des possibles sont bien reproduites, à celle que pourrait avoir la société.

On doit à Orcutt (1957) le premier projet d'application dans le domaine des sciences sociales, c'est-à-dire une utilisation de la microsimulation à des fins d'études des propriétés démographiques et économiques d'une nation. Orcutt se soucie à l'époque de la pauvreté de l'analyse distributionnelle offerte par les modèles macroéconomiques appliqués. Il propose alors une nouvelle voie fondée sur une modélisation des comportements individuels. L'aspect dynamique n'est véritablement traité qu'à partir des années 1970 dans le cadre de la construction aux Etats-Unis du modèle DYNASIM²². Depuis les années 1990, les physiciens (cf. Castiglione, 2001) emploient également des techniques similaires de microsimulation dynamique pour étudier les microstructures des marchés financiers (*econophysics*).

Dans cette partie, les informations nécessaires à la construction d'un modèle de microsimulation dynamique appliquée à l'analyse des retraites sont détaillées, avant d'aborder la nature technique de l'exercice ainsi que les interrogations sur la fiabilité des résultats. En guise d'illustration des problèmes liés à l'emploi d'une méthode de simulation stochastique, deux exemples élémentaires de microsimulation dynamique font l'objet d'une étude précise dans une annexe.

2.1. Les informations nécessaires à la construction d'un modèle

Dans le domaine économique et social, l'objectif de la microsimulation dynamique est de reproduire des trajectoires de cycles de vie qui soient pertinentes au regard des processus d'évolution sociale observés et des besoins de l'analyse. Ces processus décrivent les lois qui gouvernent les destins individuels. Ils traduisent les conséquences individuelles du fonctionnement global de la société. Ces conséquences traduisent des choix d'action (mariage, départ à la retraite, etc) ou des obligations (licenciement, naissance, décès, etc). Les destins individuels sont dépendants des caractéristiques socio-économiques : l'environnement économique et l'histoire passée concourent au devenir de chacun, mais ce devenir est en partie aléatoire, c'est-à-dire qu'il est en partie inexplicable et donc soumis, à défaut, au hasard²³.

Ces processus font l'objet d'une modélisation microéconométrique préalable (spécification d'une relation puis estimation des paramètres de la relation identifiée). Le choix d'une relation consiste à déterminer les éléments connus qui expliquent le mieux, d'un point de vue logique et statistique, les évolutions des différentes valeurs prises par chaque caractéristique ainsi que la nature purement aléatoire. A ce stade, l'un des attraits de la microsimulation est de tenir compte des corrélations croisées entre les différentes caractéristiques et leurs valeurs passées. Les observations des variables se réalisent sous une forme quantitative continue (mesurable dans une unité définie préalablement, par exemple, le salaire en euro), quantitative discrète (par exemple, l'âge) ou qualitative (non mesurable et par nature discrète). En particulier, les méthodes économétriques employées pour estimer les processus qui décrivent l'évolution de variables discrètes consistent à estimer directement la

²² Pour plus de détails, voir le recensement des différents modèles existants réalisé dans la troisième partie.

²³ Dans les modèles comportementaux de cessation d'activité, il est possible d'introduire l'évolution anticipée de la législation des retraites. Il peut alors se poser des problèmes de résolution du modèle global sous l'hypothèse d'anticipations rationnelles.

probabilité de survenance d'un évènement (modèle LOGIT ou PROBIT). Lorsque le temps est considéré comme continu, les méthodes économétriques utilisées font appel à des modèles de durée ; il s'agit alors de fournir une estimation de la distribution de probabilité du temps écoulé entre deux évènements.

En règle générale, ne sont reprises que les conséquences des caractéristiques individuelles et globales (par exemple, le taux de chômage) sur les histoires individuelles, et non, les conséquences des histoires individuelles sur les variables globales (notion de bouclage macroéconomique). Cependant, les paramètres institutionnels du régime de retraite peuvent se relier facilement aux histoires individuelles dès lors que les modélisateurs ont la possibilité d'inclure des règles d'ajustements des taux de cotisation et des règles de calcul des retraites (taux de remplacement, âge de départ) qui préservent l'équilibre des finances sociales (bouclage partiel). En retour, ces ajustements peuvent jouer un rôle important lorsque les paramètres institutionnels modifient les comportements individuels (par exemple, le mode de calcul des retraites peut influencer les incitations à la cessation d'activité).

L'élément clé dans le calcul des retraites est l'histoire salariale. L'occupation d'un emploi et le montant du salaire constituent donc les deux caractéristiques socio-économiques de base qui doivent être observées et modélisées. Un tel besoin va nécessiter l'identification ou la reconstitution de carrières salariales. En outre, compte tenu des clauses de réversion dans le calcul des retraites en faveur du conjoint survivant et de l'utilisation possible de la microsimulation pour analyser la transmission patrimoniale, il peut paraître utile de connaître les liens familiaux entre les différents individus. Dans ce but, de nombreux modélisateurs utilisent d'ailleurs une population fermée, c'est-à-dire une population où l'ensemble des liens familiaux (filiation, mariage ou union libre voire divorce s'il s'agit de reproduire les ruptures de liens) entre les différents individus est conservé ou créé à l'aide d'algorithmes d'appariement. Ces caractéristiques sont souvent les principaux éléments à reproduire pour être en mesure de calculer les pensions retraites.

La description des individus comprend d'autres caractéristiques. Chacune d'elle nécessite aussi une modélisation particulière de type microéconométrie. La liste suivante récapitule l'ensemble des informations généralement intégrées dans les différents modules, dont l'interconnexion globale constitue le modèle de microsimulation (voir par exemple Insee, 1999) :

(i) La démographie : pour évaluer les évolutions de la démographie, il est nécessaire de comprendre les mécanismes de migration interne géographique (entre différents états ou régions d'une zone économique), d'immigration / émigration, ainsi que les évolutions des taux de survie et fécondité (distribution des âges pour le premier enfant puis les autres).

(ii) La nuptialité : l'unité individuelle n'est pas suffisante pour appréhender correctement la législation. Il faut en général être en mesure de connaître les caractéristiques de la cellule familiale. Il s'agit alors d'identifier les différents éléments qui expliquent la formation de liens de type mariage ou union libre ainsi que les divorces.

(iii) L'éducation et l'acquisition de l'indépendance : il s'agit d'identifier et de reproduire différents cursus scolaires et de transmission de capital humain ainsi que les mécanismes de décisions de départ du domicile parental. L'acquisition de l'indépendance est en général connectée à la carrière professionnelle.

(iv) L'état de santé : il peut apparaître important de connaître les caractéristiques des individus en termes de morbidité dans la mesure où la mortalité est reliée à la maladie et dans la mesure où le financement des dépenses de santé des plus âgés peut s'apparenter à une forme de système par répartition où les prestations sont en nature (pour un *survey* sur les modèles de microsimulation appliqués à la santé, voir Breuil-Genier, 1998).

(v) Les carrières professionnelles : la modélisation des carrières est un point crucial. Il s'agit en particulier de retracer les transitions entre les différents états d'activité (emploi, chômage) et d'inactivité (scolarité, retraite, éducation des enfants), les mécanismes de promotion salariale ou de croissance des revenus d'activités pour les indépendants, mais aussi les changements sectoriels dans la mesure où ils peuvent occasionner des changements de régimes de rattachement (par exemple, carrières mixtes dans le privé et le public).

(vi) Les revenus : pour appliquer certains barèmes fiscaux, il est souvent nécessaire de connaître l'ensemble des revenus des ménages et de distinguer les revenus issus du travail (salaire, nombre d'heures travaillées, etc) de ceux résultant de la détention d'un capital.

(vii) Le patrimoine : il s'agit d'identifier les principaux éléments qui motivent les comportements patrimoniaux (acquisition de logement ou placements financiers) ainsi que les logiques dynastiques (les héritages et les transmissions *in vivo* de patrimoine). L'évaluation du montant des patrimoines est un aspect important si on souhaite connaître les revenus complémentaires à ceux des pensions par répartition.

(viii) Les institutions publiques et la législation sociale²⁴ : il s'agit de décrire les différents instruments de taxation ou de versements de prestations ainsi que les formes d'assurances publiques (et d'affiliation) en matière de retraite, santé, survie et dépendance.

Ainsi, en appliquant des tables de mortalité conditionnelle, des comportements de choix de nuptialité et de maternité, des règles d'évolution des salaires (ou de revenus d'exploitation pour les entrepreneurs individuels), des règles de licenciement ou des choix de cessation d'activité professionnelles, etc, il va être possible de produire (ou reproduire s'il s'agit d'invention des passés) l'évolution professionnelle et familiale probable d'un échantillon d'individus pour une population donnée. On peut alors en déduire, par exemple, la progression des dépenses des branches vieillesse (ou du risque dépendance) tout en ayant des ordres de grandeurs sur l'évolution des distributions des différentes caractéristiques socio-économiques dans la population.

Il est ainsi possible d'étudier l'impact d'une réforme (modification du taux de cotisation, des conditions de départ en retraite ou du calcul de la pension) selon un certain agenda de mise en place progressive tant au niveau global (les montants totaux de dépenses et de recettes) qu'en termes d'incidence sur les inégalités intra et intergénérationnelles (par exemple en mesurant les effets des réformes par décile de revenu de cycle de vie). Le critère généralement utilisé pour mesurer la générosité du régime de retraite pour un individu est son rendement ou plus précisément son taux de rendement interne (TRI).

²⁴ Cet aspect de la modélisation constitue souvent une des étapes préliminaires du modèle de microsimulation. Il s'agit de construire une banque de données sur la législation ainsi qu'un programme capable d'évaluer le montant de la retraite pour chaque configuration possible d'histoire salariale et familiale. Un modèle de microsimulation intègre donc un modèle de cas-types.

2.2. La nature technique de l'exercice

D'un point de vue technique, un modèle de microsimulation dynamique est un générateur de trajectoires individuelles. Ces trajectoires ont une nature stochastique²⁵ puisque, à chaque période, le hasard affecte un avenir parmi un ensemble de possibles. Il permet ainsi de constituer un échantillon de trajectoires individuelles qui est susceptible de refléter, dans la mesure où le balayage aléatoire est supposé suffisamment large et non biaisé, la réalité économique tant dans ses particularismes individuels (propriétés en termes de distributions longitudinales et verticales) que dans ses tendances globales (évolution générale qui caractérise la société, et dont on peut établir des pronostics en dehors d'un modèle de microsimulation).

Chaque processus identifié se présente sous une formulation mathématique quantifiée qui traduit une relation logique avec des éléments endogènes que sont les caractéristiques présentes et passées de l'individu (qui résultent eux-mêmes des chocs passés), des éléments exogènes de nature non stochastique²⁶ (par exemple, la législation) et d'un aléa dont la distribution est connue. Ce dernier pourra être reproduit à l'aide d'un générateur de nombres au hasard (il s'agit de la nature exogène et stochastique du processus).

De façon concrète, la description de la société se réalise à l'aide d'informations de nature discrète (qualitative ou mesure quantitative) ou continue. Le cœur d'un modèle de microsimulation dynamique peut donc être schématisé de la façon suivante. Considérons des individus décrits par des caractéristiques de nature discrète (vecteur de variables qualitatives ou quantitatives noté y) et de nature continue (vecteur de variables quantitatives noté x). On note z le vecteur de variables exogènes²⁷ qui décrivent les environnements institutionnel (taux de cotisation, taux de remplacement, etc ; ces paramètres se trouvent modifiés lorsqu'une réforme est simulée) et macroéconomique (hypothèses sur le taux de chômage, le taux de croissance de la productivité, etc). La simulation de réformes des systèmes de retraite signifie l'utilisation de valeurs institutionnelles alternatives. t désigne l'indice temporel et i l'indice individuel.

A un instant donné t , on observe une distribution des valeurs prises par les différentes caractéristiques dans la population :

Les valeurs prises par les caractères continus $x_{it} \in \Omega$ sont distribuées suivant une loi caractérisée par une fonction $F_t(x)$ qui mesure la proportion d'individus i à l'instant t tels que $x_{it} \leq x$.

Les valeurs prises par les caractères discrets $y_{it} \in \{y^1, \dots, y^j, \dots, y^N\}$ sont distribuées dans les proportions $p_t^y = \text{prob}(y_{it} = y)$ pour $y = y^1, \dots, y^N$ avec $\sum_y p_t^y = 1$.

²⁵ En comparaison, on calcule généralement l'espérance du processus pour chaque individu dans le cadre de la microsimulation statique.

²⁶ C'est justement au législateur que revient la tâche de la détermination, mais le modélisateur peut introduire un bouclage (partiel) afin de reproduire la contrainte d'équilibre des finances publiques.

²⁷ Le statut d'exogénéité de ces variables reflète l'absence de bouclage macroéconomique.

Pour un environnement global (économique et institutionnel) donné z_t , l'évolution des valeurs prises par les caractéristiques est gouvernée par des processus stochastiques. Pour simplifier la présentation, seules les caractéristiques passées –l'étendu du passé étant limitée à la période précédente– sont supposées influencer sur ces processus : aucune co-dépendance directe (simultanéité) entre deux caractéristiques observées au même instant n'a été supposée. On peut représenter ces processus de la façon probabiliste suivante :

$$\left\{ \begin{array}{l} \text{La fonction } G_t(x', x, y, z_t) \text{ mesure la probabilité conditionnelle pour un individu } i \\ \text{d'avoir une réalisation } x_{i,t+1} \leq x' \text{ sachant que sa situation passée était } x_{i,t} = x \text{ et } y_{i,t} = y ; \\ \text{La fonction } q_t(y', x, y, z_t) \text{ mesure la probabilité conditionnelle pour un individu } i \\ \text{d'avoir une réalisation } y_{i,t+1} = y' \text{ sachant que sa situation passée était } x_{i,t} = x \text{ et } y_{i,t} = y. \end{array} \right.$$

Idéalement²⁸, on souhaiterait, à chaque période, calculer l'évolution exacte²⁹ de la distribution des valeurs prises en coupe (analyse transversale) par les différentes variables dans la population :

$$\left\{ \begin{array}{l} F_{t+1}(x') = \sum_y p_t^y \times \int_{\Omega} G_t(x', y, x, z_t) dF_t(x) \\ p_{t+1}^{y'} = \sum_y p_t^y \times \int_{\Omega} q_t(y', y, x, z_t) dF_t(x) \end{array} \right. ,$$

ainsi qu'à mémoriser l'intégralité les trajectoires individuelles pour connaître la distribution longitudinale des trajectoires, que l'on pourrait mesurer par exemple comme la proportion d'individus i tels que les histoires individuelles vérifient : $(x_{i0} \leq x_0; \dots; x_{it} \leq x_t)$ et $(y_{i0} = y_0; \dots; y_{it} = y_t)$.

En pratique, on ne peut pas calculer la valeur théorique de ces distributions car la procédure de simulation repose sur une sélection de trajectoires, par succession de tirages aléatoires, pour un grand nombre d'individus et non sur une méthode de calcul basée sur un *continuum* d'individus. Le *continuum*³⁰ – atomicité de chaque agent – serait une garantie pour calculer avec exactitude la distribution. Pour cela, il faudrait recourir à des méthodes d'intégration numérique. De tels calculs nécessiteraient également des ressources informatiques importantes³¹. Avec un modèle de microsimulation, les distributions sont

²⁸ Par exemple, Stiglitz (1971), en pionnier, s'est intéressé à la dynamique des inégalités d'un point de vue théorique et en équilibre général. Dans cet article, il se préoccupe des propriétés théoriques de la dynamique des inégalités et étudie la convergence de la distribution en coupe. Sur ce sujet, voir également Bourguignon (1981).

²⁹ Au cas où la taille d'échantillon nécessaire pour assurer une telle mesure serait plus grande que la taille de la population, il pourrait aussi être intéressant de connaître les propriétés d'échantillonnage pour des échantillons de taille égale à celle de la population étudiée.

³⁰ La notion de *continuum* peut sembler extrême car la population est, par nature, de taille finie.

³¹ Certains modèles d'équilibre général calculable (MEGC) avec hétérogénéité d'agents utilisent des techniques d'intégration numérique pour évaluer la dynamique de la distribution des richesses sous hypothèse de *continuum*. Par exemple, Den Haan (1997) simule la dynamique agrégée d'une économie en équilibre général avec différentes tailles d'échantillon qu'il compare aux résultats obtenus avec un continuum. Dans un modèle dynamique très simple d'évaluation du prix des actifs financiers avec chocs individuels non assurable à la Lucas (1994), il trouve une certaine convergence de la dynamique et des propriétés sur les distributions de richesse sous hypothèse de *continuum* et celles obtenues avec une population composée de 300.000 individus. Il conclut en faveur de l'intégration numérique : "This suggests that obtaining accurate solutions of this model using the Monte Carlo techniques [...] requires a large number of agents which would make the algorithm computationally very expensive".

seulement calculées à partir d'un nombre fini d'individus. Sous certaines hypothèses, la loi des grands nombres peut assurer l'annulation tendancielle de la variance de l'estimateur de la moyenne mais la taille de l'échantillon nécessaire peut être considérable. Tout dépend alors de la vitesse de convergence vers 0 de la variabilité de l'estimateur du moment considéré. Il en ressort que les résultats agrégés (les différents moments des lois de distribution des données longitudinales ou en coupe) issus des simulations peuvent suivre un processus aléatoire complexe avec une variance conditionnée par la taille de l'échantillon (voir les éléments de discussion sur deux exemples en annexe). L'horizon temporel de simulation joue également un rôle car les erreurs s'accumulent et la variance de fluctuation de l'agrégat peut croître avec la durée de la projection. Par ailleurs, pour les distributions longitudinales, de nombreuses trajectoires ne seront jamais simulées. Il y a certainement matière à réfléchir sur l'importance de ces « disparitions » dans l'utilisation des résultats agrégés³², notamment pour calculer des indicateurs d'inégalités intra et intergénérationnelles ou de générosité du système.

2.3. Fiabilité des résultats

La microsimulation dynamique donne deux principaux résultats : une évaluation du montant des droits à la retraite à un niveau individuel ainsi qu'à un niveau agrégé. Il est important de s'interroger sur la pertinence de ces mesures. La fiabilité des résultats³³ s'apprécie en comparant les objectifs idéaux de la microsimulation et les contraintes techniques (qualité des données utilisées, taille des échantillons, choix de spécification et puissance de calcul des ordinateurs). Les objectifs idéaux de la microsimulation sont de fournir des estimations de la réalité probable dont la qualité de précision se mesure par les faiblesses du biais en espérance et de la variance³⁴. Les sources d'imprécision dans ces estimations proviennent de la nature même de la microsimulation (emploi de banques de données et de techniques de modélisation économétrique), mais aussi de son aspect dynamique (prise en compte du temps et utilisation d'une méthode de simulation stochastique).

2.3.1. Les sources d'imprécision liées à la microsimulation en général

Les sources d'imprécision liées à la microsimulation en général sont de quatre types :

- (i) La construction d'un modèle de microsimulation nécessite le recours à de nombreuses banques de données tant pour caractériser les individus que pour identifier et quantifier les processus qui déterminent les changements d'état de chacun d'entre eux. Il peut donc apparaître deux types d'erreur d'échantillonnage³⁵ qui vont concerner les individus sur lesquels porte précisément la microsimulation, mais aussi, les individus dont l'observation a permis l'estimation des paramètres du modèle.

³² On peut noter que le calcul d'un indicateur comme le revenu de cycle de vie permet d'introduire un deuxième lissage en plus de la sommation sur les individus : celui de la sommation des salaires sur l'ensemble de la durée de la vie.

³³ Klevmarken (2002a, 2002b) et Wolf (2000a) donnent quelques éléments d'analyse concernant la fiabilité des estimations.

³⁴ Il existe des techniques statistiques pour réduire la variance des estimateurs. Ces techniques ne sont pas présentées ici.

³⁵ La note de Wolf (2000a) fournit une discussion intéressante sur les erreurs d'échantillonnage.

- (ii) Les spécifications retenues : le recours à une spécification de comportement ou d'une relation causale repose toujours sur un certain arbitraire et expose donc à l'oubli de variables ou à l'emploi de formulations inadaptées (par exemple, modélisation linéaire plutôt que non-linéaire, traitement des anticipations individuelles, etc).
- (iii) Les contraintes d'observation : les spécifications des processus peuvent en raison d'observations manquantes ignorer l'importance de certaines variables.
- (iv) Certains modèles utilisent une population (ou cohorte) fictive, qui reproduit, seulement et en principe, certaines propriétés distributionnelles (en termes de moments et de corrélations croisées) de la population (ou cohorte) réelle. Pour ces types d'échantillons, il est très difficile de connaître les propriétés d'échantillonnage.

2.3.2. *Les sources d'imprécision spécifiques à la microsimulation dynamique*

Les sources d'imprécision spécifiques à la microsimulation dynamique proviennent de l'aspect dynamique et de l'utilisation d'une technique de simulation stochastique.

- (i) La reconstitution des passés : lorsque les données sur le passé sont manquantes, il est nécessaire de rétropoler, et de la même façon que l'on peut se tromper sur les estimations futures, on peut commettre des erreurs sur les passés.
- (ii) La discrétisation du temps³⁶ (fréquence annuelle en général) peut conduire à des approximations sur la durée effective des différents événements (par exemple, on n'est pas forcément chômeur toute l'année, mais seulement une partie).
- (iii) Les estimations de tendance dans l'évolution des comportements ou de certaines variables d'environnement macroéconomique (par exemple, le taux de chômage, l'allongement de la durée de la vie, l'allongement de la durée des études, l'augmentation de l'offre de travail féminine, l'évolution du progrès technique) peuvent être erronées.
- (iv) L'absence de bouclage macroéconomique³⁷ : les modèles de microsimulation produisent des supports d'analyse prospective tout en restant en équilibre partiel (environnement global prédéfini). Autant, dans une analyse statique, on peut considérer dans un premier temps (en première approximation) que le problème étudié de façon isolée n'a pas d'influence sur le reste de l'économie ; autant dans une perspective dynamique sur longue période (entre 40 ans et un siècle), le raisonnement

³⁶ Pour une discussion sur les intérêts respectifs des modèles en temps discret ou des modèles en temps continu, on peut se référer à Galler (1997).

³⁷ Une utilisation jointe d'un modèle de microsimulation et d'un modèle macroéconométrique est en principe possible. Dans cette configuration, le caractère limité des résultats est à reconsidérer. Toutefois, il y a lieu de s'interroger sur la pertinence économique et statistique d'une telle utilisation jointe.

avec scénario macroéconomique inchangé, lorsqu'on introduit des réformes, est plus délicat. Parmi, les évolutions macroéconomiques susceptibles d'être également influencées par les réformes, on peut noter le taux de chômage (sensible à l'évolution de l'offre de travail), l'accumulation de capital productif et la demande globale (par exemple, évolution de la consommation des vieux par rapport à la consommation des jeunes).

- (v) L'imperfection des générateurs de nombres au hasard : *a priori*, cet élément est marginal, compte tenu des études préalables de validation qu'ils suscitent.
- (vi) La méthode de simulation stochastique s'apparente à une technique d'échantillonnage de trajectoires, et l'insuffisance des tailles d'échantillons introduit une variabilité artificielle dans les résultats agrégés³⁸. Cette variabilité se traduit par des évolutions erratiques. Cette insuffisance peut rendre délicate l'exploitation des distributions longitudinales pour évaluer le montant total des droits à pension à un niveau agrégé ou traiter des questions d'équité.

Pour beaucoup de ces sources d'imprécision, il est difficile d'établir une mesure de la variance supplémentaire et des biais qu'elles peuvent induire. Pour certains types d'imprécision, il est toujours possible de recourir à des méthodes d'inférence statistique, mais celles-ci nécessitent d'importantes ressources informatiques. Etablir une marge de confiance globale semble donc plutôt difficile³⁹. C'est pourquoi, les marges réelles de précision des estimations que les économètres ont l'habitude de présenter lorsqu'ils réalisent des prévisions macroéconomiques à l'aide de modèles usuels de séries temporelles, ne peuvent pas vraiment être fournies. Par ailleurs, la fiabilité des estimateurs, obtenus par simulation stochastique, des différents moments des distributions peut également poser un soucis important. Ce point d'interrogation fait l'objet d'une discussion particulière dans le cadre de l'étude, qui figure en annexe, de deux processus stochastiques élémentaires.

3. La microsimulation dynamique en pratique : Un recensement des principaux modèles

L'introduction de la microsimulation dans le domaine des sciences sociales est due à Orcutt (1957). Partant du constat de l'incapacité des modèles macroéconomiques de renseigner sur l'impact des politiques publiques sur, par exemple, la répartition des revenus

³⁸ Cf. Blanchet et Chanut (1997) : « Le principal inconvénient réside dans le caractère stochastique du résultat, qui vient de ce que la simulation des évolutions individuelles fait intervenir des tirages aléatoires. Le problème que pose ce caractère stochastique dépend néanmoins du type de résultats considérés. Si on s'intéresse à des indicateurs de stock très agrégés, leur évolution sera peu sensible à ce caractère stochastique même si la microsimulation porte sur un fichier de taille modérée. L'évolution des variables de flux ou des variables relatives à de petites sous-populations sera en revanche beaucoup plus instable. Dans ce cas, deux solutions existent. L'une est de répéter les simulations à l'aide de tirages indépendants et d'en effectuer des moyennes (mais ceci n'élimine pas les effets d'échantillonnage tenant à la composition de l'échantillon de départ), l'autre est de travailler sur une base individuelle plus large. Dans les deux cas, les temps d'exécution des programmes sont allongés d'autant. »

³⁹ Par exemple, Pudney et Sutherland (1996) fournissent des intervalles de confiance de résultats, mais ces derniers sont obtenus avec un modèle comportemental de microsimulation statique.

des agents dans l'économie, Orcutt proposa de développer des outils de simulation basés sur des données microéconomiques. Tout d'abord statiques⁴⁰, les modèles de microsimulation ont vu l'introduction d'éléments dynamiques au cours des années 1970. Toutefois, bien que la construction du premier modèle de microsimulation dynamique (DYNASIM) ait été initiée dès 1969 aux Etats-Unis, ce n'est que depuis une quinzaine d'années que cet outil d'analyse économique connaît son essor. L'accroissement des capacités de calcul et la disponibilité de données microéconomiques sont les principaux facteurs de son développement dans les années 90. Dans cette partie, nous tentons de dresser un inventaire des principaux modèles existants. Les articles de O'Donogue (2000) et de Zaidi et Rake (2001) sont deux articles de référence qui fournissent de nombreuses informations sur le contenu des modèles. O'Donoghue (2000) inventorie et donne des détails techniques précis sur les modèles. Zaidi et Rake (2001) fournissent un recensement moins complet mais tentent d'énoncer un certain nombre de règles et bilans sur les conditions de faisabilité et de fiabilité d'un modèle. Dans une perspective d'analyse statistique, on peut noter l'article de Wolf (2000) qui présente un bilan critique de techniques économétriques d'estimation employées dans la construction des modèles de microsimulation. Un recensement des principaux modèles opérationnels dans le monde est tout d'abord effectué. Le cas français avec le modèle DESTINIE est ensuite abordé plus en profondeur.

3.1. Les principaux modèles opérationnels

Aujourd'hui, on recense des modèles de microsimulation dynamique dans la quasi-totalité des pays développés. Dans les tableaux 1, 2 et 3, un résumé des particularités des principaux modèles est donné. Ce descriptif comprend six rubriques : le pays d'étude, l'institution qui assure la construction du modèle, les références bibliographiques majeures, le type de base de données utilisé, la nature (population réelle ou fictive) et la taille de l'échantillon ainsi que la dimension du temps (discret ou continu). Les modèles ont été regroupés par zone géographique : l'Amérique du Nord, l'Europe et l'Asie-Océanie.

Par son rôle pionnier et sa richesse en matière de bases de données, les Etats-Unis sont incontestablement le pays le mieux doté. Les modèles DYNASIM 2 et CORSIM sont certainement les plus aboutis. Ce dernier est à la base du modèle canadien DYNACAN et du modèle suédois SVERIGE. Bien que la construction des modèles reste majoritairement dans le domaine de la recherche, les institutions administratives sont de plus en plus partie prenante (DYNACAM, MINT, SFB3, DESTINIE, PENSIM, MOSART, SESIM), ce qui révèle l'intérêt que présente la microsimulation dans l'évaluation des politiques publiques. La quasi-totalité des modèles présentés est utilisée afin d'analyser les systèmes de retraite. On notera néanmoins que certains d'entre eux permettent également d'étudier les politiques de santé (DYNAMOD, HARDING, LifePaths, NEDYMAS). L'Europe est la région du monde qui regroupe le plus de modèles, mais cette particularité résulte de la multitude des régimes de retraite. L'Italie, le Royaume-Uni et la Suède sont les trois pays les mieux dotés, avec plusieurs modèles opérationnels. En Australie, le modèle NATSEM a un rôle dominant et réalise de nombreux travaux méthodologiques sur le sujet, qui sont diffusés sous la forme de *technical papers* (par exemple, Galler, 1997 ; Baekgaard, 2002).

⁴⁰ Voir Sutherland (1998) pour une étude récente des modèles de microsimulation statique.

Tableau 1 - Modèles de microsimulation dynamique en Amérique du Nord

Pays	Institution	Modèle	Références	Base de données	Taille de l'échantillon	Temps discret (D) ou continu (C)
Canada	Canada Pension Plan (CPP)	DYNACAN	Caldwell et Morrison (2000)	1% de l'échantillon du recensement 1971	212.000 individus	C
	Statistics Canada	LifePaths	Rowe et Wolfson (2000)	Population fictive (ensemble de cohortes)	variable	C
	Statistics Canada	DEMOGEN	Wolfson (1988)	Cohorte fictive	1.000 à 5.000 individus	D
Etats-Unis	Urban Institute	DYNASIM I et II	Orcutt et alii (1976) Hendriks et Holden (1976) Werttheimer et alii (1986)	Current Population survey (CPS) 1973 + Social Security Administration	60.000 individus	mixte
	Cornell University	CORSIM	Caldwell (1996) Caldwell et Morrison (2000)	1% de l'échantillon du recensement 1960	180.000 individus	D
	Social Security Administration (SSA)	MINT	Butrica et alii (2001)	Survey of Income and Program Participation 1990-93 + Social Security Administration	85.000 individus	mixte
	US Department of Labour	PENSIM/2	Holmer et alii (2001)	Cohorte fictive	variable	C
	ICF/Lewin Inc.	PRISM	Kennell et Sheils (1986)	Current Population survey (CPS) 1978-79 + Social Security Administration	28.000 individus	D

Tableau 2 - Modèles de microsimulation dynamique en Europe

Pays	Institution	Modèle	Références	Base de données	Taille de l'échantillon	Temps discret (D) ou continu (C)
Allemagne	Universités de Frankfurt, Mannheim and Berlin	SFB3 (Population)	Galler et Wagner (1986)	Fichier de données micro 1969	69.000 ménages	D
		SFB3 (Cohorte)	Hain et Hellberger (1986)	Cohorte fictive	7.300 individus	D
Belgique	DULBEA, Université libre de Bruxelles	Modèle Retraite	Joyeux et alii (1996)	Population fictive basée sur données micro	Variable	D
France	INSEE	DESTINIE 1997	Blanchet et Chanu (1997) INSEE (1999)	Enquête Actifs financiers 1991	37.000 individus	D
		DESTINIE 2001	Walraet et Vincent (2002)	Enquête Patrimoine 1998	50.000 individus	D
Irlande	University of Cambridge	Modèle dynamique	O'Donoghue (2000)	Living in Ireland survey 1994	4.500 ménages	D
Italie	Université de Pennsylvanie	ANAC	Ando et Nicoletti Altimari (1999)	Enquête sur les revenus des ménages 1993	67.000 ménages	D
	Banque d'Italie	DYNAMITE	Ando et alii (2000)	Enquête sur les revenus des ménages 1993	67.000 ménages	D
	University of Bologna	Modèle de cohorte	Baldini (1997)	Cohorte fictive	4.000 individus	D
Norvège	Statistics Norway	MOSART	Fredriksen (1998)	Données administratives 1989	40.000 individus	D
Pays-Bas	Université de Tilburg	NEDYMAS	Nelissen (1993,2000) Andreassen et alii (1996)	Population fictive basée sur le recensement 1947	10.000 individus	D
Royaume-Uni	London School of Economics (LSE)	LIFEMOD	Falkingham et Johnson (1995)	Cohorte fictive	4.000 individus	D
	Department of Work and Pensions	PENSIM	Handcock et alii (1992)	Retirement survey 1988 + social change and economic life initiative survey 1986 + Family expenditure survey 1988	5.000 allocataires	C
Suède	Uppsala University	MICROHUS	Klevmarken et Olovsson (1996)	HUS income distribution 1984	environ 2000 ind.	C
	Ministère des finances	SESIM	Ericson et Hussenius (1998, 1999)	HINK survey 1992	30.000 individus	D
	Spatial Modelling Center	SVERIGE	Holm et alii (2002)	Données administratives 1985-1995	9.000.000 ind.	D

Tableau 3 - Modèles de microsimulation dynamique en Asie et Océanie

Pays	Institution	Modèle	Références	Base de données	Taille de l'échantillon	Temps discret (D) ou continu (C)
Australie	NATSEM, Canberra University	DYNAMOD I et II	Antcliff (1993) Bracher (1996) King et alii (1999)	1% du recensement 1986	150.000 individus	mixte
		HARDING	Harding (1993)	Cohorte fictive	4.000 individus	D
	University of Melbourne	Modèle de cohorte de Melbourne	Van de Ven (1998)	Population fictive	50.000 hommes (mariés ou non)	D
Japon	National Institute of Research Advancement (Tokyo)	Modèle de cohorte	Ando (1996)	Cohortes fictives	4.000 individus	D
Nouvelle Zélande	Business and Economic Research Ltd (BERL)	MIDAS	Stroombergen et alii	Population fictive basée sur le recensement 1991	10.000 individus	D

Depuis l'origine, les comparaisons entre les différents modèles sont fréquentes, et elles ont conduit à plusieurs publications collectives. Parmi celles-ci, on peut remarquer celles sous la direction de Orcutt *et alii* (1986), Hancock et Sutherland (1992), Harding (1993, 1996) et Milton *et alii* (2000). Les besoins de concertation se traduisent également par l'apparition de groupes de travail internationaux. Par exemple, le *ESRC-Sage International Network*⁴¹ rassemble autour du Sage (Royaume-Uni) des constructeurs en provenance de quatre pays européens : le SFI (Institut indépendant et affilié au Ministère des Affaires sociales) pour le Danemark, l'INSEE pour la France, le ESRI (*Economic and Social Research Institute*) pour l'Irlande ainsi que l'université de Lund et le *National Insurance Board* pour la Suède. Le SAGE est un centre de recherche affilié à la *London School of Economics*.

3.2. L'expérience française en particulier

Le travail fondateur de la microsimulation dynamique en France est attribuable à Blanchet et Chanut (1997) et est présentée de façon détaillée dans Insee (1999). Le modèle DESTINIE a été bâti autour de l'enquête Actifs Financiers 1991 ; il a été récemment mis à jour à l'aide des données de l'enquête Patrimoine 1998. Seules les retraites publiques de base (régime général) des salariés du privé sont étudiées. Dans la population reproduite par le modèle, il n'y a ni agents de l'état, ni artisans-commerçants, ni professions libérales, ni chefs d'entreprise. Ce modèle a depuis engendré une série de travaux particuliers de microéconométrie sur les ménages en matière de comportement de fécondité (Robert-Bobée, 2001), de transferts entre les générations (Bonnet et Mahieu, 1999), de choix de cessation d'activité (Mahieu et Sédillot, 2002) et de carrières salariales (Colin, 1999).

On notera cinq références majeures d'utilisation du modèle à des fins d'évaluation de réformes (Pelé et Ralle, 1997 ; Blanchet et Chanut, 1998 ; Bardaji, Sédillot et Walraet, 2002) ou des propriétés redistributives du système de retraite français (Bonnet et Mahieu, 1999 ; Walraet et Vincent, 2002).

Pelé et Ralle (1997)⁴² utilisent le modèle DESTINIE pour étudier les conséquences de la réforme de 1993, qui a augmenté la durée de cotisation nécessaire à une retraite à taux plein (160 trimestres au lieu de 150), et accru le nombre d'années pris en compte dans le calcul des pensions (25 au lieu de 10). L'analyse porte sur la génération 1960, qui est la première génération à laquelle s'appliquera intégralement la réforme. Les auteurs trouvent que la réforme de 1993 produit des effets importants en termes de report de l'âge de la retraite puisque 35% de l'échantillon repousserait son départ à la retraite. Par ailleurs, selon les situations particulières retenues, la baisse moyenne de la pension du régime général varie entre 3,5% et 16,8%.

Blanchet et Chanut (1998) s'intéressent également à la réforme de 1993. Leurs résultats permettent d'illustrer les conséquences de cette réforme en termes de variation des revenus sur le cycle de vie. Ils ne fournissent pas d'estimations chiffrées des conséquences de la réforme sur les inégalités de cycle de vie. Mais, ils remarquent que la microsimulation met fortement en évidence un effet, souvent peu commenté, qui est que l'allongement de la durée de référence pour le calcul du salaire moyen n'a pas d'effets pour les travailleurs accomplissant une carrière salariale toujours au dessus du plafond de la sécurité sociale. Par

⁴¹ <http://www.lse.ac.uk/Depts/sage/internat.htm>

⁴² Voir aussi Pelé (1997).

ailleurs, ils remarquent que les résultats en termes d'inégalités sont très sensibles aux hypothèses de mobilité salariale.

Bardaji, Sédillot et Walraet (2002) étudient trois réformes du régime général : l'allongement de 2,5 ans de la durée de cotisation ; l'abaissement de l'âge minimal de liquidation à 58 ans ; l'indexation des pensions sur les prix + 0,8%. La microsimulation permet de prévoir sur 40 années les conséquences de ces réformes en termes d'âge moyen à la liquidation, de pension moyenne et durée moyenne validée. Une ventilation des résultats par sexe permet par exemple de montrer que l'abaissement de l'âge minimal de liquidation profite davantage aux hommes qu'aux femmes pour les premières générations étudiées, mais, en raison de l'allongement de la durée des études et d'une participation féminine plus grande, la tendance s'inverse pour les générations suivantes.

Bonnet et Mahieu (1999) étudient l'influence des évolutions démographiques, économiques et institutionnelles sur la redistribution intergénérationnelle inhérente au système par répartition. Les auteurs évaluent en particulier l'évolution des taux de cotisations et des taux de rendement du système de retraite selon les générations et selon un certain nombre d'hypothèses concernant l'évolution future du salaire moyen, du taux de chômage à long terme, du taux de mortalité ou bien encore de l'âge minimal de liquidation. Ils montrent par exemple que si les salaires augmentaient de 3 % par an, le taux de cotisations n'aurait à augmenter que de 3,3 points (sous le plafond) au lieu de 11,5 points dans le scénario central (croissance de 2 % des salaires). De même, une augmentation de l'âge légal de la retraite de 60 à 63 ans (pour les générations nées à partir de 1941) permettrait de réduire le taux de cotisations en 2040 de 5,2 points (en dessous du plafond). Une telle mesure ne modifierait pas le taux de rendement à long terme puisque la baisse des cotisations compenserait la diminution de la durée pendant laquelle les pensions seraient reçues. Mais le taux de rendement serait fortement détérioré pour les générations de transition puisqu'elles paieraient le même taux de cotisation tout en recevant des pensions moins longtemps.

Walraet et Vincent (2002) s'intéressent quant à eux à la question de la redistribution intragénérationnelle. Ils utilisent le modèle DESTINIE pour tester dans quelle mesure le système de retraite français corrige les inégalités, notamment d'opportunité salariale liée à la maternité. Les résultats montrent clairement que la redistribution profite assez largement aux femmes disposant de bas revenus et aux couples ayant élevé des enfants. Les calculs des auteurs donnent également une évaluation des bénéfices supplémentaires pour les femmes qui proviennent du seul fait que le système de retraite ne tient pas compte des différences d'espérance de vie (à ceci près, que dans un couple, le survivant peut prétendre à une pension de réversion).

Ces travaux témoignent de l'intérêt de la microsimulation dynamique pour prévoir le futur et évaluer l'impact de réformes possibles. Elle permet de prendre en considération de façon le plus réaliste possible la structure de la population et des revenus, qu'il est nécessaire de connaître compte tenu des non-linéarités des modes de calcul des retraites dans les systèmes existants. Ce type de recherche illustre le potentiel de la microsimulation pour réfléchir au caractère adapté ou non d'une formule de calcul des retraites.

Une mise en perspective

Dans cet article, nous avons dressé un portrait et un recensement des modèles de microsimulation dynamique. Au regard de cet examen, nous notons, d'un point de vue méthodologique, que les résultats agrégés de simulations dynamiques peuvent revêtir un caractère fortement aléatoire et que les résultats sur données individuelles peuvent présenter de nombreuses imprécisions. Cette spécificité plaide en faveur du développement d'une meilleure connaissance théorique et empirique des propriétés statistiques de la microsimulation. Un tel développement permettrait de mesurer les propriétés globales des modèles ainsi que de construire des indicateurs de fiabilité sur les estimations. Il faut toutefois noter que ce terrain n'est pas vierge, mais les techniques existantes (le *bootstrap* par exemple) sont souvent très coûteuses en ressources informatiques. Néanmoins, l'accroissement des puissances de calcul rendent cet argument de moins en moins pertinent.

La jeunesse de la microsimulation en France fait que la construction des modèles, quels qu'ils soient, repose souvent sur une base de données (de taille modeste en général) qui n'a pas été spécifiquement conçue pour une telle utilisation (par exemple, limitation du champs d'étude aux salariés du privé). Il est alors nécessaire d'utiliser des banques de données annexes pour accroître l'information disponible, ce qui est évidemment source d'aléas. Dans la mesure où l'intérêt de la microsimulation n'est plus à démontrer, il semble souhaitable que les besoins liés à l'emploi de cette méthode soient intégrés dès la conception des bases de données.

Dans l'avenir, on peut penser qu'au sein de chaque pays, l'apparition de plusieurs modèles engendrera une concurrence favorable à leur amélioration par confrontation, mais aussi les différences d'approches devraient permettre de fournir un certain encadrement des évolutions probables. La fiabilité⁴³ croissante des modèles de microsimulation devrait être un argument clé pour voir leur utilisation plus importante dans l'orientation des décisions en matière de finances publiques.

Dans le domaine de l'analyse des déterminants du chômage et, notamment, ceux en rapport avec l'évolution de la structure démographique de la population active, des améliorations sont clairement souhaitables. Un taux de chômage endogène paraît plus crédible. A ce niveau, se pose la question de la connexion de la microsimulation avec les outils standards de la macroéconomie.

Enfin, nous avons noté une certaine pauvreté des indicateurs d'inégalités sur le cycle de vie. A part le recours à un calcul de taux de rentabilité interne de la retraite sur le cycle de

⁴³ Sur ce point, les recommandations de Citro et Hanushek (1997), pour le *U.S. National Research Council*, affichent un scepticisme marqué : « The relevant federal agencies should consider the development of a new integrated individual level microsimulation model for retirement-related policy analysis as an important long-term goal, but construction of such a model would premature until advances are made in data, research knowledge, and computational methods [...] To respond to immediate policy needs, agencies should use limited, special-purpose models with the best available data and research findings to answer specific policy questions. Although such models may not provide very accurate estimates, the alternative of developing complex new individual-level microsimulation or employer models in advance of needed improvements in data and research knowledge has little prospect of producing better results and will likely represent, in the immediate future, a misuse of scarce resource ». Mais la fiabilité future dépend bien sûr des crédits publics alloués dès à présent pour ce type de travaux, et en ce qui concerne l'aspect financier, Harding (1999) cite en exemple deux modèles dont le coût (financement public attribué) a été d'environ 1,5 millions de dollars pour un délai de construction de l'ordre de 5 ans.

vie, les utilisateurs de modèles semblent disposer de peu d'outils de mesure des inégalités entre générations mais aussi au sein même des cohortes. On peut espérer que les travaux de recherche menés dans ce domaine seront capables de fournir des indicateurs globaux éclairants pour les choix publics.

Bibliographie

Ando A. (1996). "Microsimulation Analysis of Aggregate Savings Behaviour in Japan: Effects of Aging Population on Aggregate Savings and Design of Policies on Savings in Japan", miméo, National Institute for Research Advancement, Tokyo.

Ando A., Brandolini A., Bruno G., Cannari L., Cipollone P., D'Alessio G., Faiella I., Forni L., Marino M. R. et Nicoletti Altomari S. (2000). "The Bank of Italy's DYNAMITE: Recent Developments", miméo, Rome: Bank of Italy.

Ando A. et Modigliani F. (1957). "Tests of the Life Cycle Hypothesis of Saving: Comments and Suggestions", *Bulletin of the Oxford Institute of Statistics*.

Ando A. et Nicoletti Altomari S. (1999). "A Microsimulation Model of the Italian Household Sector", miméo, University of Pennsylvania.

Andreassen L., Fredriksen D. et Ljones O. (1996). "The Future Burden of Public Pension Benefits: A Microsimulation Study", in A. Harding (Ed.).

Antcliff S. (1993). "An Introduction to DYNAMOD: A Dynamic microsimulation", miméo, NATSEM, Canberra.

Baekgaard H. (2002). "Micro-Macro Linkage and the Alignment of Transition Processes", *Technical Paper n°5*, Natsem.

Baldini J. (1997). "Diseguaglianza e Redistribuzione nel Ciclo di Vita", Bologna: Il Mulino.

Bardaji J., Sédillot B. et Walraet E. (2002). "Evaluation de trois réformes du Régime Général d'assurance vieillesse à l'aide du modèle de microsimulation DESTINIE", *Document de travail de la DESE*, G2002/07, Insee.

Beffy P.-O., Deroyon J., Fourcade N., Gregoir S., Laïb N. et Monfort B. (2003). "Evolutions démographiques et croissance : une projection macro-économique à l'horizon 2020", *Document de travail de la DESE*, G 2003/04, Insee.

Bergmann B., Eliasson G. et Orcutt G. H. (Eds) (1980). *Micro Simulation Models: Methods and Applications*, IUI Conference Report, Industrial Institute for Economic and Social Research, Stockholm.

Barro R. J. (1974). "Are Government Bond Net Wealth?", *Journal of Political Economy*, 82 (6).

Blanchet D. (1992). "Retraite et croissance économique : un essai de simulation", *Economie et prévision*, n°105, 1-16.

Blanchet D. et Chanut J-M. (1997). "Microsimulation des retraites : Annexe technique", miméo, Insee.

Blanchet D. et Chanut J-M. (1998). "Les retraites individuelles à long terme : une projection par microsimulation", *Economie et Statistique*, n°315.

Bommier A., Magnac T. et Roger M. (2001). "Départs en retraite : évolutions récentes et modèles économiques", *Revue Française d'Economie*, 26.

Bonnet C. et Mahieu R. (1999). "Microsimulation Techniques Applied to Intergenerational Transfers", *Document de travail de la DESE*, G9906, Insee.

Bourguignon F. (1981). "Pareto Superiority of Unequalitarian Equilibria in Stiglitz' Model of Wealth Distribution with Convex Saving Function", *Econometrica*, 49 (6).

Bracher M. (1996). "Preliminary Demographic Parameters for the DYNAMOD Microsimulation Model", miméo, Natsem.

Breuil-Genier P. (1998). "Les enseignements théoriques et pratiques des microsimulations en économie de la santé", *Economie et Statistique*, n°315.

Briard K. (2003). "Typologie des carrières des retraités de droit propre du régime général et construction de profils-types", *Retraite et Société*, n°40.

Butrica B., Iams H., Moore J. et Waid M. (2001). "Methods in Modelling Income in the Near Term (MINT I)", *Working paper ORES n°91*, SSA.

Caldwell S. B. (1996). "Health, Wealth, Pensions and Life Paths: The CORSIM Dynamic Microsimulation Model", in A. Harding (Ed).

Caldwell S. B. et Morrisson R. (2000). "Validation of longitudinal microsimulation models: experience with CORSIM and DYNACAN", in MITTON *et alii* (Eds).

Castiglione S. (2001). *Microsimulation of Complex System Dynamics. Automata Models in Biology and Finance*, Inaugural-Dissertation, Universität zu Köln.

Chagny O, Dupont G., Sterdyniak H. et Veroni P. (2001). "Les réformes des systèmes de retraites en Europe", *Revue de l'OFCE*, n°78.

Charpin J.-M. (1999). *L'avenir de nos retraites*, Rapport au premier ministre, Documentation Française.

Chauvel L. (1998). *Le destin des générations*, Paris : PUF.

Chauvin V. et Plane M. (2001). "2000-2040 : Population active et croissance", *Revue de l'OFCE*, 79.

Citro C. F. et Hanushek E. A. (Eds) (1997). *Assessing Policies for Retirement Income: Needs for Data, Research, and Models*, Washington, DC: National Academy Press.

Colin C. (1999). "Modélisation des carrières salariales dans DESTINIE", *Document de travail de la DESE*, G9902, Insee.

Conseil Economique et Social (2000). *L'avenir des systèmes de retraite*, Rapport Teulade, http://www.conseil-economique-et-social.fr/ces_dat2/2-3based/base.htm.

Conseil d'Orientation des Retraites (2002). *Retraites: renouveler le contrat social entre les générations*, La documentation française.

Corsim (2001). "CORSIM 4.0 Analyst Documentation", Strategic Forecasting, Ithaca, New York..

Cornilleau G. et Sterdyniak H. (1995). "Les retraites en France, des débats théoriques aux choix politiques" in B. Cochemé et F. Legros (sous la direction de), *Les retraites, genèse, acteurs, enjeux*, Armand-Colin.

Cornilleau G. et Sterdyniak H. (2003). "Réforme des retraites : tout miser sur l'allongement de la durée de cotisation?", *Lettre de l'OFCE*, n°237.

Cottave R. (1991). *Livre blanc sur les retraites*, Rapport du CGP, Documentation Française.

Den Haan W. J. (1997). "Solving Dynamic Models with Aggregates Shocks and Heterogeneous Agents", miméo, University of California at San Diego.

Division Redistribution et Politiques Sociales (1999). "Le modèle de microsimulation dynamique DESTINIE", *Document de travail de la DESE*, G9913, Insee.

Dupont G. et Sterdyniak H. (2000). *Quel avenir pour nos retraites?*, La Découverte.

Dupuis J.-M. et El Moudden C. (2002). "Transformations du marché du travail et droits à pension, quelles perspectives ?", article présenté au colloque IRES-ENRSP, octobre.

Ericson P. et Hussenius J. (1998). "A note about SESIM – a dynamic microsimulation model", miméo, Ministry of finance, Suède.

Ericson P. et Hussenius J. (1999). "Distributional Effects of Public Student Grants in Sweden – a Presentation and an Application of the Dynamic Microsimulation Model SESIM", miméo, Ministry of finance, Suède.

Falkingham J. et Johnson P. (1995). "A Unified Funded Pension Scheme (UFPS) for Britain", in Falkingham and Hills (Eds), *The Dynamic of Welfare: The Welfare State and the Life Cycle*, Prentice-Hall.

Feldstein M. S. (1974). "Social Security, Induced Retirement, and Aggregate Capital Accumulation", *Journal of Political Economy*, 82 (4).

Fleurbaey M. et Michel P. (1992). "Quelle justice pour les retraites ?", *Revue d'économie financière*, 23.

Folster S. (1997). "Social Insurance Based on Personal Savings Accounts", *European Economy*, No. 1997/4.

Fredriksen D. (1998). *Projections of Population, Education, Labour Supply and Public Pension Benefits: Analyses with the Dynamic Microsimulation Model MOSART*, Social and Economic Studies, Statistics Norway.

Friedman M. (1957). *A Theory of the Consumption Function*, New-York.

Galler H. P. (1997). "Discrete-Time and Continuous-Time Approaches to Dynamic Microsimulation Reconsidered", *NATSEM Technical Paper*, n°13.

Galler H. P. et Wagner G. (1986). "The Microsimulation Model of the Sbf3 for the analysis of Economic and Social Policies", in G. H. Orcutt *et alii* (Eds).

Hain W. et Hellberger C. (1986). "Longitudinal Simulation of Lifetime Income", in Orcutt G. *et alii* (Eds).

Hairault J.-O. et Langot F. (2002). "Accumulation, Inégalités et Réforme des retraites : une modélisation markovienne à agents hétérogènes", miméo, Cepremap.

Hairault J.-O., Langot F. et Sopraseuth T. (2003). "Eléments d'évaluation de la réforme des retraites : le cas d'une date de départ choisie", miméo, Cepremap.

Hancock R., Mallender J. et Pudney S. (1992). "Constructing a Computer Model for Simulating the Future Distribution of Pensioner's Incomes for Great Britain", in Hancock R. et Sutherland H. (Eds).

Hancock R. et Sutherland H. (Eds) (1992). *Microsimulation Models for Public Policy Analysis: New Frontiers*, Sticerd, LSE, London.

Harding A. (Ed.) (1993). *Lifetime Income Distribution and Redistribution: Applications of a Microsimulation Model*, North-Holland.

Harding A. (Ed.) (1996). *Microsimulation and Public Policy*, North-Holland-Elsevier.

Harding A. (1999). "Discussion on "Ageing and Social Expenditure" ", in *Proceedings of the Policy Implications of the Ageing of Australia's Population Conference*, Productivity Commission.

Hendriks G. et Holden R. (1976a). *A Report on the Variance in Simulated Earnings Histories produced by DYNASIM*, Urban Institute.

Hendriks G. et Holden R. (1976b). "The Role of Microanalytic Simulation Models in Projecting OASDI costs", *Proceedings of the American Statistical Association*.

Hénin P.-Y. et Weitzenblum T. (2003). "Eléments d'évaluation de la réforme des retraites : coûts et avantages des stratégies de transition", miméo, Cepremap.

Holm E., Holme K., Mäkilä K., Mattsson-Kaupi M., Mörvik G. (2002). "The SVERIGE spatial microsimulation model : content, validation and example applications", miméo, Spatial Modelling Center.

Holmer M., Janney A. et Cohen B. (2001). "Overview of PENSIM/2", miméo, Washington D.C.: Policy Simulation Group.

Insee (1999). "Le modèle de microsimulation dynamique DESTINIE", *Document de travail de la DESE*, G 9913, Insee.

Joyeux C., Plasman R. et Scholtus B. (1996). "A Model of the Evolution of Pensions Expenditures in Belgium", *Meeting of the European HCM Network on Socio-Economic Modelling*, Cambridge.

Kennell D. L. et Sheils J. F. (1990). "PRISM: Dynamic Simulation of Pension and Retirement Income", in G. H. Lewis et R. C. Michel (Eds).

King A., Baekgaard H. et Robinson M. (1999a). "DYNAMOD-2: An Overview", NATSEM Technical Paper n°19, University of Canberra, Australie.

King A., Baekgaard H. et Robinson M. (1999b). "The Base Data for Dynamod-2", NATSEM Technical Paper n°20, University of Canberra, Australie.

Klevmarken N. A. (1997). "Behavioral Modeling in Micro Simulation Models. A Survey", *Working Paper 1997-31*, Department of Economics, Uppsala University.

Klevmarken N. A. (2002a). "Statistical Inference in Microsimulation Models: Incorporating External Information", *Mathematics and Computers in Simulation*, 59 (1-3).

Klevmarken N. A. (2002b). "Microsimulation – a Tool for Economic Analysis", miméo, Uppsala University.

Klevmarken N. A. et P. Olovsson (1996). "Direct and behavioral effects of income tax changes - simulations with the Swedish model MICROHUS", in Ann Harding (Ed.).

Kotlikoff L. J., 1995, "Privatization of Social Security: How It Works and Why it Matters", *NBER Working Paper*, n°5330.

Lagarde F., Launay J.-P. et Lenormand F. (1982). "Les effets redistributifs du système de retraites : une méthode, un constat et des voies de réformes", *Droit social*, 6.

Le Cacheux J. et Touzé V. (2002). "Les modèles d'équilibre général calculable à générations imbriquées : enjeux, méthodes et résultats", *Revue de l'OFCE*, 80, janvier.

Le Cacheux J. et Touzé V. (2003). "Vieillesse et richesse des nations", *Revue de l'OFCE*, 86, juillet.

Legris B. et Lollivier S. (1996). "Le niveau de vie par génération", *INSEE Première*, n°423.

Lewis G. H. et Michel R. C. (Eds) (1990). *Microsimulation Techniques for Tax and Transfer Analysis*, The Urban Institute Press.

Lucas R. E. Jr. (1976). "Econometric Policy Evaluation: A Critique", *Carnegie-Rochester Conference Series on Public Policy*, 1.

Lucas D.J. (1994). "Asset Pricing with Undiversifiable Income Risk and Short Sales Constraints: Deepening the Equity Premium Puzzle", *Journal of Monetary Economics*, 34.

Milton L., Sutherland H. et Weeks M. (Eds) (2000). *Microsimulation Modelling for Policy Analysis: Challenges and Innovations*, Cambridge University Press.

Nelissen J. H. .M. (1993). "The Redistributive Impact of Social Security Schemes on Lifetime Labour Income", KUB, Pays-Bas.

Nelissen J. H. .M. (2000). "The Costs and the Lifetime Redistributive Impact of Capital Funded Dutch Supplementary Pensions", *Geneva Papers on Risk and Insurance*, 25.

O'Donoghue C. (1999). "Dynamic Microsimulation: A Methodological Survey", *Brazilian Electronic Economic Journal*.

O'Donoghue C. (2000). "A Dynamic Microsimulation Model for Ireland", University of Cambridge, LSE, miméo.

Orcutt G. (1957). "A new type of economic status", *Review of Economics and Statistics* 58, 773-797..

Orcutt G., Caldwell S. et Wertheimer R. (1976). *Policy Exploration through Microanalytical Simulation*, The Urban Institute.

Orcutt G. et Glazer A. (1980). "Microanalytic Modelling and Simulation", in Bergmann B. *et alii* (eds).

Orcutt G., Merz J. et Qunike H. (Eds) (1986). *Microanalytical Simulation Models to Support Social and Financial Policy*, North-Holland.

Papoulis A. (1965). *Probability, Random Variables, and Stochastic Processes*, Series in Systems Science, McGraw-Hill.

Pelé L.-P. (1997). "Les conséquences de la réforme des retraites de 1993 : une approche par microsimulation", miméo, division Redistribution et Politiques Sociales, INSEE.

Pelé L.-P. et Ralle P. (1997). "Age de la retraite : les aspects incitatifs du régime général", *Document de travail de la DESE*, G97-18, Insee

Pelé L.-P. et Ralle P. (1999). "Les choix de l'âge de la retraite : aspects incitatifs des règles du régime général et effets de la réforme de 1993", *Economie et Prévision*, n°138-139.

Pudney S. et Sutherland H. (1996). "Statistical Reliability in Microsimulation Models with Econometrically-Estimated Behavioural Responses", chap. 21, in A. Harding (Ed).

Raynaud E. (2003). "Les retraités polypensionnés des secteurs public et privé: une analyse par cas types", *Études et Résultats*, n° 236, DREES.

Robert-Bobee I. (2001). "Modelling Demographic Behaviors in the French Microsimulation Model Destinie: An Analysis of Future Change in Completed Fertility", *Document de travail de la DESE*, G2001-14, Insee.

Rowe M. et Wolfson G. (2000). "Public Pension Reforms – Analyses Based on the LifePaths Generation Accounting Framework", miméo.

Stiglitz J. E. (1969). "Distribution of Income and Wealth Among Individuals", *Econometrica*, 37.

Stroombergen A., Rose D. et Miller J. (1995). "Wealth Accumulation and Distribution: Analysis with a Dynamic Microsimulation Model", Wellington: Business and Economic Research Ltd.

Sutherland H. (1998). "Les modèles statiques de microsimulation en Europe dans les années 90", *Economie et Statistique*, 315.

Van De Ven J. (1998). "A Dynamic Cohort Microsimulation Model", University of Melbourne, Department of Economics, Research Paper, 637.

Walraet E. et Vincent A. (2002). "Intragenerational Distributional Analysis in the French Private Sector Pension Scheme", mimeo, Insee et Delta.

Werttheimer R., Zedlewski S. R., Anderson J. et Moore K. (1986). "DYNASIM in Comparison with Other Microsimulation Models", in G. Orcutt *et alii* (Eds).

Wolf D. A. (2000a). "Note on Statistical Analysis and Microsimulation for Studying Living Arrangements and Intergenerational Transfers", miméo, Center of Policy Research, Syracuse University.

Wolf D. A. (2000b). "The Role of Microsimulation in Longitudinal Data Analysis", miméo, Center of Policy Research, Syracuse University.

Wolfson M. C. (1988). "Homemaker Pensions and lifetime redistribution", *Review of Income and Wealth*, Vol. 34, No. 3.

Zaidi A. et Rake K. (2001). "Dynamic Microsimulation Models: A Review and Some Lessons for SAGE", SAGE – London School of Economics, working paper.

Zweimuller J., Winter-Ebmer R. et Falkinger J. (1996). "Retirement of Spouses and Social Security Reform", *European Economic Review*, 40.

ANNEXE TECHNIQUE - Etude des propriétés de l'estimateur de la moyenne de deux processus stochastiques élémentaires.

Les informations recherchées dans l'emploi d'un modèle de microsimulation sont de deux types : d'un point de vue transversal, on cherche à connaître les distributions des valeurs prises par les différentes variables ; d'un point de vue longitudinal, on veut également observer les différentes trajectoires individuelles et évaluer leurs distributions. Pour cela, les simulations permettent de disposer d'estimateurs (par exemple, une moyenne) des différents moments des distributions (moyenne, écart-type, etc). Par essence, ces estimateurs sont des variables aléatoires soumises à un processus stochastique. Une question d'importance est de savoir à partir de quelle taille d'échantillons dispose-t-on de suffisamment bons estimateurs des différents moments des distributions longitudinales et transversales. *A priori*, il semble que la précision (faible variance) de ces estimateurs est d'autant plus faible que l'horizon est long. Dans cette annexe, nous tentons d'approfondir cet aspect dans le cadre de deux processus stochastiques élémentaires.

Dans les deux exemples qui suivent, on étudie des processus discret et continu. La variable discrète comprend seulement deux états : il s'agit d'un processus binaire. On suppose aussi que les deux processus sont univariés –ils peuvent donc s'étudier séparément–, ils sont parfaitement identifiés –il n'y a pas de problème de spécification, ni d'estimation– et les échantillons initiaux sont également représentatifs de la population.

Etude du processus stochastique d'une variable discrète

Considérons le processus stochastique d'une variable discrète caractérisée par deux états 1 et 0 (par exemple, 1=emploi ; 0=chômage). On note p_0 (resp. $1-p_0$) la proportion d'individus dans l'état 1 (resp. 0). Dans le cadre de ce processus discret, un individu initialement dans un état 1 (resp. 0) a une probabilité u (resp. $1-v$) d'y rester. Intéressons-nous ici seulement à l'évolution du premier moment de la distribution⁴⁴, c'est-à-dire la moyenne empirique μ_t d'un échantillon constitué de N individus. Si on note x_{it} les évolutions individuelles, la moyenne simulée $\mu_t = \frac{\sum_{i=1}^N x_{it}}{N}$, au début de la simulation, est un estimateur sans biais de la moyenne théorique qui est égale à p_t : $E_0(\mu_t) = p_t$. Ce résultat provient du fait qu'on suppose que l'échantillon initial a été choisi de façon à reproduire parfaitement la moyenne théorique : il n'y a pas de biais initial d'échantillonnage.

En revanche, à chaque période de la simulation, l'estimateur à une période de la moyenne est biaisée (biais d'échantillonnage) puisqu'on observe :

$$E_t(\mu_{t+1}) = u \cdot \mu_t + v \cdot (1 - \mu_t) \neq p_{t+1} \quad (1)$$

sauf si $\mu_t = p_t$ (égalité supposée vraie pour $t=0$). Quant à la variance de l'estimateur à l'horizon d'une période, elle évolue comme suit :

⁴⁴ Bien entendu, cette démarche pourrait s'appliquer pour les autres moments de la distribution.

$$V_t(\mu_{t+1}) = \frac{\mu_t \cdot u \cdot (1-u) + (1-\mu_t) \cdot v \cdot (1-v)}{N} \quad (2)$$

En jouant sur la propriété de l'opérateur espérance $E_t(f(\mu_{t+2})) = E_t(E_{t+1}(f(\mu_{t+2})))$ où f est une fonction continue quelconque, on détermine la variance pour un horizon de deux périodes :

$$V_t(\mu_{t+2}) = E_t(V_{t+1}(\mu_{t+2})) + V_t(E_{t+1}(\mu_{t+2})) \quad (3)$$

En itérant d'une période, on trouve :

$$V_t(\mu_{t+3}) = E_t(V_{t+2}(\mu_{t+3})) + V_t(E_{t+2}(\mu_{t+3})) \quad (4)$$

Compte tenu de l'espérance de la moyenne empirique à l'horizon d'une période, on déduit :

$$V_t(\mu_{t+2}) = E_t(V_{t+1}(\mu_{t+2})) + (u-v)^2 \cdot V_t(\mu_{t+1}) \quad (3')$$

et

$$V_t(\mu_{t+3}) = E_t(V_{t+2}(\mu_{t+3})) + (u-v)^2 \cdot V_t(\mu_{t+2}) \quad (4')$$

Dans le cas particulier où $u(1-u) = v(1-v) = \sigma^2$, on trouve $V_t(\mu_{t+1}) = \frac{\sigma^2}{N}$, et par récurrence, on déduit la variance à l'instant 0 de l'estimateur de la moyenne à un horizon T :

$$V_0(\mu_T) = \frac{\sigma^2}{N} \left(1 + (u-v)^2 + \dots + (u-v)^{2(T-1)} \right) \quad (5)$$

La variance de l'estimateur de la moyenne est une fonction croissante de l'horizon de prévision, et elle admet une limite finie⁴⁵ :

$$V_0(\mu_\infty) = \frac{1}{N} \cdot \frac{\sigma^2}{1-(u-v)^2} \quad (6)$$

L'évolution des distributions théoriques est donnée par l'équation dynamique matricielle suivante :

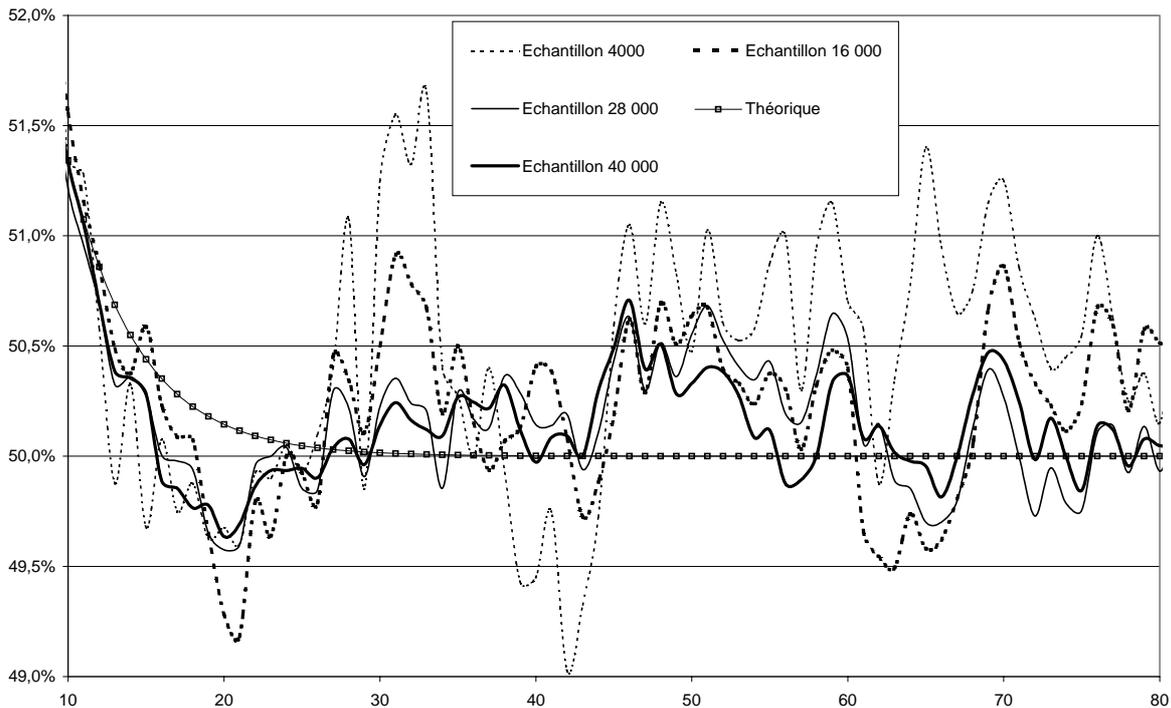
$$\begin{bmatrix} p_t \\ 1-p_t \end{bmatrix} = \begin{bmatrix} u & v \\ 1-u & 1-v \end{bmatrix} \cdot \begin{bmatrix} p_{t-1} \\ 1-p_{t-1} \end{bmatrix} \quad (7)$$

Le graphique 1 trace l'évolution de la moyenne estimée dans le cadre d'une série de simulations stochastiques pour des valeurs numériques : $u = 1-v = 0,9$ et $\mu_0 = p_0 = 0,6$. La valeur théorique est également calculée. Le graphique 1 montre que des fluctuations

⁴⁵ Cette condition renvoie à une propriété sur les valeurs propres de la matrice de transition.

marginales demeurent, même pour des tailles importantes d'échantillon. Avec ces valeurs numériques, l'écart-type du processus stochastique individuel est 0,3, la valeur stationnaire de la moyenne stationnaire est 50%, et la fourchette de précision à 2 écarts-types) de l'estimation de la moyenne de long terme est + ou - 0,6 point pour $N = 40\ 000$ contre + ou - 2 points avec $N = 4\ 000$.

**Graphique 1 – Processus stochastique d'une variable discrète :
Exemples de moyennes simulées avec différentes tailles
d'échantillon sur 80 périodes**



Pour des raisons pratiques de communication de résultats, on peut comprendre que le caractère artificiellement fluctuant de la moyenne empirique (par exemple, l'espérance de vie) n'est pas satisfaisant. En comparaison, même si une prévision macroéconomique est une variable aléatoire, cette dernière a en général un caractère lisse, surtout lorsqu'il s'agit de projections à long terme. De plus, pour des raisons de cohérence avec les pronostics macroéconomiques et démographiques, il peut être intéressant d'aligner les résultats issus de la microsimulation sur le scénario macroéconomique : par exemple, le taux de chômage simulé sur le taux de chômage tendanciel. Pour éviter ce caractère fluctuant de la moyenne simulée, les concepteurs⁴⁶ de modèles ont la possibilité de forcer les simulations en instaurant des tirages aléatoires sans remise. Cette façon de procéder n'est pas nécessairement incohérente avec la réalité économique : lorsqu'on observe la distribution d'une caractéristique dans la population, cette distribution peut autant résulter d'un tirage sans remise que de d'un tirage avec remise. Ainsi, une probabilité de 90% de survenance d'un évènement donné pour une population de 1000 personnes⁴⁷ se traduit par l'existence de 900

⁴⁶ Baekaard (2002) présente, outre la méthode de tirage sans remise, une méthode d'alignement qui consiste à réviser les poids de chaque individu de façon à reproduire les agrégats macroéconomiques.

⁴⁷ Par exemple, pour un échantillon de 60.000 concernant 60 générations, il faut compter 1000 individus en moyenne par cohorte.

opportunités de se trouver dans cette situation. La probabilité *ex ante* est bien 10%, mais au fur et à mesure de la simulation, les probabilités individuelles d'échec ou de succès ne sont pas constantes mais fluctuent⁴⁸. Toutefois, même si la moyenne ne répond pas à un processus aléatoire, les autres moments empiriques, qui ne sont pas abordés ici, demeurent des processus aléatoires. Il n'est donc, *a priori*, pas facile de contrôler l'ensemble des moments reproduits par la simulation.

Par ailleurs, le principal attrait d'un modèle de microsimulation dynamique est de conserver l'ensemble des informations concernant les trajectoires individuelles. L'exercice permet ainsi d'estimer l'évolution de la distribution des différents profils intertemporels. Se pose alors la question de la fiabilité des estimateurs de cette distribution. Si T est l'horizon de la simulation, le nombre exact de trajectoires possibles est 2^{T+1} . L'objet de la simulation est de fournir une estimation de la distribution de ces trajectoires. *A priori*, la qualité des estimateurs de la distribution des trajectoires sera d'autant plus mauvaise que le nombre de trajectoires potentielles sera grand. Par exemple, si l'horizon de simulation est 40 périodes : le nombre total de trajectoires est supérieur à 1000 milliards⁴⁹. Il est impossible de se soucier de toutes ces trajectoires. La microsimulation élimine donc beaucoup de trajectoires possibles⁵⁰, mais elle fournit *a priori* celles qui sont les plus pertinentes à étudier. Un modèle de microsimulation peut, de ce point de vue, être considéré comme un outil de sélection (échantillonnage) de trajectoires « intéressantes » (qui deviennent des trajectoires « types ») et chaque trajectoire ainsi choisie a une importance égale au poids de l'individu considéré dans la population. En comparaison avec les cas-types, le choix aléatoire d'une trajectoire a le mérite de la neutralité⁵¹.

Etude du processus stochastique d'une variable continue

Intéressons-nous au processus stochastique d'une variable continue (par exemple, l'évolution log-normalisée, centrée, réduite, et stationnarisée du salaire d'individus appartenant à la même classe d'âge et ayant la même qualification professionnelle) du type :

$$x_{i,t} = x_{i,t-1} + \varepsilon_{it} \quad (8)$$

où ε_{it} est un bruit blanc i.i.d. suivant une loi normale centrée et de variance σ^2 . Tous les processus individuels $x_{i,0}$ sont égaux et initialisés à 0. L'espérance théorique de ce processus à l'instant 0 (début de la simulation) est donc nulle. De même que dans l'exemple précédent, intéressons-nous au premier moment de la distribution des caractéristiques individuelles $x_{i,t}$,

c'est-à-dire la moyenne $\mu_t = \frac{\sum_{i=1}^N x_{it}}{N}$. On montre facilement que la moyenne théorique estimée

à l'instant 0 du processus à l'horizon T est $E_0(\mu_T) = 0$ tandis que l'écart-type à l'instant 0 de

la moyenne empirique à l'horizon T est $\sigma_T = \sqrt{V_0(\mu_T)} = \sqrt{\frac{T}{N}} \cdot \sigma$. On trouve une propriété

⁴⁸ D'ailleurs, lorsqu'on observe une certaine constance dans les processus de transition pour un groupe type, il n'est pas possible de savoir si les proportions observées s'appliquent de façon constante.

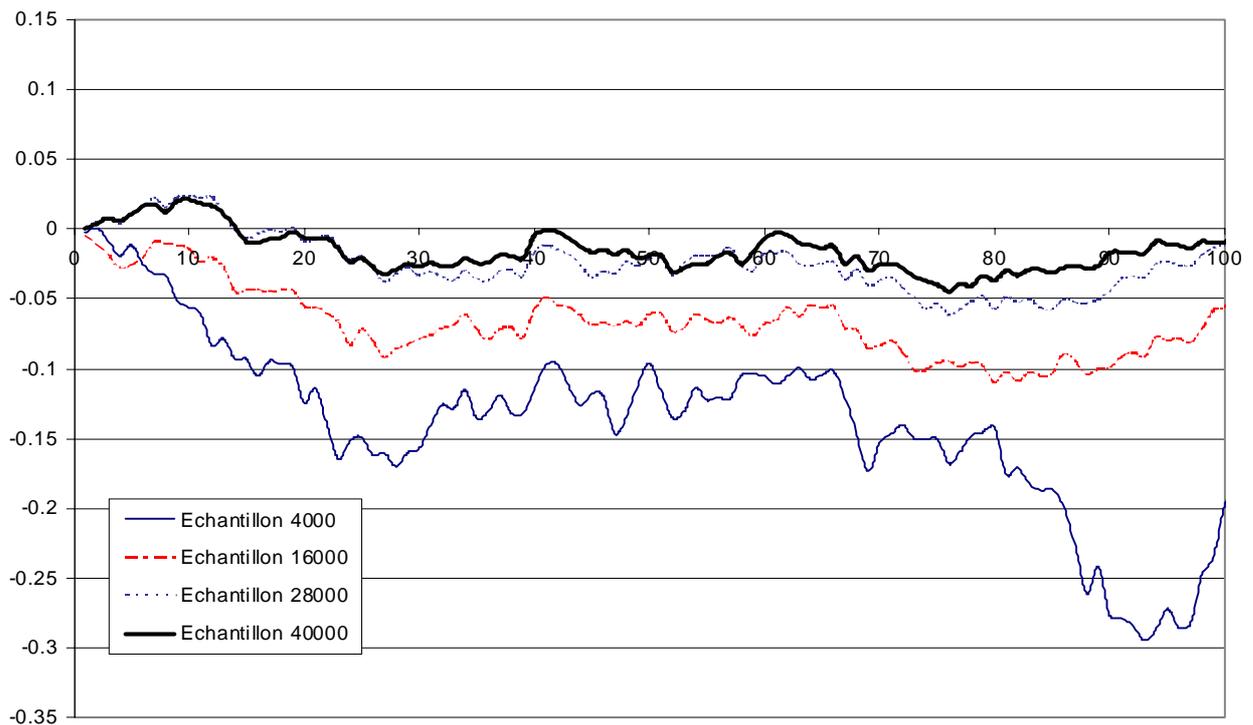
⁴⁹ Bien sûr, ce chiffre peut paraître considérable en comparaison à la taille de la population d'un pays.

⁵⁰ Dans une certaine mesure, la microsimulation sélectionne un certain nombre de « cas types » auxquels elle associe une distribution représentative (estimateur sans biais) de la réalité probable.

⁵¹ La non-neutralité dans le choix est également intéressante (*cf.* l'intérêt de construire un cas-type).

similaire au cas précédent: la fourchette de précision de l'estimateur augmente avec l'horizon de la simulation et est une fonction décroissante de la taille de l'échantillon. Le graphique 2 illustre le caractère aléatoire de la moyenne empirique pour une valeur numérique $\sigma^2 = 1$. Celle-ci suit, en effet, un mouvement brownien. Des moyennes empiriques ont été estimées dans plusieurs configurations de tailles d'échantillon. On s'aperçoit que même pour des échantillons de grande taille ($N = 40\ 000$), le caractère fluctuant n'est pas négligeable.

**Graphique 2 – Processus stochastique d'une variable continue :
Exemples de moyennes simulées avec différentes tailles
d'échantillon sur 100 périodes (moyenne théorique = 0)**



Dans le cadre d'un processus continu, une manière simple de calculer avec exactitude la moyenne est de procéder à une simulation sans tirage aléatoire, c'est-à-dire en choisissant comme valeurs simulées l'espérance des processus individuels. Par la suite, il est toujours possible de réviser la réalisation simulée avec choc pour chaque individu de façon à ce que les valeurs obtenues coïncident en moyenne à l'espérance théorique.

Dans le cas d'une variable continue, la notion de nombre de trajectoires possibles n'a pas de sens car l'ensemble concerné est un ensemble de *continuum* (\mathbb{R}^T). Cette notion peut s'approcher en discrétisant la variable. Par exemple si le choix de l'unité de discrétisation des revenus est le décile, dix groupes de revenus sont observés. Dans cette configuration, le nombre à recenser de trajectoires possibles sur une durée d'activité de 40 années est gigantesque en comparaison du cas binaire précédent.