Market Stability vs. Market Resilience: Regulatory Policies Experiments in an Agent-Based Model with Low- and High-Frequency Trading

Sandrine Jacob Leal
CEREFIGE - ICN Business School (Nancy-Metz) France

Mauro Napoletano
OFCE-Sciences Po and SKEMA Business School

April 2016
Market Stability vs. Market Resilience: Regulatory Policies Experiments in an Agent-Based Model with Low- and High-Frequency Trading

Sandrine Jacob Leal∗ Mauro Napoletano†

April 1, 2016

Abstract
We investigate the effects of different regulatory policies directed towards high-frequency trading (HFT) through an agent-based model of a limit order book able to generate flash crashes as the result of the interactions between low- and high-frequency (HF) traders. We analyze the impact of the imposition of minimum resting times, of circuit breakers (both ex-post and ex-ante types), of cancellation fees and of transaction taxes on asset price volatility and on the occurrence and duration of flash crashes. In the model, low-frequency agents adopt trading rules based on chronological time and can switch between fundamentalist and chartist strategies. In contrast, high-frequency traders activation is event-driven and depends on price fluctuations. In addition, high-frequency traders employ low-latency directional strategies that exploit market information and they can cancel their orders depending on expected profits. Monte-Carlo simulations reveal that reducing HF order cancellation, via minimum resting times or cancellation fees, or discouraging HFT via financial transaction taxes, reduces market volatility and the frequency of flash crashes. However, these policies also imply a longer duration of flash crashes. Furthermore, the introduction of an ex-ante circuit breaker markedly reduces price volatility and removes flash crashes. In contrast, ex-post circuit breakers do not affect market volatility and they increase the duration of flash crashes. Our results show that HFT-targeted policies face a trade-off between market stability and resilience. Policies that reduce volatility and the incidence of flash crashes also imply a reduced ability of the market to quickly recover from a crash. The dual role of HFT, as both a cause of the flash crash and a fundamental actor in the post-crash recovery underlies the above trade-off.

Keywords: High-frequency trading, Flash crashes, Regulatory policies, Agent-based models, Limit order book, Market volatility.

JEL codes: G12, G01, C63.

∗CEREFIGE - ICN Business School (Nancy-Metz) France. Address: ICN Business School (Nancy-Metz) 13, rue Michel Ney, 54000 Nancy (France). Tel:+33 383173776. Fax:+33 383173080. E-mail address: sandrine.jacob-leal@icn-groupe.fr
†Corresponding author. OFCE-Sciences Po, Skema Business School, Sophia-Antipolis (France), and Institute of Economics, Scuola Superiore Sant’Anna, Pisa (Italy). Address: 60, rue Dostoievski BP 85, 06902 Sophia Antipolis Cedex (France) Tel: +33 (0)4 89 73 71 06. E-mail address: mauro.napoletano@sciencespo.fr

1
1 Introduction

This paper studies the effects of a set of regulatory policies aimed at curbing the possible negative effects of high-frequency trading (HFT henceforth), and at reducing market volatility and the occurrence of flash crashes.

Over the past decade, HFT has sharply increased in US and European markets (e.g., AMF, 2010; SEC, 2010; Lin, 2012, and references therein). HFT also represents a major challenge for regulatory authorities, partly because it encompasses a wide array of algorithmic trading strategies and partly because of the big uncertainty yet surrounding the net benefits it has for financial markets. Indeed, on the one hand, some studies have highlighted the benefits of HFT as a source of an almost continuous flow of liquidity (see e.g., Brogaard, 2010; Menkveld, 2013). On the other hand, other works (see e.g., SEC, 2010; Angel, Harris, and Spatt, 2011; Lin, 2012; Kirilenko and Lo, 2013) have pointed to HFT as a source of higher volatility in markets and as a key driver in the generation of extreme events like flash crashes, whose incidence has grown in the last decades (Johnson, Zhao, Hunsader, Meng, Ravindar, Carran, and Tivnan, 2012; Golub, Keane, and Poon, 2012). The regulatory framework is complicated by the fact that - although many explanations have so far been proposed for flash crashes - no consensus has yet emerged about the fundamental causes of these extreme phenomena (see Haldane, 2011). Overall, these open issues leave policy-makers empty-handed about possible policies that could be used to mitigate the negative effects of HFT without affecting its benefits, and about policies that could prevent flash crashes and/or dampen their impact on markets.

Earlier empirical and theoretical works have already attempted to study the effect of different sets of regulatory measures (e.g., Westerhoff, 2008; Brewer, Cvitanic, and Plott, 2013; Vuorenmaa and Wang, 2014) and of some specific regulation policies such as financial transaction tax (Colliard and Hoffmann, 2013; Fricke and Lux, 2015; Lavicka, Lichard, and Novotny, 2014), minimum resting times (Hayes, Paddrik, Todd, Yang, Beling, and Scherer, 2012), market design (Budish, Cramton, and Shim, 2015), cancellation fee (Friederich and Payne, 2015), position limits (Lee, Cheng, and Koh, 2011). However, these works have either not considered the role of HFT (e.g., Westerhoff, 2008), or they have treated flash crashes as resulting from an exogenous shock (e.g., Brewer, Cvitanic, and Plott, 2013) or, finally, they have only focused on a very narrow set of policies (e.g., Hayes, Paddrik, Todd, Yang, Beling, and Scherer, 2012; Vuorenmaa and Wang, 2014).

On these grounds, we contribute to the current debate about the regulatory responses to flash crashes, and to the potential negative externalities, of HFT by studying the impact of a set of policy measures in an agent-based model where flash crashes en-
dogenously emerge out of the interplay between low- and high-frequency traders. The goal is to shed some light on which policy measures are effective to curb volatility and the incidence of flash crashes and/or to fasten the process of price-recovery after a crash. To this end, we extend the ABM developed in Jacob Leal, Napoletano, Roventini, and Fagiolo (2016) to allow for endogenous orders’ cancellation by HF traders, and we then use the model as a test-bed for a number of policy interventions directed towards HFT. This model is particularly well-suited and relevant in this case because, differently from existing works (e.g., Brewer, Cvitanic, and Plott, 2013), it is able to endogenously generate flash crashes as the result of the interactions between low- and high-frequency (HF) traders. Moreover, compared to the existing literature we consider a broader set of policies, also of various nature. The list includes market design policies (circuit breakers) as well as command-and-control (minimum-resting times) and market-based (cancellation fees, financial transaction tax) measures.

The model in Jacob Leal, Napoletano, Roventini, and Fagiolo (2016) portrays a market wherein LF agents trade a stock, switching between fundamentalist and chartist strategies according to strategies’ profitability. HF agents differ from LF ones in many respects. First, unlike LF traders, activation of HF traders is not based on chronological time but it is event-based i.e., depends on specific market conditions (see Easley, López de Prado, and O’Hara, 2012). Second, HF agents adopt low-latency directional strategies that exploit the price and volume information released in the book by LF traders (cf. Aloud, Tsang, Olsen, and Dupuis, 2012). Third, HF traders keep their positions open for very short periods of time and they pursue tight inventory management (Kirilenko, Kyle, Samadi, and Tuzun, 2011). Lastly, HF endogenously cancel their orders based on expected profits (see Kirilenko, Kyle, Samadi, and Tuzun, 2011; SEC, 2014, for a review of cancellation practices of HF traders).

After checking the ability of the model to reproduce the main stylized facts of financial markets, we run extensive Monte-Carlo experiments to test the effectiveness of policies which have been proposed and implemented both in Europe and in the US to curb HFT and to prevent flash crashes, namely the implementation of i.) trading halt facilities (both ex-post and ex-ante designs); ii.) minimum resting times, ; iii.) order cancellation fees; iv.) transaction tax. Computer simulations show that slowing down high-frequency traders, by preventing them from frequently and rapidly cancelling their orders, ought to the introduction of either minimum resting times or cancellation fees, has beneficial effects on market volatility and on the occurrence of flash crashes. Also discouraging HFT via the introduction of a financial transaction tax produces similar outcomes (although the magnitude of the effects is smaller). All these policies impose a speed limit on trading.
Thus finding that they are valid tools to cope with volatility and the occurrence of flash crashes confirms the conjectures in Haldane (2011) about the need of tackling the “race to zero” of HF traders in order to improve financial stability. At the same time, we find that all these policies imply a longer duration of flash crashes, and thus a slower price recovery to normal levels. Furthermore, the results regarding the implementation of circuit breakers are mixed. On the one hand, the introduction of an ex-ante circuit breaker markedly reduces price volatility and completely removes flash crashes. This is merely explained by the fact that this type of regulatory design prevents the huge price drop, source of the flash crash. On the other hand, ex-post circuit breakers do not have any particular effect on market volatility, nor on the number of flash crashes. Moreover, they increase the duration of flash crashes.

Overall, our results indicate the presence of a fundamental trade-off characterizing HFT-targeted policies, namely one between market stability and market resilience. Policies that improve market stability - in terms of lower volatility and incidence of flash crashes - also imply a deterioration of market resilience - in terms of lower ability of the market price to quickly recover after a crash. This trade-off is explained by the dual role that HFT plays in the flash crash dynamics of our model. On the one hand, HFT is the source of flash crashes by occasionally creating large bid-ask spreads and concentrating orders on the sell side of the book. On the other, HFT plays a key role in the recovery from the crash by quickly restoring liquidity.

The paper is organized as follows. Section 2 describes the model. In Section 3, we present and discuss the simulation results, starting with a discussion of the main features of the flash crash dynamics in our model, and then moving to the presentation of the results concerning policy experiments. Finally, Section 4 concludes. The appendix briefly discusses the ability of the model to reproduce the main stylized facts of financial markets and contains the table of the parameters’ values used in the baseline scenario.

2 The Model

We use the model, developed in Jacob Leal, Napoletano, Roventini, and Fagiolo (2016), of a stock market populated by heterogeneous, boundedly-rational traders. Agents trade an asset for $T$ periods and transactions are executed through a limit-order book (LOB) where the information about the type, the size and the price of all agents’ orders is stored (see, for instance, Maslov, 2000; Zovko and Farmer, 2002; Avellaneda and Stoikov, 2008; Bartolozzi, 2010). The market is populated by two groups of agents depending on their trading frequency (i.e., the average amount of time elapsed between two order
placements), namely $N_L$ low-frequency (LF) and $N_H$ high-frequency (HF) traders ($N = N_L + N_H$). Although the number of agents in the two groups is kept fixed over the simulations, the proportions of low- and high-frequency traders change over time, as some agents may not be active in each trading session. Furthermore, agents of both types are different not only in terms of trading frequencies, but also in terms of strategies and activation rules. A detailed description of the behavior of LF and HF traders is provided in Sections 2.1 and 2.2.

In the model, a trading session is assumed to last one minute. At the beginning of each trading session $t$, active LF and HF agents know past market prices as well as past and current fundamental values of the traded asset. Based on the foregoing information set, each trading session $t$ proceeds in the following way. First, each active LF trader submits a buy or sell order to the LOB market, specifying its size and its limit price. Then HF orders are inserted in the book after LF ones and before the matching process takes place. The size and the price of their orders are also displayed in the LOB. We assume that HF traders are able to compute the transactions that would take place given the existing book, their prices and, therefore, their expected profits. They then use the computed expected profits to decide whether to confirm or to cancel their orders from the book. We capture the last feature by assuming that the matching procedure takes place in two steps. First, orders are matched according to their price and then arrival time in a “temporary” matching procedure. A “temporary” trading session price is hence computed. HF agents then check the expected profits they would make given the current book and decide to confirm/cancel their orders. Cancelled orders are removed from the book. Second, the actual matching takes place and the actual trading session price ($\bar{P}_t$) is determined as the price of the last executed transaction in the trading session. LF and HF unexecuted orders rest in the book for the next trading sessions ($\gamma_L$ and $\gamma_H$ periods, respectively). Lastly, given $\bar{P}_t$, all agents compute their actual profits and LF agents update their strategy for the next trading session (see Section 2.1 below). Notice that the possibility that some orders are removed from the book before the actual matching process takes place implies in general a difference between temporary and actual trading prices, as well as a difference between expected and actual traders' profits.

1In particular, we assume that HF agents simultaneously compute transactions, prices and expected profits based on the same order book information.

2The assumption that HF orders are inserted after the ones of LF traders and that HF agents are able to calculate expected profits before actual matching takes place are convenient ways of capturing one of the distinctive feature of HFT i.e., their ability to rapidly process a large amount of information and to exploit low-latency strategies (see e.g., Hasbrouck and Saar, 2013).

3The price of an executed contract is the average between the matched bid and ask quotes.
2.1 Low-Frequency Traders

In the market, there are \( i = 1, \ldots, N_L \) low-frequency agents who take short or long positions on the traded asset.\(^4\) The trading frequency of LF agents is based on chronological time, which is heterogeneous across LF agents and constant over time. In particular, each LF agents’ trading speed is drawn from a truncated exponential distribution with mean \( \theta \) and is bounded between \( \theta_{\text{min}} \) and \( \theta_{\text{max}} \) minutes.\(^5\)

In line with most heterogeneous agents models of financial markets, LF agents determine the quantities bought or sold (i.e., their orders) according to either a fundamentalist or a chartist (trend-following) strategy (see, e.g., De Long, Shleifer, Summers, and Waldmann, 1990; Lux and Marchesi, 2000; Farmer, 2002; Kirman and Teyssiére, 2002; Chiarella and He, 2003; Hommes, Huang, and Wang, 2005; Westerhoff, 2008). More precisely, given the last two market prices \( \bar{P}_{t-1} \) and \( \bar{P}_{t-2} \), orders under the chartist strategy \( D_{c,t} \) are determined as follows:

\[
D_{c,t} = \alpha_c (\bar{P}_{t-1} - \bar{P}_{t-2}) + \epsilon_{c,t},
\]

where \( 0 < \alpha_c < 1 \) and \( \epsilon_{c,t} \) is an i.i.d. Gaussian stochastic variable with zero mean and \( \sigma_c \) standard deviation. If a LF agent follows a fundamentalist strategy, her orders \( D_{f,t} \) are equal to:

\[
D_{f,t} = \alpha_f (F_t - \bar{P}_{t-1}) + \epsilon_{f,t},
\]

where \( 0 < \alpha_f < 1 \) and \( \epsilon_{f,t} \) is an i.i.d. Gaussian random variable with zero mean and \( \sigma_f \) standard deviation. The fundamental value of the asset \( F_t \) evolves according to a geometric random walk:

\[
F_t = F_{t-1}(1 + \delta)(1 + y_t),
\]

with i.i.d. \( y_t \sim N(0, \sigma^y) \) and a constant term \( \delta > 0 \). After \( \gamma_L \) periods, unexecuted orders expire, i.e. they are automatically withdrawn from the book. Finally, the limit-order price of each LF trader is determined by:

\[
P_{i,t} = \bar{P}_{t-1}(1 + \delta)(1 + z_{i,t}),
\]

where \( \delta > 0 \) and \( z_{i,t} \) measures the number of ticks away from the last market price \( \bar{P}_{t-1} \) and it is drawn from a Gaussian distribution with zero mean and \( \sigma^z \) standard deviation.\(^6\)

\(^4\)We assume that LF traders are not able to employ low-latency trading since they process information and respond to market events with a scale that is equal or higher than the one of the trading session.

\(^5\)See also Alfarano, Lux, and Wagner (2010) for a model with different time horizons in a setting different from ours.
In each period, low-frequency traders can switch their strategies according to strategy’s profitability. At the end of each trading session $t$, once the market price $\bar{P}_t$ is determined, LF agent $i$ computes her profits ($\pi_{st}^{i,t}$) under chartist ($st = c$) and fundamentalist ($st = f$) trading strategies as follows:

$$\pi_{st}^{i,t} = (\bar{P}_t - P_{i,t}) D_{st}^{i,t}. \quad (5)$$

Following Brock and Hommes (1998), Westerhoff (2008), and Pellizzari and Westerhoff (2009), the probability that a LF trader will follow a chartist rule in the next period ($\Phi_{c,i,t}^c$) is given by:

$$\Phi_{c,i,t}^c = \frac{e^{\pi_{c,i,t}/\zeta}}{e^{\pi_{c,i,t}/\zeta} + e^{\pi_{f,i,t}/\zeta}}, \quad (6)$$

with a positive intensity of switching parameter $\zeta$. Accordingly, the probability that LF agent $i$ will use a fundamentalist strategy is equal to $\Phi_{f,i,t}^f = 1 - \Phi_{c,i,t}^c$.

### 2.2 High-Frequency Traders

As mentioned above, the market is also populated by $j = 1, \ldots, N_H$ high-frequency agents who buy and sell the asset.\(^6\) Contrary to LF agents, HF traders employ low-latency technologies which enable them to place their orders with high speed. Moreover, HF agents differ from LF ones not only in terms of trading speed, but also in terms of activation and trading rules. In particular, contrary to LF strategies, which are based on chronological time, HF agents adopt trading rules framed in event time (see e.g., Easley, López de Prado, and O’Hara, 2012),\(^7\) i.e., the activation of HF agents depends on the extent of the last price change observed in the market. As a consequence, HF agents’ trading speed is endogenous. More specifically, each HF trader has a fixed price threshold $\Delta x_j$, drawn from a uniform distribution with support bounded between $\eta_{\text{min}}$ and $\eta_{\text{max}}$. This determines whether she will participate in the trading session $t$ (see Aloud, Tsang, and Olsen, 2013, for a similar attempt in this direction):

$$\left| \frac{\bar{P}_{t-1} - \bar{P}_{t-2}}{P_{t-2}} \right| > \Delta x_j. \quad (7)$$

\(^6\)We assume that $N_H < N_L$. The proportion of HF agents vis-à-vis LF ones is in line with empirical evidence (Kirilenko, Kyle, Samadi, and Tuzun, 2011; Paddrik, Hayes, Todd, Yang, Scherer, and Beling, 2011).

\(^7\)On the case for moving away from chronological time in modeling financial series see Mandelbrot and Taylor (1967); Clark (1973); Ané and Geman (2000).
Active HF agents submit buy or sell limit orders with equal probability \( p = 0.5 \) (Maslov, 2000; Farmer, Patelli, and Zovko, 2005).

HF traders adopt directional strategies that try to profit from the anticipation of price movements (see SEC, 2010; Alould, Tsang, Olsen, and Dupuis, 2012) and exploit the price and order information released by LF traders and by other HF traders (if any).

First, HF traders account for current order flows to determine their order size \( D_{j,t} \). More specifically, HF traders’ order size is drawn from a truncated Poisson distribution whose mean depends on volumes available in the sell-side (buy-side) of the LOB, if the order is a buy (sell) order.\(^8\) The ability of HF traders to adjust the volumes of their orders to the ones available in the book reflects their propensity to absorb LF agents’ orders. Moreover, in order to account for empirical evidence indicating that HF traders do not accumulate large net positions (CFTC and SEC, 2010; Kirilenko, Kyle, Samadi, and Tuzun, 2011), we add two additional constraints to HF order size. On the one hand, HF traders’ net position is bounded between +/-3,000. On the other hand, HF traders’ buy (sell) orders are smaller than one quarter of the total volume present in the sell (buy) side of the book (see, for instance, Kirilenko, Kyle, Samadi, and Tuzun, 2011; Bartolozzi, 2010; Paddrik, Hayes, Todd, Yang, Scherer, and Beling, 2011).\(^9\)

Second, HF traders account for current best ask and bid prices to set their order limit price (see Section 2.2 below). In particular, in each trading session \( t \), HF agents trade near the best ask \( (P_{\text{ask}}^t) \) and bid \( (P_{\text{bid}}^t) \) prices available in the LOB (see e.g., Paddrik, Hayes, Todd, Yang, Scherer, and Beling, 2011).\(^10\) Accordingly, HF buyers and sellers’ limit prices are formed as follows:

\[
P_{j,t} = P_{\text{ask}}^t (1 + \kappa_j) \quad P_{j,t} = P_{\text{bid}}^t (1 - \kappa_j),
\]

where \( \kappa_j \) is drawn from a uniform distribution with support \( (\kappa_{\text{min}}, \kappa_{\text{max}}) \). Figure 1 illustrates HF order placement using directional strategy based on a spread of one tick.

A key characteristic of empirically-observed high-frequency trading is the high order cancellation rate (CFTC and SEC, 2010; Kirilenko, Kyle, Samadi, and Tuzun, 2011). We introduce such a feature in the model in two ways. First, we assume that HF agents’ unexecuted orders are automatically removed from the book after a given time period.

\(^8\)In the computation of the mean of the Poisson distribution, the relevant market volumes are weighted by the parameter \( 0 < \lambda < 1 \).

\(^9\)Our assumption about HF orders’ size reflects empirically-observed HF characteristics, namely HF traders are few firms in the market but represent more than 30% of total trading volume (Kirilenko, Kyle, Samadi, and Tuzun, 2011; Aldridge, 2013).

\(^10\)This assumption is consistent with empirical evidence on HF agents’ behavior, which suggests that most of their orders are placed very close to the last best prices (SEC, 2010).
\( \gamma^H \) (i.e., exogenous cancellation), which is shorter than LF agents’ one, i.e. \( \gamma^H < \gamma^L \). Moreover, in each period, HF traders are able to process and to use available market information to decide whether to cancel their orders, while these orders could be executed (i.e., endogenous cancellation). More precisely, we assume that - once new orders have been inserted in the book - HF traders are able to simultaneously compute the volumes and prices of the transactions that would take place conditionally on the existing book, and on this basis, they are able to compute their expected profits. HF agents will cancel their orders when expected profits are negative. Instead, they will confirm their orders when expected profits are non-negative. More formally, let \( \pi^E_{j,t} \) be the expected profits of HF trader \( j \) conditional on the book available in period \( t \), we get:

\[
\begin{align*}
\pi^E_{j,t} < 0, & \quad \text{cancel order} \\
\pi^E_{j,t} \geq 0, & \quad \text{confirm order}
\end{align*}
\]

where \( \pi^E_{j,t} \) is determined by:

\[
\pi^E_{j,t} = (\bar{P}^\text{temp}_t - P^\text{temp}_{j,t})D_{j,t}. 
\] (9)

where \( \bar{P}^\text{temp}_t \) and \( P^\text{temp}_{j,t} \) are the temporary market price and the temporary transaction price of agent \( j \), respectively, and \( D_{j,t} \) is the size of the order of agent \( j \). Lastly, at the end of each trading session, HF traders’ profits \( \pi_{j,t} \) are computed as follows:

\[
\pi_{j,t} = (\bar{P}_t - P_{j,t})D_{j,t}. 
\] (10)

Figure 1: Directional strategy order placement.
where $P_{j,t}$ is her actual transaction price and $\bar{P}_t$ is the actual market price. As already mentioned at the beginning of the section, given that HF traders can intentionally cancel orders, the final order book would be different from the one before HF traders’ endogenous cancellation. Accordingly, expected profits of HF traders could be different from expected ones.

3 Simulation Results

We investigate the properties of the model presented in the previous section via Monte-Carlo simulations. More precisely, we carry out $MC = 50$ Monte-Carlo iterations, each one composed of $T = 1,200$ trading sessions using the baseline parametrization, described in Table 8 (see Appendix A.2). The value of the parameters employed in our simulations are in line with existing works.\(^{11}\)

As a first step in our analysis of simulation results, we verify that our ABM is able to jointly reproduce the main stylized facts of financial markets with the same configuration of parameter values (see details in Appendix A.1). We then assess the properties of the model in generating flash crashes and we investigate the key determinants of flash crashes, distinguishing the initial sharp price drop and the subsequent price recovery (see Section 3.1). Lastly, we investigate the effectiveness of a set of regulatory policies on market volatility, the frequency and the duration of flash crashes (see Section 3.2).

3.1 HFT and the Anatomy of Flash Crashes

In line with empirical evidence (CFTC and SEC, 2010; Kirilenko, Kyle, Samadi, and Tuzun, 2011), we identify flash crashes as drops in the asset price of at least 5% followed by a sudden recovery of at most 30 minutes (corresponding to thirty trading sessions in each simulation run). Applying such a definition, in line with Jacob Leal, Napoletano, Roventini, and Fagiolo (2016), we find that our model is able to endogenously generate flash crashes as an emergent property resulting from the interactions between low- and high-frequency traders (see Table 1). Indeed, we find that flash crashes emerge only when HF traders are present in the market and their frequency is significantly higher than one. In contrast, when the market is only populated by LF traders, flash crashes do not emerge.

\(^{11}\)More precisely, for the LF trading strategies equations, we chose the same values employed in previous ABM works (e.g., Westerhoff, 2008). In addition, following Paddrik, Hayes, Todd, Yang, Scherer, and Beling (2011), several values of the parameters concerning HF traders’ behavior (e.g., order size) were selected in order to be consistent with the evidence reported in Kirilenko, Kyle, Samadi, and Tuzun.
Table 1: Market volatility ($\sigma_P$) and flash crashes statistics in the baseline scenario with HF traders and in the scenario with only LF traders.

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_P$</th>
<th>Number of flash crashes</th>
<th>Avg. duration of flash crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>0.016 (0.002)</td>
<td>4.636 (0.398)</td>
<td>7.139 (0.484)</td>
</tr>
<tr>
<td>Only-LFT</td>
<td>0.002 (0.000)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Figure 2: Complementary cumulative distributions of bid-ask spreads in different market phases. Pooled sample from 50 independent Monte-Carlo runs.

What are the main drivers of the emergence of flash crashes in our model? First, the directional strategies employed by HF traders can lead to large bid-ask spreads, setting the premises for the emergence of flash crashes. Figure 2 shows the distributions of bid-ask spreads conditioned on different market phases, i.e., normal times, crash and recovery phases. We observe that the mass of the distribution of bid-ask spreads is significantly shifted to the right during crashes, clearly indicating the presence of large (2011).

In particular, we construct the pooled samples (across Monte-Carlo runs) of bid-ask spread values singling out “normal time” phases and decomposing “flash-crash” periods in “crash” phases (i.e. periods of sharp drops in the asset price) and the subsequent “recovery” phases (i.e. periods when the price goes back to its pre-crisis level). Next, we estimate the complementary cumulative distributions of bid-ask spreads in each market phase using a kernel-density estimator.
bid-ask spreads at the time of the price fall. The emergence of large bid-ask spreads is explained by the different strategies employed by high- and low-frequency traders in our model. Active LF traders set their order prices “around” the price of the last trading session, which tends to fill the existing gap between the best bid and ask prices at the beginning of a given trading section. In contrast, active HF traders, who submit their orders after LF agents, place large buy (sell) orders just few ticks above (below) the best ask (bid), which tends to widen the bid-ask spread in the LOB.

HFT-induced large bid-ask spreads are therefore one key driver of flash crashes in our model. This is further confirmed by the analysis of agents’ aggressiveness behaviour conditional on different market phases. First Figure 3, shows that the model generates a positive relationship between agents’ orders aggressiveness and the size of the bid-ask spreads. Thus, higher orders’ aggressiveness (of any agents’ type) leads to larger bid-ask spreads in our model. However, the degree of aggressiveness of HF and LF agents markedly differ across market phases. Table 2 shows average orders’ aggressiveness ra-

---

13 The plot in Figure 3 shows the contour of the theoretical function between bid-ask spreads and HF and LF orders’ aggressiveness ratios that is implied by the dynamics of the model. The function was generated by interpolating the scattered data of aggressiveness ratios of HF and LF traders and bid-ask spread pooled across Monte-Carlo simulations in the baseline scenario. The interpolation was performed by using the scatteredInterpolant function in Matlab.

---
Table 2: Orders’ aggressiveness ratios for different categories of traders and different market phases. Values are averages across 50 independent Monte-carlo runs. Monte-carlo standard errors in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>LFT buy</th>
<th>HFT buy</th>
<th>LFT sell</th>
<th>HFT sell</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal times</td>
<td>0.086</td>
<td>0.130</td>
<td>0.041</td>
<td>0.108</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Crashes</td>
<td>0.000</td>
<td>0.013</td>
<td>0.000</td>
<td>0.831</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.008)</td>
<td>(0.000)</td>
<td>(0.031)</td>
</tr>
<tr>
<td>Recovery</td>
<td>0.083</td>
<td>0.106</td>
<td>0.004</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.009)</td>
<td>(0.002)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Unconditional values</td>
<td>0.086</td>
<td>0.130</td>
<td>0.041</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.002)</td>
</tr>
</tbody>
</table>

Aggressiveness ratios for both types of agents and book sides, and conditional on different market phases (i.e., “normal times”, “crash”, “recovery”). As the table shows, in all market phases, aggressiveness ratios of LF traders are much lower than HF traders ones. Moreover, order aggressiveness of HF traders is low during normal market phases. For instance, nearly 90% of buy and sell orders placed by HF traders do provide liquidity to the market, which contributes to keep bid-ask spreads low. In contrast, orders’ aggressiveness of HF agents increases abruptly in the crash phase (see Table 2). In particular, in such a phase, most HF sell orders are aggressive (about 85%) and thus remove liquidity from the market and generate large bid-ask spreads. Instead, HF aggressiveness is very low on the buy side.

To sum up, the above discussion shows that flash crashes in our model are the result of:

i.) large bid-ask spreads and

ii.) concentration of aggressive orders on the sell-side of the book.

It is worth noticing that these explanations for the emergence of flash crashes are in line with the empirical evidence about the market dynamics observed, for instance, in empirical studies like Kirilenko, Kyle, Samadi, and Tuzun (2011) and Baron, Brogaard, Hagströmer, and Kirilenko (2015). In these studies, an incoming order is considered “aggressive” if it is matched against an order that is resting in the book, i.e., if it removes liquidity from the market. In contrast, an order provides liquidity on the market if it fills the book of resting orders. Finally, it has no effect on market liquidity if it is matched against another incoming order in the same trading session.

14Alike Jacob Leal, Napoletano, Roventini, and Fagiolo (2016), we use the definition provided by trading platforms (e.g., CME Globex), and widely used in the empirical literature (Kirilenko, Kyle, Samadi, and Tuzun, 2011; Baron, Brogaard, Hagströmer, and Kirilenko, 2015). An incoming order is considered “aggressive” if it is matched against an order that is resting in the book, i.e., if it removes liquidity from the market. In contrast, an order provides liquidity on the market if it fills the book of resting orders. Finally, it has no effect on market liquidity if it is matched against another incoming order in the same trading session.

15In Jacob Leal, Napoletano, Roventini, and Fagiolo (2016), we also show that such an asymmetry is further confirmed by the distribution of overall orders of HF traders across market sides. There, we also explain how HF orders’ synchronization is an emergent property related to the event-time strategy of HF traders and may emerge even if the choice of each HF agent between selling or buying is a Bernoulli distributed variable with probability p = 0.5.
during the flash crash of May 6th, 2010 (CFTC and SEC, 2010; Kirilenko, Kyle, Samadi, and Tuzun, 2011). Moreover, computer simulations highlight the key role that high-frequency trading has in generating such extreme events in financial markets. Indeed, the emergence of periods of high market illiquidity is endogenous and intimately related to the pricing strategies of HF traders (see Eq. 8). In that, flash crashes are therefore not simply generated by large orders and thus cannot be associated with “fat finger” explanations (see Haldane, 2011, for a discussion of the different proposed explanations of flash crashes). Finally, our simulation results confirm on the one hand that - in line with recent empirical evidence (see e.g. Brogaard, 2010; Menkveld, 2013) - HF traders may have a beneficial effect on markets during normal times, by providing non-aggressive orders and therefore contributing to keep bid-ask spread low. On the other hand, they also show that the liquidity provided by HFT is extremely fragile and that orders of HF agents can occasionally be extremely aggressive, removing liquidity from the market, and generate abrupt and large drops in the market price.

The above discussion has made clear that HFT plays a key role in causing the significant price falls, which are the footprint of all flash crashes. However, HFT also actively contributes to the quick recovery after the crash. Table 2 shows indeed that the orders’ aggressiveness ratios of HF agents are much lower during the recovery phase of the flash crash. In addition, orders’ aggressiveness ratios are symmetric between the buy and sell sides of the book. Thus, orders of HF agents contribute to restore liquidity in such a phase, thereby favoring the recovery of the price. The return to normal liquidity conditions during the recovery is also documented by the behavior of the conditional bid-ask spread distribution (cf. Figure 2). The distribution of the bid-ask spreads during recoveries is indeed not statistically different from the one during normal times.

Two factors explain the positive role played by HFT in favoring price-recovery after the crash. The first is that wide variations in asset prices trigger the activation of a large number of high-frequency traders which leads to a surge in order volumes of HF agents. In addition, as each HF trader is either a buyer or a seller with probability $p = 0.5$, when the number of active HF agents is large, HF orders will tend to be equally split between the sell- and buy-side of the LOB, which explains the symmetry in HF orders’ aggressiveness ratios observed during the recovery (see Table 2). The second element supporting the rapid price recovery is the order-cancellation rate of HF traders. Indeed, high order cancellation implies a short duration of HF orders in the book. As a matter of fact, this also implies that the HF bid and ask quotes will tend to reflect current market conditions. Such a memory effect of HF orders explains the low time persistence of high bid-ask spreads after a crash and contributes to the quick replenishment of market
liquidity and price.

3.2 Regulatory policies experiments

In the previous section, we have documented that the model is able to robustly reproduce the main stylized of financial markets and to endogenously generate flash crashes as the result of the trading activity of HF traders. We pointed out the very reasons underlying both phases of a flash crash, namely the sharp price drop and the swift recovery of the price. In this context, we now turn to use the model as a test-bed to investigate the effectiveness of a set of regulatory policies which have been so far implemented and proposed to cope with the possible negative effects of high-frequency trading and to curb flash crashes. We focus on the following policies: i.) circuit breakers, ii.) minimum resting times, iii.) cancellation fees, iv.) financial transaction taxes. Moreover, we study the impact of the aforementioned policies on price-returns volatility as well as on the number and duration of flash crashes.

3.2.1 Circuit breakers

Section 3.1 provides insights about the mechanisms through which HFT may be a source of episodic price instability and systemic risk. Regulators have recently taken proactive steps to avoid flash crashes and to deal with periodic illiquidity in markets. In particular, in the aftermath of May 2010 Flash Crash, the CFTC and the SEC proposed several measures to prevent this type of extreme events such as, for instance, updated circuit breakers (SEC, 2011b, 2012) and limit up/limit down mechanisms (also known as ex-ante circuit breakers, see SEC, 2011a, 2012; Haldane, 2011). Indeed, extreme price fluctuations are likely to exacerbate execution uncertainty and discourage trading (Greenwald and Stein, 1991; Subrahmanyam, 2012). Instead, trading halts should allow for a “cool-down” period, improve market liquidity and reduce volatility (Greenwald and Stein, 1991; Kodres and O’Brien, 1994; Ackert, 2012). Circuit breakers (or impediments to trade), i.e., mechanisms designed to reduce the risk of a price collapse by means of trading halts in presence of excessive price volatility, have been implemented for long time in many exchanges, both in Europe and in the US (CFTC and SEC, 2010; Furse, Haldane, Goodhart, Cliff, Zigrand, Houstoun, Linton, and Bond, 2011; Prewitt, 2012; Gomber and Haferkorn, 2013). However, they were traditionally market-wide and triggered only by large price movements. They were therefore conceived only as ex-post reactions to excessive price volatility. After the events of May 6, 2010, new and more sensitive stock-specific systems, which work on an ex-ante basis, have therefore been im-
plemented (SEC, 2012). Nowadays, circuit breakers can take many forms, from trading halts in single stocks or in entire markets to limit up and down prices with a variety of percent price change and different reference points, and restrictions on one trading venue or across multiple venues (Furse, Haldane, Goodhart, Cliff, Zigrand, Houstoun, Linton, and Bond, 2011; Subrahmanyam, 2012).

The empirical evidence on the efficacy of circuit breakers\[^{16}\] and price limits\[^{17}\] is still limited. Accordingly, it is still not clear what type of breakers are the most effective. In this section we try to contribute to the existing literature on circuit breakers by performing a computational test of their impact on volatility and the duration of flash crashes.

We focus on two distinct types of circuit breakers, namely: i. a trading halt in single stock triggered by a certain percent price change from the last price, as implemented, for instance, at the NYSE-Euronext, the London Stock Exchange and the Deutsche Bourse (i.e., an ex-post device); ii. a limit up/limit down price mechanism in place e.g., on the NYSE and NASDAQ, Tokyo Stock Exchange and Korea Exchange (i.e., an ex-ante device). More precisely, we first study the effect of introducing an ex-post trading halt mechanism in response to substantial price drops which is intended to stop trading in the exchange for a time period (\(np\)). In this Monte-Carlo experiment, the circuit breaker is triggered by a relative price change from the last price that is in absolute value larger than \(\beta\%), where \(\beta = 5\%\),\[^{18}\] and the trading halt is assumed to last for \(np = 5\) periods. Notice that, by construction, this type of circuit breaker leaves unaffected the number of flash crashes, but it could have an impact on the duration of flash crashes and on market volatility. In the second type of experiment, we introduce a limit up/limit down price of \(\beta = 5\%\) which, when triggered, stops trading in the exchange for \(np = 5\) periods. In this case, the trading halt occurs before the trading session price is formed. In such a case, the imposition of limit up/limit down price completely removes flash crashes from the dynamics.

The results of both experiments are shown in Table 3.\[^{19}\] We only report results for market volatility and the average duration of flash crashes. First, we find that the introduction of ex-post circuit breakers has a negligible effect on market volatility.

\[^{16}\]Few examples include e.g., Lauterbach and Ben-Zion (1993); Santoni and Liu (1993); Goldstein and Kavajecz (2004); Brugler and Linton (2014).

\[^{17}\]Some examples include, for instance, Kim and Rhee (1997); Cho, Russell, Tiao, and Tsay (2003); Diacogiannis, Patsalis, Tsangarakis*, and Tsiritakis (2005); Bildik and Güleryüz (2006); Stamatou (2008).

\[^{18}\]Notice that, in our parametrization, the threshold for the trading halt activation corresponds to the one used to identify flash crashes.

\[^{19}\]We also ran simulations for alternative values of \(\gamma^H\) (\(\gamma^H = 1\) and \(\gamma^H = 1200\)) and for \(\beta = 3\%\). Results are however consistent with the ones presented in Table 3.
Table 3: The effect of different types of circuit breakers on price volatility and flash crash statistics when $\gamma^H = 20$ and $\beta = 5\%$. Values are averages across 50 independent Monte-Carlo runs. Monte-Carlo standard errors in parentheses. ($\sigma_P$): price returns volatility.

<table>
<thead>
<tr>
<th></th>
<th>$\sigma_P$</th>
<th>Avg. duration of flash crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>No circuit breaker baseline</td>
<td>0.016</td>
<td>7.139</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.484)</td>
</tr>
<tr>
<td>ex-post circuit breaker</td>
<td>0.010</td>
<td>13.345</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.609)</td>
</tr>
<tr>
<td>ex-ante circuit breaker</td>
<td>0.005</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>-</td>
</tr>
</tbody>
</table>

Moreover, this defensive regulation has a detrimental effect on the duration of the flash crash, since the trading halt merely slows down the price recovery. Indeed, we observe that the price would have recovered sooner without the imposition of the circuit breaker, i.e., if HF traders would have been able to fully play their role in the recovery phase of the flash crash. This is mainly explained by the positive role played by HF traders in the recovery from the crash (see Section 3.1). The imposition of a trading halt instead prevents HF traders from providing the required liquidity after the crash and thus leads to longer flash crashes.

How do results change if we turn from ex-post to ex-ante circuit breakers? First, besides removing flash crashes in our parametrization, ex-ante trading halts lead to a reduction in price volatility compared to the baseline (compare first and third row of the first column in Table 3). This is mainly explained by the fact that this device is triggered before trade is actually performed and therefore it prevents extreme price fluctuations.

Overall, and in line with earlier works (see e.g. Subrahmanyam, 2013; Apergis, 2014), our results show that breakers should be used with caution, especially when they represent impediments to trade and deteriorate the trading process within a particular stock. In particular, our findings indicate that ex-ante circuit breakers are a much more effective tool than ex-post trading halts, because they completely remove extreme drops in price from the market and they significantly dampen market volatility. In contrast, ex-post trading halts have only a limited impact on volatility. In addition, they may introduce important distortions in the natural process of recovery from a crash that would otherwise took place.
3.2.2 Minimum resting times

Minimum resting times specify a minimum time that a limit order must remain in the book i.e., it cannot be cancelled within a given time span ($\gamma^H$). The impetus for imposing this command-and-control regulatory instrument is that markets operating at high speed are characterized by a large number of orders that are cancelled very quickly after submission. Orders’ cancellation is a inherent feature of many HF traders strategies and has raised many critiques against HFT (CFTC and SEC, 2010; Kirilenko, Kyle, Samadi, and Tuzun, 2011). Indeed, the ability of HF traders to quickly cancel their orders could render market liquidity misleading (Kirilenko, Kyle, Samadi, and Tuzun, 2011; Prewitt, 2012; Breckenfelder, 2013; Friederich and Payne, 2015), and it could favor price short-term volatility (Hanson, 2011; Bershova and Rakhlin, 2013; Breckenfelder, 2013). Furthermore, rapid order cancellations are likely to increase the cost of monitoring the market for all participants and reduce the predictability of a trade’s execution quality, given that the quotes displayed may have been cancelled by the time the new order hits the resting order (Furse, Haldane, Goodhart, Cliff, Zigrand, Houstoun, Linton, and Bond, 2011). Nevertheless, the net benefits of minimum resting times are still unclear (Furse, Haldane, Goodhart, Cliff, Zigrand, Houstoun, Linton, and Bond, 2011). On the one hand, minimum resting times can increase the likelihood of a viewed quote being available to trade and therefore make the order book dynamics more transparent. In addition, longer expiration times create liquidity that reduces price variance in the market (Brewer, Cvitanic, and Plott, 2013). Lastly, by “slowing down” markets, minimum resting times may favor participation, especially if some traders (e.g., small retail investors) feel that high speed makes market unfair and hursts market integrity (see, for instance, Haldane, 2011). On the other hand, minimum resting times can impinge upon hedging strategies that operate by placing order across markets and expose liquidity providers to increased “pick-off risk” due to the inability to cancel stale orders (Oxera, 2012). Liquidity provision may be even more impeded during times of high volatility, when it is particularly expensive to post limit orders. Furthermore, this measure may also change the dynamics of the market by attracting more aggressive HFT (Farmer and Skouras, 2013). Lastly, market quality may be diminished due to higher transaction costs for the end users and lower price efficiency.

In this context, and given that the empirical evidence about the effects of minimum resting times is still limited,\(^{20}\) in this section, we aim at shedding some light on the

\(^{20}\)For instance, the work of Furse, Haldane, Goodhart, Cliff, Zigrand, Houstoun, Linton, and Bond (2011) reports only two cases for the implementation of minimum resting times. Namely ICAP which introduced a minimum quote lifespan on its electronic in June 2009 and the Istanbul Stock Exchange
impact of minimum resting times on market dynamics by investigating the effects of such a measure on market volatility as well as on the number and the duration of flash crashes. To this end, we run a Monte-Carlo experiment where we impose that HF orders cannot deliberately cancel their orders for a number of periods equal to the expiration time $\gamma^H$. We then increase the expiration time while keeping all other parameters at their baseline values (see Appendix A.2). In this experiment, we vary the parameter $\gamma^H$ from 1 to 60 periods/minutes. The results of this experiment are reported in Table 4. Table 4 reveals that increasing minimum resting times (i.e., making $\gamma^H$ higher) dampens market volatility (see second column in Table 4). The beneficial effect on volatility is one of the purported primary effects of the measure (see in particular SEC, 2010) and which did not allow the cancellation of limit orders during continuous auction mode until mid-2011. However, it is not clear what one can really learn from these two experiments.
is consistent with earlier works (see in particular Hayes, Paddrik, Todd, Yang, Beling, and Scherer, 2012). This outcome is mainly explained by the fact that minimum resting times slow down HF traders and prevent them from aggressively trading on the most recent news and information disclosed in the LOB.

Minimum resting times (i.e., higher $\gamma^H$) have also a beneficial effect on the number of flash-crash episodes (see third column of Table 4). This outcome again stems from the lower aggressiveness that such a measure imposes on HF trading strategies. In contrast, we find that the duration of flash crashes is inversely related to the duration of minimum resting times (cf. fourth column of Table 4). This finding is explained by the fact that stricter rules on orders’ expiration of HF orders also imply a longer memory effect (cf. Section 3.1). In fact, as $\gamma^H$ increases, the bid and ask quotes posted by HF agents stay longer in the LOB and therefore large bid-ask spreads persist more. Furthermore, less HF traders participate in the market based on the most recent market information. This slows down the replenishment of market liquidity and prevents the quick price recovery. Lastly, the number of contracts traded at prices close to the flash-crash one rises which prevents the price rebound.

Overall, the above results thus indicate that the imposition of minimum resting times can be a very effective tool in order to dampen market volatility and to reduce the incidence of flash crashes. In that, they bring support to earlier works advocating for such a measure (Haldane, 2011; SEC, 2010). At the same time, they also hint to the presence of a trade-off between volatility and incidence of extreme events, on the one hand, and price-resilience (because of longer recoveries) on the other hand. 21 As we shall discuss in the next sections, such a trade-off is also inherent the market-based measures on which we focus on, namely cancellation fees and financial transaction taxes.

### 3.2.3 Cancellation fees

We now turn to investigate the effect of the imposition of cancellation fees on price volatility, the frequency of flash crashes and their duration. Both US and EU regulators have called for the imposition of cancellation fees. However, they have been only incompletely enforced in a couple of exchanges since 2012 (Nasdaq and Direct Edge, Borsa Italia and Deutsche Borse stock exchanges). Cancellation fees are primarily intended to prevent overload in the exchange computer systems and to discourage the most flagrant excessive cancellations which represent unnecessary messages that do not result in trades and which, rather, come along with higher volatility (Prewitt, 2012). A portion of such

---

21 Haldane (2011) also points to the presence of a similar trade-off when deciding whether to impose resting rules or not (market efficiency versus stability).
traffic is likely to be inefficient and may raise costs to other investors who try to mon-
itor the market. Such fees would therefore discourage traders from posting orders that
are not intended to be executed (Prewitt, 2012). They will also discourage manipula-
tive HFT strategies (like stuffing and spoofing) that involve massive order cancellations
by rendering them uneconomical (Biass and Woolley, 2011; Prewitt, 2012). At the same
time, rapid reaction to new information is often a way for market makers to minimize the
risks of offering prices to other traders, and contributes to lower trading costs (Copeland
and Galai, 1983; Foucault, Rœll, and Sandás, 2003). In that, the imposition of cancel-
lation fees could instead discourage the activity of active market makers and liquidity
providers, and lead to an increase in transaction costs.

In our experiment, HF traders who deliberately decide to cancel their orders before
\( \gamma^H \) periods are charged a fee \( c \). As a result, a HF agent will cancel her order if expected
losses from trade are higher in magnitude than the cancellation cost, \( i.e., \) when \( \pi^E_{j,t} < 
- c.D_{j,t} \). The policy exercise was carried under three scenarios: \( i. \) when \( \gamma^H = 1 \), HF
traders can deliberately decide to cancel their orders before the expiration date \( (\gamma^H \) periods). However, given that the expiration date, in this case, is very small \( (i.e., \gamma^H = 1) \), HF traders will have to pay the cancellation fee only on very fast order
cancellation. This scenario represents a very soft policy measure where most cancelled
orders are not charged the fee; \( ii. \) when \( \gamma^H = 20 \), HF traders can decide to deliberately
cancel their orders before their expiration date \( (\gamma^H \) periods), which leads to the payment
of the cancellation fee \( c \). However, older unexecuted HF orders which are automatically
withdrawn from the book after \( \gamma^H \) are not charged the cancellation fee. This scenario
represents a moderate policy measure where not all cancelled orders are charged the fee;
\( iii. \) when \( \gamma^H = 1200 \) (i.e. it corresponds to the length of a Monte-Carlo iteration in our
setting), HF orders can only be intentionally cancelled by HF agents. In this case, HF
unexecuted orders stay in the LOB until the end of the simulation \( (i.e., T = \gamma^H = 1200) \)
and the cancellation fee is charged on all cancelled HF orders. This scenario represents
the imposition of a very stringent policy measure. Furthermore, given the wide variety of
fee levels currently used worldwide, we tested the effect of different levels of cancellation
fees, \( c \) varying from 0.01% to 1%. The results of the above experiments are reported in
Table 5.

First, and not surprisingly, we find that, when \( \gamma^H = 1 \), the imposition of a cancel-
lation fee is not effective in dealing with volatility and flash crashes, whatever the size
of the cancellation fee is. Indeed, in this case, price volatility, the frequency and the
duration of flash crashes are not significantly different with respect to the baseline case.
This is mainly explained by the fact that, when \( \gamma^H = 1 \), HF traders frequently cancel
Table 5: HF traders’ order cancellation fees, price volatility and flash crash statistics for different values of $\gamma^H$ and different values of $c$. Values are averages across 50 independent Monte-Carlo runs. Monte-Carlo standard errors in parentheses. ($\sigma_P$): price returns volatility.

<table>
<thead>
<tr>
<th>$\gamma^H$ = 1</th>
<th>$c$</th>
<th>$\sigma_P$</th>
<th>Number of flash crashes</th>
<th>Avg. duration of flash crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.017</td>
<td>4.652</td>
<td>(0.002)</td>
<td>(0.390) (0.575)</td>
</tr>
<tr>
<td>0.01%</td>
<td>0.015</td>
<td>6.262</td>
<td>(0.001)</td>
<td>(0.568) (0.632)</td>
</tr>
<tr>
<td>0.1%</td>
<td>0.017</td>
<td>6.808</td>
<td>(0.002)</td>
<td>(0.793) (0.576)</td>
</tr>
<tr>
<td>1%</td>
<td>0.016</td>
<td>7</td>
<td>(0.002)</td>
<td>(0.720) (0.568)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\gamma^H$ = 20</th>
<th>$c$</th>
<th>$\sigma_P$</th>
<th>Number of flash crashes</th>
<th>Avg. duration of flash crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.016</td>
<td>4.636</td>
<td>(0.002)</td>
<td>(0.398) (0.484)</td>
</tr>
<tr>
<td>0.01%</td>
<td>0.006</td>
<td>2.200</td>
<td>(0.000)</td>
<td>(0.238) (0.957)</td>
</tr>
<tr>
<td>0.1%</td>
<td>0.006</td>
<td>1.750</td>
<td>(0.001)</td>
<td>(0.123) (1.234)</td>
</tr>
<tr>
<td>1%</td>
<td>0.007</td>
<td>2.115</td>
<td>(0.001)</td>
<td>(0.231) (1.162)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\gamma^H$ = 1200</th>
<th>$c$</th>
<th>$\sigma_P$</th>
<th>Number of flash crashes</th>
<th>Avg. duration of flash crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.014</td>
<td>3.909</td>
<td>(0.001)</td>
<td>(0.389) (0.531)</td>
</tr>
<tr>
<td>0.01%</td>
<td>0.002</td>
<td>1.000</td>
<td>(0.000)</td>
<td>(0.000) (0.200)</td>
</tr>
<tr>
<td>0.1%</td>
<td>0.003</td>
<td>1.000</td>
<td>(0.000)</td>
<td>(0.000) (1.794)</td>
</tr>
<tr>
<td>1%</td>
<td>0.002</td>
<td>–</td>
<td>(0.000)</td>
<td>– (–)</td>
</tr>
</tbody>
</table>
their orders because they are not penalized by the cancellation fee.

In contrast, we find that, in scenarios (ii.) and (iii.), the introduction of a cancellation fee may be an effective tool to deal with market volatility and the number of flash crashes. Furthermore, the level of the fee matters, since we observe that the higher is the cancellation fee, the greater are the effects on price volatility and the occurrence of flash crashes. In particular, under the most stringent scenario (i.e., $\gamma^H = 1200$), flash crashes completely vanish for high values of the cancellation fee and this regulatory instrument is thus very effective to deal with such extreme events. In the mild scenario (i.e., $\gamma^H = 20$) this type of policy measure is still effective to curb HFT and to mitigate flash crashes. These findings confirm one common claim against HFT according to which HF high cancellation rates may destabilize markets (SEC, 2014). Accordingly, preventing HF traders from quickly cancelling their orders decreases market volatility and completely removes flash crashes from the market.

Furthermore, the introduction of a cancellation fee tends to significantly reduce HF orders’ aggressiveness. Table 6 shows the (buy and sell) orders’ aggressiveness ratios for both HF and LF traders in the mild scenario when $\gamma^H = 20$ and $c = 0.01$. Reported values are unconditional and for different market phases. The table also compares orders’ aggressiveness ratios with and without a cancellation fee to the one that emerge in presence of a financial transaction tax (see next section). As this table reveals, the introduction of the cancellation fee generates a situation where the aggressiveness of HF traders is significantly lower than the one of LF traders, both unconditionally as well as in the normal times and recovery phases. Not surprisingly, when $\gamma^H = 20$, the average bid-ask spread is significantly lower than in the baseline (1.022 versus 1.577). These outcomes are explained by the fact that the existence of the cancellation fee effectively discourages HF traders to frequently cancel their orders, since they have an incentive to keep orders with a lower expected profit.

However, and similarly to minimum resting times and circuit breakers, the beneficial effects of cancellation fees come at the cost of a longer duration of flash crashes. Again, this outcome is explained by the fact that in presence of a cancellation fee, HF orders stay longer in the book. This does not only prevent the activation of HF traders in the recovery. It also implies that HF quotes in the book tend to reflect close-to-crash conditions. We therefore point out that preventing HF traders from quickly modifying and cancelling their orders slows down the price recovery, since HF orders do not reflect the most recent market conditions. As a result, the positive role HFT plays in the recovery from the crash is significantly dampened. This is further supported by the fact that, when $\gamma^H = 20$, the trading to book volume ratio is significantly lower than in the
Table 6: Orders’ aggressiveness ratios for different categories of traders and different market phases when \( \gamma_H = 20 \) and \( c = ft t = 0.01 \). Values are averages across 50 independent Monte-carlo runs. Monte-carlo standard errors in parentheses.

<table>
<thead>
<tr>
<th></th>
<th>Cancellation fee</th>
<th></th>
<th>Financial transaction tax</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LFT orders</td>
<td>HFT orders</td>
<td>LFT orders</td>
<td>HFT orders</td>
</tr>
<tr>
<td>Normal times</td>
<td>0.157</td>
<td>0.036</td>
<td>0.133</td>
<td>0.159</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td>(0.003)</td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Crashes</td>
<td>0.000</td>
<td>0.999</td>
<td>0.000</td>
<td>0.866</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.039)</td>
</tr>
<tr>
<td>Recovery</td>
<td>0.332</td>
<td>0.109</td>
<td>0.069</td>
<td>0.215</td>
</tr>
<tr>
<td></td>
<td>(0.021)</td>
<td>(0.015)</td>
<td>(0.007)</td>
<td>(0.032)</td>
</tr>
<tr>
<td>Unconditional</td>
<td>0.159</td>
<td>0.038</td>
<td>0.131</td>
<td>0.161</td>
</tr>
<tr>
<td>baseline</td>
<td>(0.009)</td>
<td>(0.003)</td>
<td>(0.006)</td>
<td>(0.005)</td>
</tr>
</tbody>
</table>

Overall, we suggest that HF traders’ high cancellation rates are harmful for the market since such a behavior favors market volatility and the occurrence of flash crashes. The imposition of a cancellation fee is effective in reducing market volatility and to mitigate flash crashes. Nevertheless, given the positive influence of HF traders during the recovery phase, this type of regulatory policy may prevent HF traders from participating to the recovery process and it may lengthen the duration of flash crashes.

3.2.4 Financial transaction taxes

To conclude our investigation of regulatory measures, we investigate the effects of the introduction of a financial transaction tax (FTT). So far different schemes and levels of taxes have been implemented all over the world. Examples are the stamp duty in the UK, the French financial transaction tax on high-frequency trading and the pricing scheme introduced on NYSE Euronext. In this work, we assume that HF executed orders are charged a fee \( ft t > 0 \). Accordingly, HF traders will intentionally cancel their orders whenever \( \pi_E^{j,t} < ft t \cdot D_{j,t} \).

Although its recent introduction in some markets has mainly been motivated by the goal of raising revenues in response to major financial crises (IMF, 2010; Pollin, Baker, and Schaberg, 2003), financial transaction taxes have traditionally been indicated as a

\(^{22}\)Note that values are averages across 50 independent Monte-Carlo runs.
possible effective tool to discourage short-term speculation (Tobin, 1978), to curb negative effects of HFT practices and to improve systemic resilience of financial markets (Griffith-Jones and Persaud, 2012). Nevertheless, the effectiveness of a financial transaction tax is still a controversial and highly debated topic among academics (see, for instance, McCulloch and Pacillo, 2011, for a review of existing works on financial transaction taxes). On the one hand, empirical evidence on the relationship between FTT and market quality delivers mixed results (see, for instance, Roll, 1989; Umlauf, 1993; Jones and Seguin, 1997; Habermeier and Kirilenko, 2001; Hau, 2006; Gomber, Haferkorn, and Zimmermann, 2015), although some studies (e.g. Colliard and Hoffmann, 2013) find that an FTT may have a permanent positive effect on low-latency trading, due to lower order aggressiveness and fewer rapid cancellations. On the other hand, many theoretical works suggest that an FTT can have a stabilizing effect (Ehrenstein, 2002; Westerhoff, 2003, 2004; Westerhoff and Dieci, 2006). However, other theoretical works also point out that such a stabilizing role is highly dependent on some important conditions such as market liquidity (Haberer, 2004), the level of the tax (Giardina and Bouchaud, 2004; Dupont and Lee, 2007; Demary, 2010; Fricke and Lux, 2015), the structure of the market (Pelizzari and Westerhoff, 2009). Lastly, many scholars view HFT as the main providers of liquidity in modern markets (Hendershott, Jones, and Menkveld, 2011; Menkveld, 2013). In this view, a financial transaction tax would not be beneficial because it would hurt the functioning of markets and reduce market quality (Dupont and Lee, 2007).

We therefore contribute to the above debate by running Monte-Carlo experiments where we impose different levels of financial transaction tax as a percentage of HF orders’ size, and we then investigate the resulting impact on market volatility as well as on the occurrence and the duration of flash crashes. Table 7 shows the results of this experiment when $\gamma_H = 20$. This table shows that the introduction of an FTT has a beneficial impact on market stability and on the occurrence of flash crashes. When the FTT is implemented in the market, we observe a reduction in price volatility and in the number of flash crashes. Again, these positive effects come at the cost of a longer duration of flash crashes. However, the effectiveness of financial transaction taxes is much milder compared to other policy measures discussed so far (e.g. minimum resting times and cancellation fees). In particular the reductions in volatility and in the number of flash crashes with respect to the baseline are much lower than the one obtained with cancellation fees of the same level as the tax (compare results in Table 7 to the results

\footnote{For a different view see the work of Mannaro, Marchesi, and Setzu (2008).}

\footnote{Notice that we ran the above experiments for other values of $\gamma_H$ and different levels of ftt. However, simulation results are consistent with the ones presented in Table 7.}
Table 7: The effect of different transaction tax levels on price volatility and flash crash statistics when $\gamma^H = 20$. Values are averages across 50 independent Monte-Carlo runs. Monte-Carlo standard errors in parentheses. ($\sigma_P$): price returns volatility.

<table>
<thead>
<tr>
<th>$\gamma^H = 20$</th>
<th>ftt</th>
<th>$\sigma_P$</th>
<th>Number of flash crashes</th>
<th>Avg. duration of flash crashes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0%</td>
<td>0.016</td>
<td>(0.002)</td>
<td>4.636</td>
<td>(0.398) (0.484)</td>
</tr>
<tr>
<td>0.05%</td>
<td>0.010</td>
<td>(0.000)</td>
<td>3.279</td>
<td>(0.271) (0.843)</td>
</tr>
<tr>
<td>0.5%</td>
<td>0.009</td>
<td>(0.000)</td>
<td>2.697</td>
<td>(0.249) (0.800)</td>
</tr>
<tr>
<td>1%</td>
<td>0.009</td>
<td>(0.000)</td>
<td>3.094</td>
<td>(0.306) (0.690)</td>
</tr>
<tr>
<td>10%</td>
<td>0.004</td>
<td>(0.000)</td>
<td>1.429</td>
<td>(0.107) (1.147)</td>
</tr>
<tr>
<td>50%</td>
<td>0.002</td>
<td>(0.000)</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
in Table 6 with the scenario $\gamma^H = 20$). Significant improvements are obtained only with draconian tax rates (i.e. 10% or 50%, see Table 7). Moreover, the introduction of a financial transaction tax does not lead to lower HF orders’ aggressiveness, as it was the case for the introduction of a cancellation fee (see Table 6). On the contrary, with an FTT HF orders’ aggressiveness is higher than the one of LF traders, especially in the recovery phase.

The above outcomes are explained by the different mechanisms through which cancellation fees and transaction taxes transmit their effects in markets. As we discussed in Section 3.2.3, a cancellation fee encourages HF traders to keep their orders in the book. As a result, orders’ cancellation is reduced as well as orders’ aggressiveness. In contrast, an FTT boosts order cancellation by increasing the required expected profit threshold to keep an order in the book. Without a transaction tax, a HF trader has the incentive to maintain the order in the book if the expected profit is non-negative $\pi_{j,t}^F \geq 0$. In our model, with a transaction tax, an order is kept in the book if $\pi_{j,t}^F \geq ftt \cdot D_{j,t}$. Thus the higher is the transaction tax rate, the larger is the amount of HF orders removed from the book in each trading session. However, sufficiently large amounts of order cancellations (as e.g., it is the case for draconian tax rates) have the paradoxical effect of almost removing HF traders from the market, thus reducing volatility and leading flash crashes to vanish.

Overall, the above results cast doubts on the effectiveness of financial transaction taxes on HFT, especially if its validity is compared to the one other market-based measures such as cancellation fees (or command-and-control ones like minimum resting times). Indeed, besides exhibiting the same trade-off between stability and resilience already highlighted for the other policy measures, financial transaction taxes can achieve significant reductions in volatility and have some incidence on financial crashes only if they implemented at sufficiently high rates.

4 Concluding Remarks

We developed an agent-based model of a limit-order book (LOB) market based on Jacob Leal, Napoletano, Roventini, and Fagiolo (2016) to analyze the effectiveness of a set of regulatory policies on market volatility, and on the occurrence and the duration of flash crashes. In the model, low-frequency (LF) traders interact with high-frequency (HF) agents. The former can switch between fundamentalist and chartist strategies. HF traders instead employ low-latency directional strategies to exploit the order book information released by LF agents. In addition, LF trading rules are based on chronolog-
ical time, whereas HF ones are framed in event time, i.e., the activation of HF traders endogenously depends on past price fluctuations. Finally, HF traders can endogenously cancel their orders from the book based on expected profits. In this framework, we analyzed via Monte-Carlo simulations, the impact of policies like i.) trading halt facilities (both ex-post and ex-ante designs); ii.) minimum resting times; iii.) order cancellation fees; iv.) transaction taxes. These policies have been proposed and implemented both in Europe and in the US to mitigate the possible damaging effects of HFT and to prevent flash crashes.

Computer simulations reveal that, policies slowing down the order cancellation of high-frequency traders, like the implementation of minimum resting times or cancellation fees lead to significant improvements in terms of lower market volatility and incidence of flash crashes. Also the introduction of a financial transaction tax, by discouraging HFT, can improve market stability, although the effectiveness of such a measure is much lower compared to policies targeting order cancellation, and effects are relevant only for high values of the tax. These results are all consistent with the remarks in Haldane (2011), who conjectures that the above set of policies are effective because they tackle the “race to zero” of HFT at source by imposing a speed limit on trading. At the same time, all these policies are characterized by a trade-off between market stability (in terms of lower volatility and number of flash crashes) and market resilience (in terms of longer recoveries from a crash). This trade-off emerges because of the positive role played by HFT in quickly restoring good liquidity conditions after a crash. Regulatory policies introduce important distortions in such a process, thereby contributing to lengthen the duration of price-recoveries. The beneficial impact of HFT on price resilience also underlies the results concerning the study of the impact of circuit breakers, and in particular explain why ex-post circuit breakers have no effect on volatility and have a negative impact on the duration of flash crashes. In contrast, we find that ex-ante circuit breakers are very effective, as they markedly reduce price volatility and completely remove flash crashes.

Overall, our results suggest that regulatory policies can have quite complex effects on markets populated by low and high-frequency traders. From the viewpoint of policy design, our analysis highlights in particular the importance of understanding the different transmission mechanisms through which the effects of regulatory policies unfold. Moreover, it points out the need of taking into account the fundamental dual role played by high-frequency traders. On the one hand, high-frequency trading can be the source of extreme events like flash crashes by placing aggressive sell orders and removing liquidity from the market. On the other hand, it can play a leading role in the recovery from the crash, by quickly restoring liquidity.
Our analysis could be extended in several ways. First, we could enlarge the set of policies considered, by including measures such as make/take fees, restrictions on tick size, position limits. Second, so far, we have only considered one asset market in the model. However, regulatory authorities should also focus on the linkages across markets, recognizing that some coordination is needed to ensure the effectiveness of regulatory interventions (see CFTC and SEC, 2010; Furse, Haldane, Goodhart, Cliff, Zigrand, Houstoum, Linton, and Bond, 2011), especially in high frequency markets, where HF traders can rapidly process and profit from the information stemming from different exchanges (e.g., Wah and Wellman, 2013). The May 6, 2010 highlighted, for instance, the importance of the the interconnectedness of equities and derivatives markets.

Acknowledgments

We are grateful to Andrea Roventini, Giorgio Fagiolo, Sylvain Barde, Jean-Luc Gaffard, Charles Noussair, Olivier Scaillet, Nobuyuki Hanaki, Simone Alfarano and Francesco Saraceno for stimulating discussions and fruitful comments. We also thank the participants of the WEHIA conference in Sophia Antipolis, May 2015, to the FEBS conference in Nantes, June 2015, to the MAGECO 2015 workshop in Paris, November 2015, and to the PFMC conference in Paris, December 2015, and to the Workshop on “Complex Evolving System Approach in Economics”, in Sophia-Antipolis, January 2016 where earlier versions of this paper were presented. The authors also acknowledge the financial support of the EU H2020 Project “Distributed Global Financial Systems for Society” (DOLFINs) grant n. 640772. All usual disclaimers apply.

Appendix A

A.1 Stylized Facts of Financial Markets

We follow an indirect calibration approach to the validation of our agent-based model (see Windrum, Fagiolo, and Moneta, 2007, for a discussion of this approach) by checking its ability to jointly reproduce several stylized facts of financial markets with the same configuration of parameter values.

First, in line with the empirical evidence (e.g., Fama, 1970; Pagan, 1996; Chakraborti, Toke, Patriarca, and Abergel, 2011, and references therein), we find that our model generates zero autocorrelation values of price-returns (calculated as logarithmic differences, see Figure 4). In contrast to price returns, the autocorrelation functions of absolute returns display a slow decaying pattern (cf. Figure 5), thus indicating the presence of volatility clustering in our simulated data (Mandelbrot, 1963; Cont, Potters, and Bouchaud, 1997; Lo and MacKinlay, 1999).

Another widely-studied property of financial markets is the presence of fat tails in the distribution of price returns. We plot in Figure 6 the density of pooled returns across Monte-Carlo runs (stars) together with a normal density (solid line) fitted on the pooled sample. As the figure shows, the distribution of price returns significantly departs from
Figure 4: Price-returns sample autocorrelation function (solid line) together with 95% confidence bands (dashed lines). Values are averages across 50 independent Monte-Carlo runs.

Figure 5: Sample autocorrelation functions of absolute price returns (solid line) together 95% confidence bands (dashed lines). Values are averages across 50 independent Monte-Carlo runs.
Figure 6: Density of pooled price returns (stars) across 50 independent Monte-Carlo runs together with a Normal fit (solid line). Logarithmic scale on y-axis. Densities are estimated using a kernel density estimator using a bandwidth optimized for Normal distributions.

Figure 7: Complementary cumulative distribution of negative price returns (circles) together with power-law fit (dashed line). Double-logarithmic scale.
the Gaussian benchmark (Mandelbrot, 1963; Cont, 2001). Moreover, Figure 7 shows the tail of the distribution of (negative) price returns together with a power-law fit.\textsuperscript{25} In line with empirical evidence (see Lux, 2006, and references therein), the power law distribution provides a good approximation of the simulated data of tail returns.

A.2 Parameters

Table 8: Parameters values in the baseline scenario

<table>
<thead>
<tr>
<th>Description</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monte Carlo replications</td>
<td>$MC$</td>
<td>50</td>
</tr>
<tr>
<td>Number of trading sessions</td>
<td>$T$</td>
<td>1,200</td>
</tr>
<tr>
<td>Number of low-frequency traders</td>
<td>$N_L$</td>
<td>10,000</td>
</tr>
<tr>
<td>Number of high-frequency traders</td>
<td>$N_H$</td>
<td>100</td>
</tr>
<tr>
<td>LF traders’ trading frequency mean</td>
<td>$\theta$</td>
<td>20</td>
</tr>
<tr>
<td>LF traders’ min and max trading frequency</td>
<td>$[\theta_{min}, \theta_{max}]$</td>
<td>[10,40]</td>
</tr>
<tr>
<td>Chartists’ order size parameter</td>
<td>$\alpha_c$</td>
<td>0.04</td>
</tr>
<tr>
<td>Chartists’ shock standard deviation</td>
<td>$\sigma_c$</td>
<td>0.05</td>
</tr>
<tr>
<td>Fundamentalists’ order size parameter</td>
<td>$\alpha_f$</td>
<td>0.04</td>
</tr>
<tr>
<td>Fundamentalists’ shock standard deviation</td>
<td>$\sigma_f$</td>
<td>0.01</td>
</tr>
<tr>
<td>Fundamental value shock standard deviation</td>
<td>$\sigma_y$</td>
<td>0.01</td>
</tr>
<tr>
<td>Fundamental value price drift parameter</td>
<td>$\delta$</td>
<td>0.0001</td>
</tr>
<tr>
<td>LF traders’ price tick standard deviation</td>
<td>$\sigma_z$</td>
<td>0.01</td>
</tr>
<tr>
<td>LF traders’ intensity of switching</td>
<td>$\zeta$</td>
<td>1</td>
</tr>
<tr>
<td>LF traders’ resting order periods</td>
<td>$\gamma_L$</td>
<td>20</td>
</tr>
<tr>
<td>HF traders’ resting order periods</td>
<td>$\gamma_H$</td>
<td>20</td>
</tr>
<tr>
<td>HF traders’ activation threshold distribution support</td>
<td>$[\eta_{min}, \eta_{max}]$</td>
<td>[0,0.2]</td>
</tr>
<tr>
<td>Market volumes weight in HF traders’ order</td>
<td>$\lambda$</td>
<td>0.625</td>
</tr>
<tr>
<td>size distribution</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HF traders’ order price distribution support</td>
<td>$[\kappa_{min}, \kappa_{max}]$</td>
<td>[0,0.01]</td>
</tr>
</tbody>
</table>

References


\textsuperscript{25}The power-law exponent was estimated using the freely available “power-law package” and based on the procedure developed in Clauset, Shalizi, and Newman (2009).


Hanson, T. A. (2011), “The effects of high frequency traders in a simulated market”, *Available at SSRN 1918570*.


