ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Impact of Heterogeneity of Noise Trader Agents on Volatility Clustering

Master Thesis

Basile Despond

August 31, 2019

Advisors: Prof. Dr. Didier Sornette, Rebecca Westphal
Department of Physics, ETH Zürich

Abstract

The impact of the memory length of noise traders on the frequency and size of bubbles, on volatility clustering and on market endogeneity is analyzed in an agent-based model where a risky and a risk-free asset are traded by fundamentalists and noise traders. A heterogeneity of memories is later introduced and its impact quantified.

Noise traders with a 1 day memory tend to create more bubbles than all other memories, but with the lowest average drawdown. This memory produces the least volatility clustering and the least endogenous markets. Longer memory lengths reduce the number of bubbles, enlarge their drawdown, produce more volatility clustering and more endogenous markets. We show that noise traders with no memory behave as if their memory was infinite.

Introducing a heterogeneity of two memories increases the number of bubble by $25.2 \pm 0.2\%$ and decreases their drawdown by 30%, compared to a market with all the noise traders sharing the heterogeneous market average memory. Making the same comparison, such a heterogeneity decreases the volatility clustering by 36% when quantified through the decay of the autocorrelation coefficients of absolute returns. This heterogeneity decreases the market endogeneity by $4.54 \pm 0.07\%$, when computed through a non parametric approximation of the branching ratio for self-exciting Hawkes point processes. Higher heterogeneities behave closer to a market with all noise traders sharing the heterogeneous market average memory.

Acknowledgements

I would like to sincerely thank you Prof. Sornette for giving me the opportunity to be a part of your team for the last six months. Your inputs were always very constructive and your knowledge invaluable.

Also, thank you Rebecca for supervising me throughout this work. Your constructive feedback, guidance and your help with C++ were always very much helpful.

Thank you Scott, François, Ludovic for your help with stats questions, proof reading and telling me to stop working too much.

Most importantly, thank you Regina for your constant and unconditional love and support. I could not have done it without you.

Contents

Introduction

Which role plays the memory length of speculative traders on the frequency and size of financial bubbles? How much does their memory influence volatility clustering and market endogeneity? Understanding the causes of these phenomena observed in financial returns is essential to build models describing most accurately the market dynamics.

These questions are tackled by the present work with the agent-based model (ABM)* developed by Kaizoji et al. (2015). In the artificial market considered, fundamentalist traders establish their positions based on a value analysis of the market, while speculative traders herd and act on price momentum. The model is able to reproduce financial bubbles regimes as well as the main stylized facts displayed in financial markets.

The current ABM offers a natural framework to study the impact of the look back window of speculative traders on the frequency and size of bubbles. Indeed, the long standing research of the current work advisor on critical phenomena (Sornette (1998), Sornette (2006)) and their application to the prediction of financial bubbles (Sornette et al. (1996), Sornette (2017), Jiang et al. (2010), Harras and Sornette (2011)) stresses the impact of herding of the market participants on the growth and subsequent crash of financial bubbles.

Volatility clustering, the tendency of large absolute returns to be followed by large absolute returns and small absolute returns followed by small absolute returns, is present throughout many asset classes (Bollerslev et al. (1992), Diebold and Nerlove (1989)). Investor focused on the long term outlook of the market interacting with speculative traders concentrated on short term gains may be a rationale for such a behavior. Understanding its exact origin is primordial for building models describing most accurately the underlying dynamics and lessen the resulting instabilities. This stylized fact has been widely researched through computational models and offers a vast literature. Among them, LeBaron (2000) introduces multiple time scales and shows that both long and short term memory lengths of the market participants entail higher volatility. Lux and Marchesi (2000) demonstrate that volatility clustering is linked to temporary regimes of market instability in an ABM where all traders can switch between a fundamentalist or a speculative approach. LeBaron

*An ABM is a computational model to simulate the interactions of participants (agents). By varying the rules governing the agents' behaviors, analytical insight on the consequences of their interplay can be gained.

(2006) shows that at least two different time scales are necessary to produce volatility persistence. Hommes (2006), LeBaron (2006), Samanidou et al. (2007) and Cont (2007) offer a meta analysis of the models studying volatility clustering.

The systematic study of the impact of different memory lengths of the speculative traders on the frequency and size of bubbles and on the amount of volatility clustering is not present in the literature. As noted in LeBaron (2000), the challenge of complex agent-based models is to clarify which aspect of the model is responsible for the stylized facts observed. In this regard, by only varying the memory length of the speculative traders while keeping all the other market variables fixed, its impact can be quantified. Subsequently, the outcome of a hierarchical structure of memories of the speculative traders is analyzed by extending the ABM.

The next chapter introduces the market model. Chapter 2 offers a new metric to pinpoint bubble peaks out of sample. The impact of varying the memory of the noise traders on the frequency and size of bubbles is then analyzed. Chapter 3 quantifies the volatility clustering generated by 18 different memory lengths through the analysis of the decay of the autocorrelations coefficients of the absolute returns. A more innovative approach to quantify volatility clustering, via the moments of the frequency distribution of large absolute returns, is presented in chapter 4. The tools developed are then used to probe the endogeneity of the market through a non parametric approximation of the branching ratio. Then, chapter 5 extends the ABM model to allow for a heterogeneity of time scales in the investment decision of the noise traders. The metrics developed in the preceding chapters are used to analyze the impact of six levels of market heterogeneity on the frequency and size of bubbles, volatility clustering, and on market endogeneity. An appendix offers an introduction to time series analysis.

Chapter 1

The Market Model

The current chapter presents the agent-based model (ABM) developed by Kaizoji et al. (2015), composed of two assets and two types of agents. A risky asset paying a dividend and a risk free asset paying a fixed interest rate are traded by fundamentalists and noise traders. These assets and trader types are described in the following chapters.

1.1 The two assets and the dividend process

The risk free asset can be viewed as a government bond with perfect elastic supply, returning a fixed risk-free annual interest rate R_f. The risky asset, which can be considered as a stock, is defined to pay a dividend d_t at each period t. The dividend process is represented as a multiplicative stochastic growth process as in Kohrt (2016)

$$d_{t+1} \coloneqq (1 + r_t) \cdot d_t \ , \tag{1.1}$$

with the stochastic growth rate r_t being a Gaussian process with mean value $r_d \in (0, \infty)$ and variance σ_t^2

$$r_t \coloneqq r_d + \sigma_r u_t \ . \tag{1.2}$$

The random variable u_t is i.i.d. with mean zero and unit variance

$$u_t \sim \mathcal{N}(0, 1) \ . \tag{1.3}$$

The expected value and variance of d_{t+1} are then

$$\mathbb{E}_t[d_{t+1}] = d_t(1 + r_d) \tag{1.4}$$

$$\mathrm{Var}_t[d_{t+1}] = d_t^2 \sigma_r^2 \ . \tag{1.5}$$

The excess return obtained by investing in the risky asset versus in the risk free rate R_f is defined as

$$R_{\text{excess},t+1} \coloneqq R_{t+1} - R_f + \frac{d_{t+1}}{P_t} \,, \tag{1.6}$$

where P_t is the price of the risky asset at time t, and $R_{t+1} \coloneqq \frac{P_{t+1}}{P_t} - 1$ is the return of capital. As can be seen in Equation 1.6 two revenue streams contribute to the excess return of the risky asset versus the risk free rate: the return generated by the capital and the one generated by the dividend of the risky asset.

1.2 The fundamentalist trader

The fundamentalist trader is a risk adverse rational value investor investing at each time step a proportion x_t^f - called the risky fraction - of its current wealth in the risky asset, and the remaining wealth in the risk-free asset. The risky fraction is found by maximizing the expected utility of the expected wealth W_{t+1}^f under a constant relative risk aversion (CRRA) utility function. As all the fundamentalist traders have the same utility function and expectations, they are considered to be identical. Their behavior is represented by a single fundamentalist.

The evolution of the wealth can be decomposed into a risky and risk free component. Considering the current wealth W_t^f, the price P_{t+1} of the risky asset, the fraction of the wealth invested in the risky asset x_t^f and in the risk free asset $(1 - x_t^f)$, the wealth of the fundamentalist at the next time step W_{t+1}^f is

$$\begin{aligned} W_{t+1}^f &= W_t^f + \underbrace{R_f(1 - x_t^f)W_t^f}_{W_t^{\text{risk free}}} + \underbrace{\Big(\frac{P_{t+1} - P_t}{P_t} + \frac{d_{t+1}}{P_t}\Big)x_t^f W_t^f}_{W_t^{\text{risky}}} \\ &\overset{1.6}{=} W_t^f + W_t^f[R_f + x_t^f R_{\text{excess},t+1}] \,. \end{aligned} \tag{1.7}$$

To determine the risky fraction, the fundamentalist maximizes the utility of its expected wealth at the next period

$$\max_{x_t^f} \mathbb{E}_t[U(W_{t+1}^f)] \,, \tag{1.8}$$

where the utility function $U(W)$ is defined as

$$U(W) = \begin{cases} \log(W) & \text{for } \gamma = 1 \\ \dfrac{W^{1-\gamma}}{1-\gamma} & \text{for } \gamma \neq 1 \end{cases} \tag{1.9}$$

with the constant relative risk aversion parameter γ

$$\gamma := -\frac{W_t^f U''(W_t^f)}{U'(W_t^f)} \,. \tag{1.10}$$

Expanding $\mathbb{E}_t[U(W_{t+1}^f)]$ in a Taylor expansion and using Equation 1.7 gives, neglecting higher order terms,

$$\begin{aligned}\mathbb{E}_t[U(W_{t+1}^f)] &= \mathbb{E}_t[U(W_t^f) + U'(W_t^f)(W_{t+1}^f - W_t^f) + \frac{1}{2}U''(W_t^f)(W_{t+1}^f - W_t^f)^2] \\ &= \mathbb{E}_t[U(W_t^f) + U'(W_t^f)W_t^f(R_f + x_t^f R_{\text{excess},t+1}) \\ &\quad + \frac{1}{2}U''(W_t^f)(W_t^f)^2(R_f + x_t^f R_{\text{excess},t+1})^2] \,.\end{aligned} \tag{1.11}$$

The wealth W_t^f is assumed to not change much over one time step. Thus $R_f, R_{\text{excess},t+1} \ll 1$ holds and it follows

$$\begin{aligned}\mathbb{E}_t[U(W_{t+1}^f)] &= U(W_t^f) + U'(W_t^f)W_t^f(R_f + x_t^f \mathbb{E}[R_{\text{excess},t+1}]) \\ &\quad + \frac{1}{2}U''(W_t^f)(W_t^f)^2 \underbrace{\mathbb{E}[(R_f + x_t^f R_{\text{excess},t+1})^2]}_{\text{Var}_t[R_{\text{excess},t+1}]} \,.\end{aligned} \tag{1.12}$$

The last expected value can be rewritten as a variance due to the above assumption of small returns. Maximizing this equation with regards to x_t^f yields

$$x_t^f = \frac{1}{\gamma}\frac{\mathbb{E}_t[R_{\text{excess},t+1}]}{\text{Var}_t[R_{\text{excess},t+1}]} \,, \tag{1.13}$$

In Equation 1.13, the numerator expresses the expected value of the excess return of capital (and dividend) over an investment in the risk free rate. The fundamentalists are considered to be myopic traders that do not learn and only invest w.r.t. the fundamental valuation. Based on the long term behavior of stock markets, they expect a constant growth rate. This implies that the expected value of the price return $\mathbb{E}_{R_t} := \mathbb{E}_t[R_{t+1}]$ is constant. It follows

$$\mathbb{E}_t[R_{\text{excess},t+1}] = \mathbb{E}_{R_t} - R_f + \frac{\mathbb{E}_t[d_{t+1}]}{P_t} \overset{1.4}{=} \mathbb{E}_{R_t} - R_f + \frac{d_t(1+r_d)}{P_t} = \text{constant} \,. \tag{1.14}$$

As is discussed in Boswijk et al. (2007) and Chiarella et al. (2009), the variance can be assumed to be constant as long as the fundamentalist stays myopic,

$$\mathrm{Var}_t[R_{\mathrm{excess},t+1}] \coloneqq \sigma^2_{\mathrm{excess}} \,. \tag{1.15}$$

The risky fraction is then, as presented in Ollikainen (2016),

$$x^f_t = \frac{\mathbb{E}_{R_t} - R_f + \frac{d_t}{P_t}(1+r_d)}{\gamma\sigma^2_{\mathrm{excess}}} = x^f_{min} + \frac{d_t}{P_t}\frac{1+r_d}{\gamma\sigma^2_{\mathrm{excess}}} \,, \tag{1.16}$$

with the minimum risky fraction

$$x^f_{min} \coloneqq \frac{\mathbb{E}_{R_t} - R_f}{\gamma\sigma^2_{\mathrm{excess}}} \,. \tag{1.17}$$

Equation 1.16 shines a light on the behavior of the fundamentalist trader: a minimal fraction of his wealth is invested in the risky asset at each time step regardless of the market condition. This is based on his long term view of the market and is not influenced by short term market fluctuations. Moreover, the fundamentalist analyzes the dividend-price ratio at each time step to see if the market currently under- or overvalues the risky asset. In case of a high $\frac{d_t}{P_t}$ (low price) the fundamentalist buys more of the risky asset, and sells it at a low $\frac{d_t}{P_t}$ (high price).

1.3 The noise trader

The second agent represents the "noise traders". The term was coined by Kyle (1985) to describe irrational market participants whose actions are not based on a fundamental analysis. Modeling their impact on financial bubbles and crashes was first done by de Long et al. (1990) where they showed that fundamentalist actions set off positive feedback trading by noise traders leading to bubble regimes and excess volatility. In their model, this generates positive auto-correlation of returns over a length determined by the look back period of noise traders. Campbell and Kyle (1993) have also shown that noise traders interactions with fundamentalist move the price away from its fundamental value and may explain the excess volatility.

In the present model, each individual noise trader has a bipolar nature: he is either fully invested in the risky asset and not at all in the risk free asset or the opposite. His investment decision relies on social influence and historical price movements. In this sense, he represents the lack of diversification puzzle (Statman (2004)), the trend-following (Hurst et al. (2017)) and herding mentality (Bikhchandani and Sharma (2000)) observed in the financial market. As opposed to the fundamentalists that all share the same utility function and have the same expectations, each individual noise trader may behave differently.

The number of noise traders is referred to as N_n. At time t, $N^+_t \in [0, N_n]$ are fully invested in the risky asset and $N^-_t \coloneqq N_n - N^+_t$ in the risk free asset. At each time step, every trader can independently and probabilistically switch from one asset class to the other. The switching probabilities are influenced by the social interaction, which is divided in two categories: the price momentum and the imitating behavior of the trader. These two factors are now detailed.

The price momentum is defined as an exponential moving average of past price returns

$$H_{t+1} = \theta H_t + (1-\theta) R_{t+1} \, . \tag{1.18}$$

The higher the smoothing factor $\theta \in [0,1)$, the longer the trader memory τ

$$\tau \sim {}^{1}/_{(1-\theta)} \, . \tag{1.19}$$

Throughout the next chapters, the impact of τ on the frequency and size of bubbles and on the volatility clustering will be analyzed.

To quantify the imitating behavior, the opinion index

$$s_t := \frac{N_t^+ - N_t^-}{N_n} \in [-1,1] \tag{1.20}$$

is introduced. It represents the average opinion among all noise traders w.r.t. the risky asset. For $s_t > 0$ an overall bullish spirit reigns, whereas bearishness prevails for $s_t < 0$.

The strength of the social interaction is represented by the time dependent coupling strength κ_t, also named herding propensity. It can either be constant, or modeled as an Ornstein-Uhlenbeck (OU) process which per definition satisfies the stochastic differential equation

$$\kappa_{t+1} = \kappa_t + \eta_\kappa(\mu_\kappa - \kappa_t) + \sigma_\kappa v_{t+1} \, , \tag{1.21}$$

with $\eta_\kappa > 0$ the mean reversion rate, μ_κ the mean reversion level and σ_κ the step size of the Wiener process generated by the i.i.d. random variables $v_{t+1} \sim \mathcal{N}(0,1)$.

The probabilities to switch from the risky to the risk free asset (p_t^+) and from the risk free to the risky asset (p_t^-) for the interval $(t, t+1)$ are

$$p_t^\pm := \frac{p^\pm}{2}(1 \mp \kappa_t(s_t + H_t)) \, , \tag{1.22}$$

where $p^\pm \in [0,1]$ controls the average holding time of each asset class when neither imitation nor trend following occurs ($\kappa_t = 0$). In this case, a position has an expected length of $\sim {}^{2}/_{p^\pm}$.

Harras and Sornette (2011) and Kaizoji et al. (2015) have shown that the dynamics of such noise traders can be understood as an Ising model. Disorganized phases occur when the opinions (to buy or sell the risky asset) of all the individual noise traders cancel each other out, leading to a price path based on the fundamental value. The organized phases are due to all noise traders having the same opinion due to a positive feedback mechanism, sustaining either a super-exponential growth of the risky asset or its collapse. The presence of this spontaneous symmetry breaking and the proximity to the critical point has been shown to have an influence on the excess volatility (Harras et al. (2012)).

The critical point in the present model is embodied by the average holding time p. In this regard, the interesting case occurs when the mean reversion rate is just below the average holding time: $\mu_k \lesssim p^{\pm}$, allowing the social interaction κ_t to enter the regime of super-exponential growth $\mu_k > p^{\pm}$ due to the stochastic nature of the OU process. As an OU process is a stationary Gaussian process, it is fully described by its first two moments, which converge over long periods so that κ_t is specified by

$$\kappa_t \sim \mathcal{N}\Big(\mu, \frac{\sigma_\kappa}{\sqrt{2\eta}}\Big) \,. \tag{1.23}$$

The expected time needed Δt for the social imitation strength to exit the critical regime (in which $\kappa_t \equiv \kappa_0 > p$) and revert to the subcritical regime ($\kappa_t < p$) can be derived from the first moment. Citing Kaizoji et al. (2015),

$$\Delta T = \frac{1}{\eta} \log\Big(\frac{\kappa_0 - \mu_\kappa}{p - \mu_\kappa}\Big) \,. \tag{1.24}$$

How does the system evolve in time? Let us first look at the number of noise traders invested in the risky asset at time t, N_t^+. At time t, each individual noise trader belonging to this group has the possibility to stay invested in the risky asset, or sell all of it to be completely invested in the risk free asset. For a noise trader k in N_t^+, their decision to switch or not is embodied in the i.i.d. Bernoulli random variable $\xi_k(p_t^+)$, with takes the value 1 (sell the risky asset, buy the risk free one) with probability p_t^+, and value 0 (stay invested in the risky asset) with probability $(1 - p_t^+)$.

The situation is similar for the noise traders invested in the risk free asset at time t, N_t^-. The behavior of a noise trader j that does not own the risky asset at time t is represented by $\xi_j(p_t^-)$, which takes the value 1 (switch to a risky investment) with probability p_t^-, and the value 0 (stay invested in the risk free asset) with probability $(1 - p_t^-)$.

Combining these behaviors results in

$$\begin{aligned} N_{t+1}^+ &= \sum_{k=1}^{N_t^+} [1 - \xi_k(p_t^+)] + \sum_{j=1}^{N_t^-} \xi_j(p_t^-) \\ N_{t+1}^- &= \sum_{k=1}^{N_t^+} \xi_k(p_t^+) + \sum_{j=1}^{N_t^-} [1 - \xi_j(p_t^-)] \,. \end{aligned} \tag{1.25}$$

Here, even though each noise trader can act independently from all the other noise traders, we are not interested in describing their discrete behavior. Rather, we aim to derive their cumulative impact. We thus consider a representative noise trader agent with wealth W_t^n, with each noise trader having an equal portion of it.

From a statistical mechanics point of view, each individual noise trader can be seen as a microstate defining the macrostate x_t^n, the representative noise trader agent risky fraction

$$x_t^n := \frac{N_t^+}{N_t^+ + N_t^-} = \frac{N_t^+}{N_n} \in [0,1] \,. \tag{1.26}$$

This translates to a continuous macroscopic investment weight. We see that x_t^n represents the fraction of their total wealth invested in the risky asset. This allows to gauge their bullishness on the market, as did the opinion index 1.20.

The risky fraction evolves as

$$x_{t+1} = \frac{1}{N_n}\left(\sum_{k=1}^{N_t^+}[1-\xi_k(p_t^+)] + \sum_{j=1}^{N_t^-}\xi_j(p_t^-)\right) . \tag{1.27}$$

Equation 1.27 and Equation 1.22 show that the noise traders do not take the price at time $t+1$ into consideration, which differs strongly from the approach of the fundamentalist (Equation 1.16).

The total wealth of the noise trader agent at the next time step W_{t+1}^n, considering the current wealth W_t^n, the price P_{t+1}, the fraction of his wealth invested in the risky asset x_t^n and in the risk free asset $(1-x_t^n)$ can be derived very similarly to Equation 1.7. We have

$$\begin{aligned} W_{t+1}^n &= W_t^n + R_f(1-x_t^n)W_t^n + \Big(\frac{P_{t+1}-P_t}{P_t} + \frac{d_{t+1}}{P_t}\Big)x_t^n W_t^n \\ &\overset{1.6}{=} W_t^n + W_t^n[R_f + x_t^n R_{\text{excess},t+1}] \,. \end{aligned} \tag{1.28}$$

1.4 The equilibrium market price

The price of the risky asset P_{t+1} is defined by the market clearing condition: the total excess demand summed over the noise and fundamentalist traders must vanish. We thus postulate a perfect balance of supply and demand, where the buy orders of the noise traders compensate the sell orders of the fundamentalists.

The number of shares of the risky asset owned at time t is

$$n_t^i := \frac{x_t^i W_t^i}{P_t} \,, \tag{1.29}$$

where $i \in \{f, n\}$ represents both types of traders.

The excess demand of the risky asset for trader i from time t to $t+1$ is then

$$\begin{aligned}\Delta D^i_{t\to t+1} &= n^i_{t+1} - n^i_t \\ &= \frac{x^i_{t+1}W^i_{t+1}}{P_{t+1}} - \frac{x^i_t W^i_t}{P_t} \\ &\overset{1.7,1.28}{=} W^i_t\left\{\frac{x^i_{t+1}}{P_{t+1}}(1+R_f+x^i_t R_{\text{excess},t+1}) - \frac{x^i_t}{P_t}\right\}. \end{aligned} \tag{1.30}$$

The market clearing condition can now be written as

$$\Delta D^n_{t\to t+1} + \Delta D^f_{t\to t+1} = 0\,. \tag{1.31}$$

Plugging 1.30 in 1.31 and substituting for the fundamentalist risky fraction with 1.16 gives a quadratic equation for the equilibrium market price P_{t+1}

$$a_{t+1}P^2_{t+1} + b_{t+1}P_{t+1} + c_{t+1} = 0\,, \tag{1.32}$$

with the parameters (as in Ollikainen (2016))

$$a_{t+1} = \frac{1}{P_t}\left[\nu^{nf}_t x^n_t(x^n_{t+1}-1) + x^f_t\left(x^f_{min}-1\right)\right] \tag{1.33}$$

$$b_{t+1} = x^f_t\frac{1}{\gamma\sigma^2_{R_{ex}}}\frac{d_{t+1}(1+r_d)}{P_t} + x^f_{min}\left[x^f_t\left(\frac{d_{t+1}}{P_t}-R_f\right)+R_f\right] + \nu^{nf}_t x^n_{t+1}\left[x^n_t\left(\frac{d_{t+1}}{P_t}-R_f\right)+R_f\right] \tag{1.34}$$

$$c_{t+1} = \frac{d_{t+1}(1+r_d)}{\gamma\sigma^2_{R_{ex}}}\left[x^f_t\left(\frac{d_{t+1}}{P_t}-R_f\right)+R_f\right], \tag{1.35}$$

where

$$\nu^{nf}_t := \frac{W^n_t}{W^f_t} \tag{1.36}$$

defines the wealth ratio. Equation 1.32 has the two solutions

$$P^{\pm}_{t+1} = \frac{-b_{t+1} \pm \sqrt{b^2_{t+1} - 4a_{t+1}c_{t+1}}}{2a_{t+1}}\,. \tag{1.37}$$

We refer to (Conti, 2018, p. 16-18) for the derivation of the signs of the parameters. It follows that the only solution yielding positive prices is

$$P_{t+1} = \frac{-b_{t+1} - \sqrt{b_{t+1}^2 - 4a_{t+1}c_{t+1}}}{2a_{t+1}} = \frac{b_{t+1} + \sqrt{b_{t+1}^2 + 4|a_{t+1}|c_{t+1}}}{2|a_{t+1}|} \,. \tag{1.38}$$

The causal flow for the new price P_{t+1} goes as follows. First, the noise traders update their risky fraction based on price momentum H_t and the imitating behavior s_t, but independently of the price P_t. Simultaneously, the fundamentalists update their risky fraction based on maximization of the CRRA utility function, subject to the price P_t, thus taking part in a Walrasian auction. Secondly, the noise and the fundamentalist traders give their risky fraction to the market. With this information, the new price P_{t+1} is fixed through Equation 1.38.

1.5 Numerical parameters of the simulations

The numerical values chosen for the simulations are taken from in Kaizoji et al. (2015) and Westphal and Sornette (2019) and listed in Table 1.1. Notably, the parameters r_f, d_0, r_d, σ_d, P_0 and σ_r are chosen such that one time step corresponds to 1 day. A simulation length of 20000 time steps represents 80 trading years, and 500000 time steps 2000 trading years. The asymmetry between p^+ and p^- (Equation 1.22) represents the fact that there are more positive bubbles than negative ones in real markets. Their values (≈ 0.2) infer an average holding time of each asset class of 10 days, when neither imitation nor trend following occurs ($\kappa_t = 0$). The parameter κ_t (Equation 1.23) is chosen close to $p^\pm$ allowing the social interaction κ_t to enter the regime of super-exponential growth $\kappa_k > p^\pm$ due to the stochastic nature of the OU process. The mean reversion rate η_κ and the step size of the Wiener process σ_κ are chosen such that a deviation of $2\sigma_\kappa$ of κ_t above μ_k will revert from the supercritical regime in $\Delta T = 21$ days (Equation 1.24). Both representative agents have the same initial wealth and the same initial fraction of their wealth invested in the risky asset ($x_0^n = x_0^f = 0.3$).

An snapshot of the simulation for seed 113641 and with the noise traders memory set to 610 days is shown in Figure 1.1. The top panel presents the price path between the time steps 5000 and 10000, corresponding to 20 trading years. The fourth panel displays the evolution of the risky fraction of the noise traders x_t^n and of the risky fraction of the fundamentalist x_t^f. The panel shows that the regimes where x_t^n displays a rapid growth are linked to the risky asset being more highly valued (3rd panel displaying the dividend price ratio d_t/P_t), and a decrease of the fraction of the wealth of the fundamentalist invested in the risky asset x_t^f.

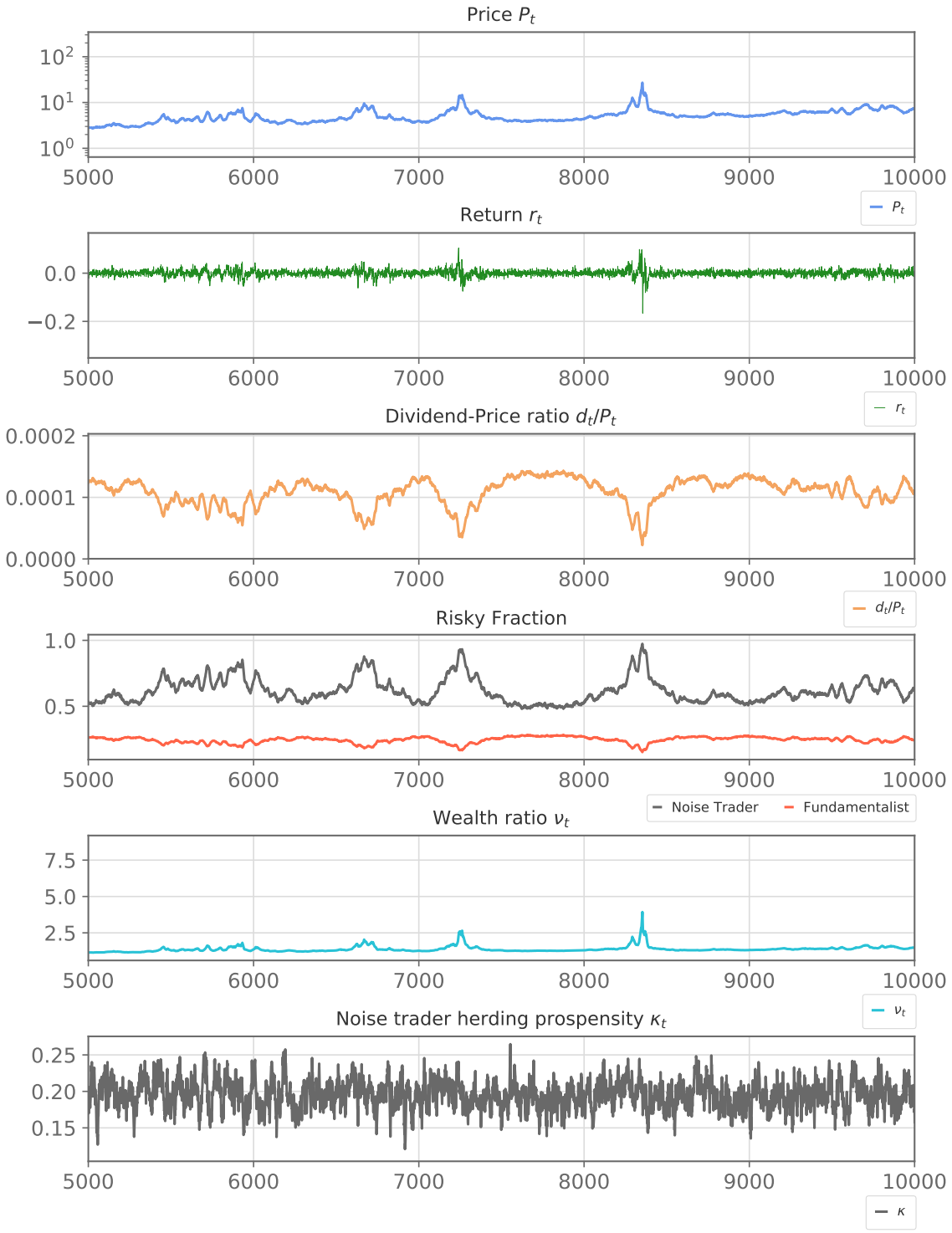


Figure 1.1: Overview of the simulation with seed 113641 and with the noise traders memory set to 610 days. The top panel presents the price path between time steps 5000 and 10000, corresponding to 20 trading years. The fourth panel displays the evolution of the risky fraction of the noise traders x_t^n and of the risky fraction of the fundamentalist x_t^f. This panel shows that the regimes where x_t^n displays a rapid growth are linked to the risky asset being more highly valued (3rd panel displaying the dividend price ratio d_t/P_t), and a decrease of the fraction of the wealth of the fundamentalist invested in the risky asset x_t^f. The lower panel shows the coupling strength κ_t modeled as an Ornstein-Uhlenbeck process satisfying Equation 1.21, with the numerical values of the parameters given in Table 1.1.

Parameter name	Explanation	Value
Market		
T	Simulation length	20000, 500000
seed	Random seed	6 digit number
Assets		
r_f	Risk free interest rate	0.01/250=0.00004
d_0	Initial dividend	0.04/250=0.00016
r_d	Expected growth rate of the dividend	0.04/250=0.00016
σ_d	Expected standard deviation of the dividend growth rate	0.000016
P_0	Initial price of the risky asset	1
σ_r	Expected standard deviation of the risky asset price	$\sqrt{0.10/250} = 0.02$
N^r	Number of risky assets	1
Noise traders		
x_o^n	Initial fraction of the risky asset held by the noise traders	0.3
W_0^n	Initial wealth of the noise traders	10^9
p_+	Switching probability when holding the risky asset	0.199375
p_-	Switching probability when not holding the risky asset	0.200625
$\boldsymbol{\theta}$	**Memory parameter**	**0.8, variable**
H_0	Initial momentum	0.00016
N^n	Number of noise traders	5000
N_G^n	Number of groups of noise traders	1
Fundamentalists		
x_0^f	Initial fraction of the risky asset held by the fundamentalists	0.3
W_0^f	Initial wealth of the fundamentalists	10^9
E_{R_t}	Expected return of the risky asset	0.00016
Social coupling strength		
κ_0	Initial social coupling strength	$0.98 \cdot 0.199375$
μ_κ	Mean of the OU social coupling strength	$0.98 \cdot 0.199375$
η_κ	Mean reversion of the OU social coupling strength	0.11
σ_κ	Standard deviation of the OU social coupling strength	0.001

Table 1.1: Parameters used in the simulations of the ABM. The values of r_f, d_0, r_d, σ_d, P_0 and σ_r (listed in the the Assets row) are given as per time step. Their values imply that each time step corresponds approximately to 1 day. A simulation length of 20000 time steps represents thus 80 trading years, and 500000 time steps 2000 trading years. The impact of the memory parameter of the noise traders (bold) will be extensively analyzed in the following chapters

Chapter 2

Impact of the Noise Traders' Memory on the Frequency and Size of Bubbles

In a financial bubble, the expectations of future returns and social influence shadow the assets real value in the decision making process of participants. Kaizoji et al. (2015) showed the influence of social interaction on the price formation in an agent based model formed of fundamentalists and noise traders, as described in chapter 1. Within this framework, we demonstrate the impact of the memory τ (1.18) of the noise trader agent on the frequency and size of bubbles. To do so, 18 different memory lengths are chosen. Their values, listed in Table 2.1, follow the Fibonacci sequence. This choice has the advantage to produce a converging scaling between two successive lengths.

The metric used in Westphal and Sornette (2019) to detect bubbles out of sample is mentioned. A more representative metric is then introduced. The robustness and stability of the new metric parameters are investigated by generating 1000 simulations of 20000 time steps for each memory τ. Section 2.5 presents the results for simulations of 495000 time steps computed with the robust parameters of the metric.

2.1 Identification of peaks and valleys and computation of drawdowns

In Westphal and Sornette (2019), a peak occurs at the time steps t_i if the price P_{t_i} is the maximum of the 250 previous and future prices

$$P_{t_i} \geq P_{t_j} \quad \forall t_j \in [t_i - 250, t_i + 250] \, . \tag{2.1}$$

This insures a minimum of 250 days between two consecutive peaks. However, this does not imply that a peak is found at every 500 time steps. If the price shows an upwards trend, there will be a higher price in the next 250 time steps for all prices leading to the real peak however long the upward trend is. Yet, in a flat market this approach will detect peaks that are not part of a real

*Equation 1.19 does not yield a finite value in this case. Thus the price momentum defined in Equation 1.22 is set to $H_t = 0$.

Memory length in days τ	**Value of the memory parameter** θ **in Table 1.1**
0	$H_t = 0$ *
1	0
2	0.5
3	0.666666666666667
5	0.8
8	0.875
13	0.923076923076923
21	0.952380952380952
34	0.970588235294118
55	0.981818181818182
89	0.98876404494382
144	0.993055555555556
233	0.995708154506438
377	0.997347480106101
610	0.998360655737705
987	0.998986828774063
1597	0.999373825923607
2584	0.999613003095975

Table 2.1: List of the 18 different memory parameters τ (Equation 1.19) used in this chapter, chapter 4 and chapter 3. The numbers of days corresponding to the value of τ follow the Fibonacci numbers, producing a converging scaling between two successive memory lengths. The numbers for θ are given with a 15 digits precision such that the program rounds the value of τ, obtained by Equation 1.19, to the value of τ listed here.

bubble but just local maxima of the flat market. Using this metric results in counting too many peaks and yielding an average drawdown per peak not representative of the drawdown following the real bubbles in the time series studied. Comparing the results from many simulations using such a metric is coherent, as long as no absolute numbers is presented such as the average number of days between two consecutive bubbles. Also, arguing that a model yields a similar average drawdown as bubbles do in real markets does not hold as the metric skews the drawdown of the true bubbles generated by the model.

In this regard a metric a_t to detect peaks out of sample is proposed. A bubble is defined as an abnormal rise and a following drawdown. The value of the metric a_t is computed as

$$a_{t_i}(s) = \frac{1}{P_i}\frac{1}{2s}\sum_{j=1}^{s}(2P_i - P_{i-j} - P_{i+j}) \quad \forall t_i \in [s, T-s] \,, \tag{2.2}$$

for a price time series of length T, and a maximal bubble length s . The factor $\frac{1}{P_i}$ allows to interpret a_{t_i} as the average percentage difference of the price P_i w.r.t. its s left and s right neighbors. In addition, the metric a_t scales with the price and does not generate more bubbles in latter years.

The metric a_{t_i} is subsampled by keeping only the time steps t_i that have a value higher than a given threshold *thresh*. Subsequently, a_{t_i} is subsampled again such that the minimum number of days between two bubbles is 250 trading days, as in Westphal and Sornette (2019). With the definition of a_t, the parameter *thres* represents the minimal percentage difference between the detected peak and the average price of its $2s$ neighbors. The metric proposed is related to a negative acceleration. A peak has a maximum downward curvature (maximum negative acceleration), which is quantified through a_t (see Ardila et al. (2015) for the stronger explanatory power of the acceleration versus the momentum)

As in Westphal and Sornette (2019), the troughs are defined as the minimum price between two consecutive peaks, and the drawdown between a peak and the following trough is computed as

$$d_{t_i} = \log(P_{t_{peak}=t_i}) - \log(P_{t_{valley}}) \ . \tag{2.3}$$

2.2 Robustness of the metric

The implications and robustness of the parameter s on the number of bubbles generated by the metric presented in Equation 2.2 are investigated on time series of 20000 time steps.

A bubble is detected at time t_i if the metric a_{t_i} is the local maximum on the range $[P_{t_i} - s, P_{t_i} + s]$ and the price P_{t_i} is the maximum over the range $[t_i - 250, t_i + 250]$ (minimum distance between two bubbles fixed at $k = 250$ days). The parameter s can thus be interpreted as a cut-off for the maximal length of a bubble. As will be shown, this parameter has a significant impact on the number of bubbles detected. As no more than one bubble can occur in 250 days, $s \leq 250$.

A first problem due to a small s can be seen for $s = 50$ in Figure 2.1: at least two clear bubbles between time steps 2500 and 5000 as well as one bubble around time-step 10500 are not detected. Figure 2.2 is a zoom of Figure 2.1 with the value of a_t plotted in green. With a $s = 50$, each price is compared to its 50 left and right neighbors. The bubble around time-step 10500 takes more than 50 days to form and explode, hence a_t has a lower value. From the construction of a_t, a small s is not equipped to detect bubbles forming over periods larger than s. This impacts the number of bubbles, the average time between bubbles and the average drawdown per bubble. If the bubbles' build up and consequent crash were symmetric, the average drawdown would not be affected. However the evolution of the bubbles generated by the present ABM is found to be asymmetric[†].

A second artifact of a small s is revealed when n bubbles occur in a row. Then, only the peak preceding the largest of the n drawdowns will be defined as the peak of the bubble regime. The peak so defined is not representative of the true maximum price of the bubble regime. Such a

[†]As can be deduced from the asymmetric switching probabilities, which comes from the impact of the dividend on the price momentum.

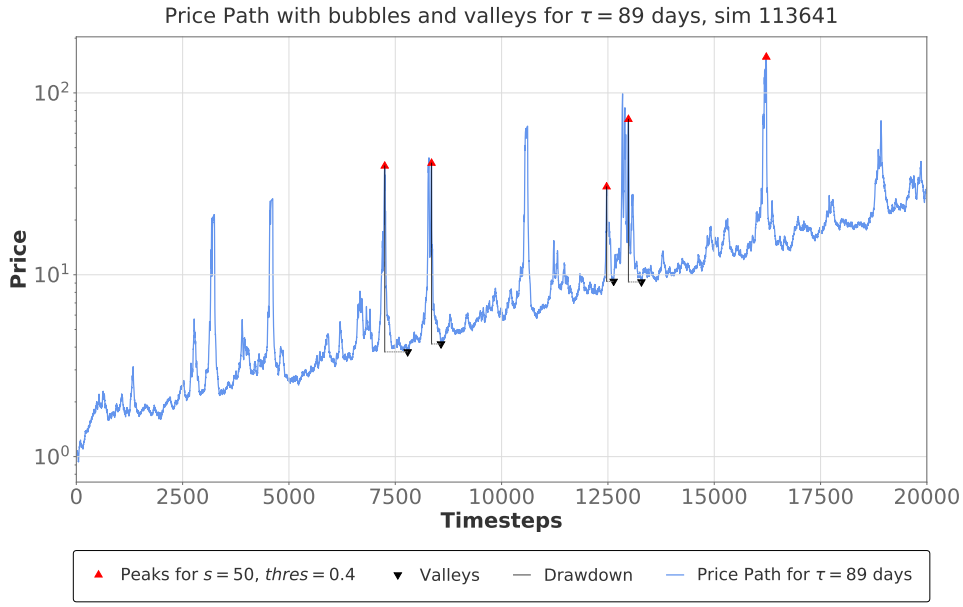


Figure 2.1: A price path generated with the noise traders' memory set to 2 days. The peaks and valleys are selected with $thres = 40\%$, $s = 50$ days and the minimal amount of time steps between two consecutive bubbles $k = 250$. Clear bubbles between time steps 2500 and 5000 as well as one bubble around time-step 10500 are not detected. Note also the bubble regime around time-step 13000 where only the last and lowest bubble peak is counted. The simulation is run for the values listed in Table 1.1, with the random seed 113641 and the memory parameter set to 89 days.

behavior can be seen Figure 2.1 and Figure 2.2 in the bubble regime around time-step 13000. First, note that this bubble regime has 3 peaks. As they all occur within 250 days, a maximum of one bubble can be counted. Secondly, we see on Figure 2.2 that there is a relatively smaller drawdown following the first peak compared to the drawdown of the rightmost peak of this bubble regime (pinpointed with a red triangle). As a_{t_i} is the average percentage difference of the price P_i w.r.t. its s left and s right neighbors, its value is smaller for the first two peaks than for the third.

Let us now compare these results with the similar price path, same threshold $thres = 0.4$ but $s = 250$ (the maximum allowed value for a maximum of one bubble every 250 days, and the value used in Westphal and Sornette (2019)). Figure 2.3 and the corresponding zoom in Figure 2.4 show that a higher value for s resolves both mentioned artifacts. We conclude that for a given threshold, a larger s will yield more bubbles and more importantly will be more representative.

The parameters selected in Table 1.1 imply that the price will revert from a deviation to the supercritical regime in $\Delta T = 21$ days. On symmetry ground, it can be argued that the minimum duration of a bubble is 42 days. However, through the herding behavior of the noise traders, organized phases tend to last longer.

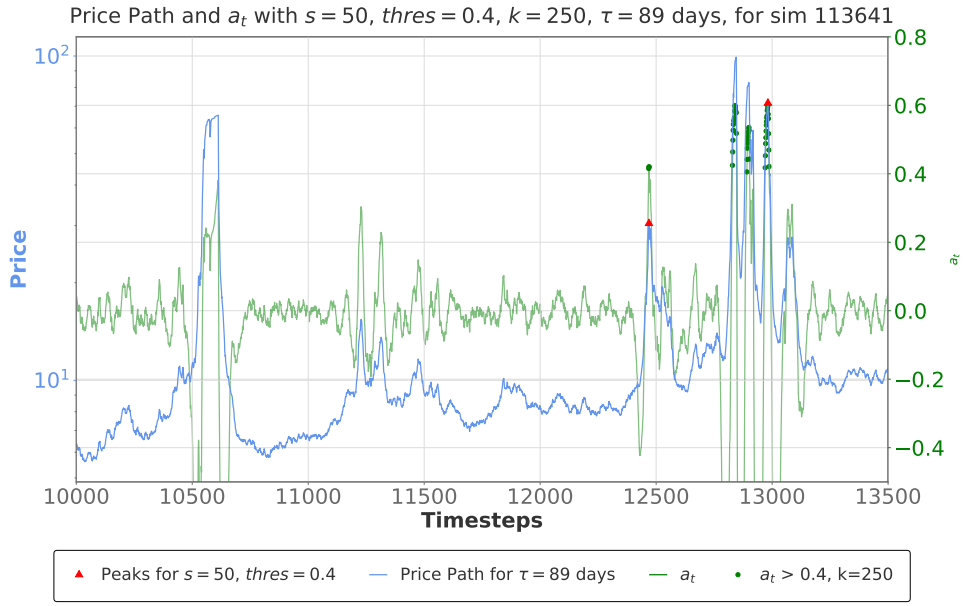


Figure 2.2: Zoom of Figure 2.1. The value of a_t is computed as defined in Equation 2.2. The values of a_t above the threshold $thres = 0.4$ are highlighted. As described in the text, the value of a_t is on average lower for smaller s, which makes it miss bubbles forming over longer periods than s time steps. The simulation is run for the values listed in Table 1.1, with the random seed 113641 and the memory parameter set to 89 days.

2.3 Impact of the metric on the bubble count and drawdown sizes

We demonstrated in the previous section the importance and impact of the parameter s on the number of peaks detected. To quantify this impact, we fix $thres = 0.4$ and compute the average number of peaks for $s \in \{50, 100, 150, 200, 250\}$ over 1000 different price paths for each memory presented in Table 2.1. Each price path comprising of 20000 time steps. We thus have an average number of bubbles for the 18 different memory lengths and the 5 different s values. For each value, the error is computed as the standard error of the mean as we are in a counting process. The results presented in Figure 2.5 show the statistically significant impact of s. Except for a memory of 0 day, the shortest memory lengths seem to create the most bubbles. Numerical values will be presented and commented in section 2.5.

As shown in section 2.2, only bubbles narrower than a given s days will be detected at all. This influences the average drawdown per bubble: for small s only very narrow and thus steep bubbles will be detected. Smaller, slower forming bubbles are not detected which artificially yields a higher average drawdown. Moreover, a small s will misattribute the true peak of a large bubble regime. These two effects have thus an impact on the average drawdown per peak. The results are presented in Figure 2.6, where the drawdowns are computed as in Equation 2.3. Overall, a memory length of 1 day seems to create the smallest drawdown. For each value, the error bar is computed as the standard deviation of the distribution of the log drawdown. Even though the difference between

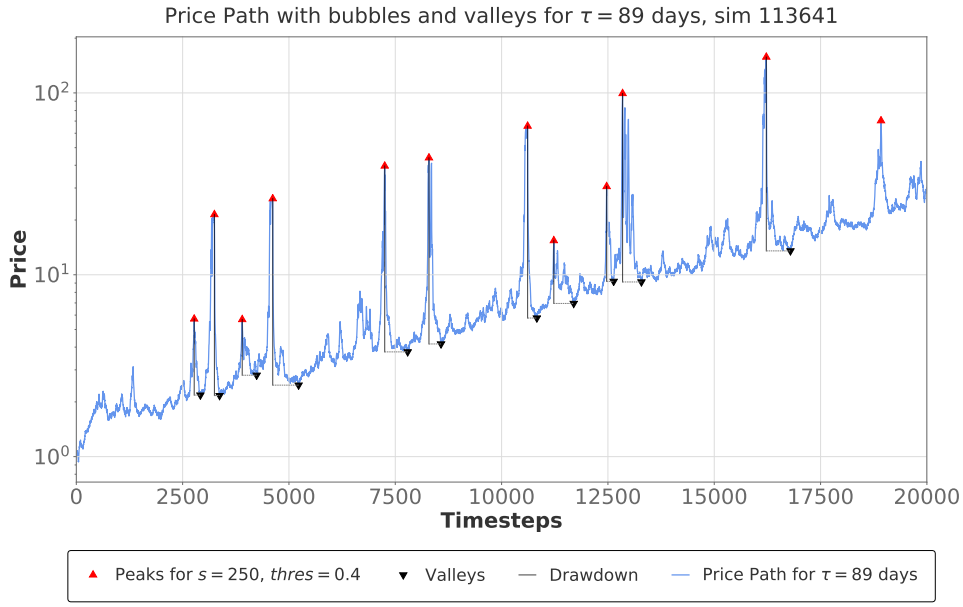


Figure 2.3: A price path generated with the noise traders' memory set to 89 days. The peaks and valleys are selected with $k = 250$ days, $thres = 40\%$, and $s = 250$ days. In comparison with the $s = 50$ case in Figure 2.2 where 5 bubbles were detected, 12 bubbles are detected here. Also, the peak of the bubble regime around time-step 13000 is more precisely pinpointed. The simulation is run for the values listed in Table 1.1, with the random seed 113641 and the memory parameter set to 89 days.

different s are not statistically significant, we note a systematic clear and non-overlapping trend. The standard deviation is thus too pessimistic.

2.4 Impact of the threshold parameter on the bubbles count and drawdown sizes

With our redefinition of Equation 2.2, interpreting the parameter *tresh* is straightforward: it represent the minimal percentage difference between the detected peak and its $2s$ neighbors. Setting $thres = 0.4$ selects the bubbles whose peak has a price at least 40% higher than the average price of its $2s$ neighbors. Thus, a higher *thres* yields less bubbles but with a higher drawdown average.

Figure 2.7 demonstrates the impact of the threshold parameter on the average number of bubbles, where the parameter s is set to 250 days. We proceed as in section 2.3 to quantify the impact of the *thres* parameter. We fix $s = 250$ and compute the average number of peaks for $thres \in \{0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6\}$ over 1000 different price paths for each memory presented in Table 2.1, with each price path comprising of 20000 time steps. We thus have an average number of bubbles for the 18 different memory lengths and the 11 different *thres* values. For each value, the error is computed as the standard error of the mean as we are in a counting

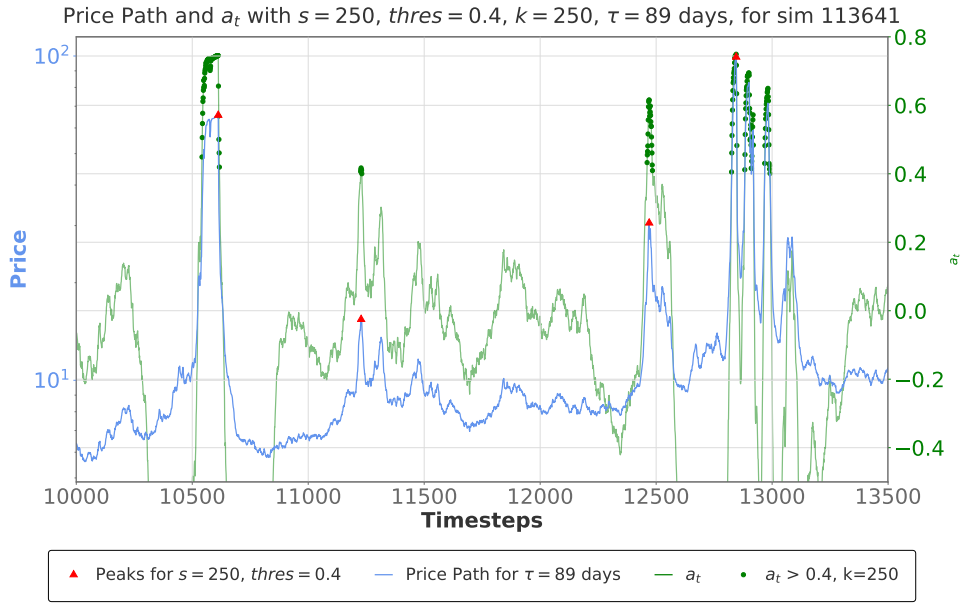


Figure 2.4: Zoom of Figure 2.3. The value of a_t is computed as defined in Equation 2.2. The values of a_t above the threshold $thres = 0.4$ are highlighted. In comparison with Figure 2.2 where $s = 50$, the value of a_t with $s = 250$ depicts much more truthfully the sharp movements of the underlying price path. The simulation is run for the values listed in Table 1.1, with the random seed 113641 and the memory parameter set to 89 days.

process. We disregard the results for $thres = 0.6$ as not enough bubbles are formed for this regime in a price path of 20000 time steps. For $thres \in [0.2, 0.5)$, the shorter memories (greater than 0 day) yield the most bubbles. We conclude that for a simulation of length 20000, the number of bubbles computed with the metric a_t is robust to a varying $thres \in [0.2, 0.5)$.

Similarly, Figure 2.8 demonstrates the impact of the threshold parameter on the average drawdown per bubble. The parameters s and k are set to 250 days. For each value, the error bars are computed as the standard deviation of the distribution of the log drawdown. As for the previous section, the error bars are too pessimistic as the trend is smooth and the mean drawdowns for each *thres* do not overlap another. As expected, a smaller *thres* yields smaller drawdown. More importantly, the behavior for noise traders with memory 0 and 1 day presented in section 2.3 are robust from a threshold of 0.2 to 0.6.

An important difference between the impact of s and *thres* is that the *thres* parameter subsets the bubbles that have a peak with price at least *thres*% above the average price of its $2s$ neighbors. The choice of *thres* is not relevant for the robustness of the analysis, as long as enough bubbles form the subset (based on the results displayed in Figure 2.7 we would not recommend going above 0.5 with such a simulation length). The parameter *thres* does not subset the bubbles by the regime they are in or the time they take to reach their peak. The parameter s however is more problematic as a small s hides bubbles from the analysis, such as the ones depicted in Figure 2.1. Within our framework of one trade per time-step, $s = 250$ depicts the best the reality.

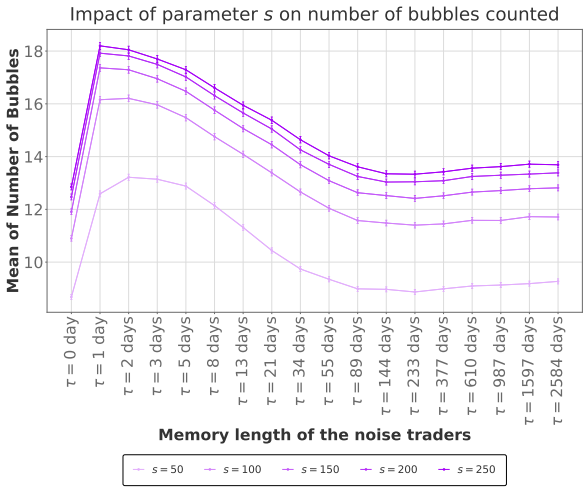


Figure 2.5: For a fixed *thres* $= 0.4$, the number of bubbles counted depends strongly and very significantly on the parameter s. Each point is the arithmetic mean of the number of bubbles over 1000 price paths. The error bars represent the standard error of the mean. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

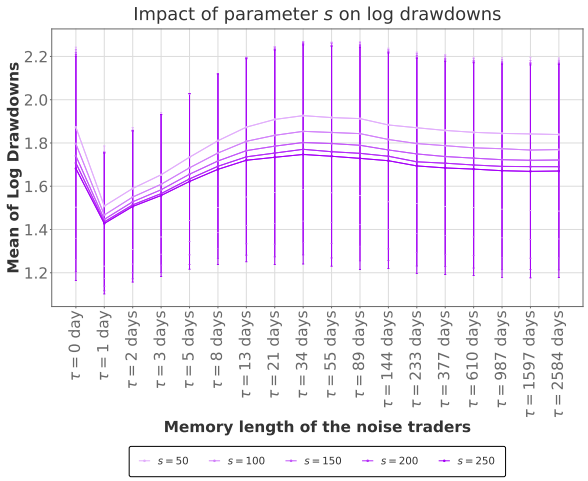


Figure 2.6: For a fixed *thres* $= 0.4$, the average drawdown for a given s and a given noise trader's memory is computed. Each point is the average drawdown per bubble. The error bars represent the standard deviation as error of the drawdown. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

2.5 Number of bubbles for longer simulations

We here present the number of bubbles obtained by running our analysis on 1000 seeds for each τ parameter, with each seed generating a price path of 495000 steps[‡] as (compared to 20000 for the preceding subsections). These results bring more significance as the price paths are 25 times longer, and it will allow us to compare on the same footing the results obtained in the next chapters. We here fix $s = 250$, *thres* $= 0.4$, $k = 250$.[§]

The number of bubbles created for each τ is plotted in Figure 2.9, with the numerical values given in Table 2.2. We note that the traders with shorter memory (bigger than 0 day) tend to create more bubbles with a maximum for 1 day with 473.3 ± 0.7 bubbles. The memories that create the least bubbles are 0 day (350.4 ± 0.6 bubbles), 144 days (361.1 ± 0.6 bubbles) and 233 days (361.8 ± 0.6 bubbles). The bubbles created by noise traders with a memory length of 1 day have a relatively lower

[‡]We generate 500000 steps from which we disregard the first 5000 (burn in period).

[§]The coupling strength of the noise traders depend on the memory length τ. Hence, the duration of the bubbles may also depend on τ. A smaller memory could tend to create bubbles faster as the noise traders agent only remember the immediate past. Hence, a smaller value of s might already be significant for shorter strictly positive memories. A variable s for each τ was not analyzed. The results shown in Figure 2.5 prove that memory lengths longer than 250 (the value of s) create more bubbles than memories of 89, 144 and 233 days.

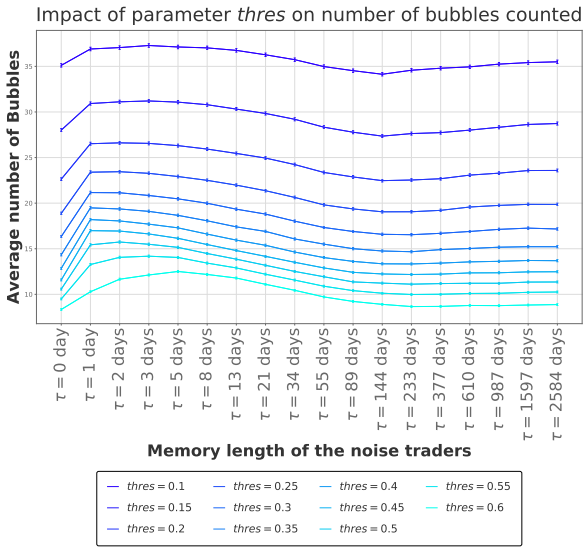


Figure 2.7: Evolution of the number of bubble w.r.t. the *thres* parameter for $s = 250$. For $thres \in [0.2, 0.5)$, the shorter memory (greater than 0 day) yield the most bubbles. The curve for $thres = 0.4$ is the same as the curve $s = 250$ in Figure 2.5. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

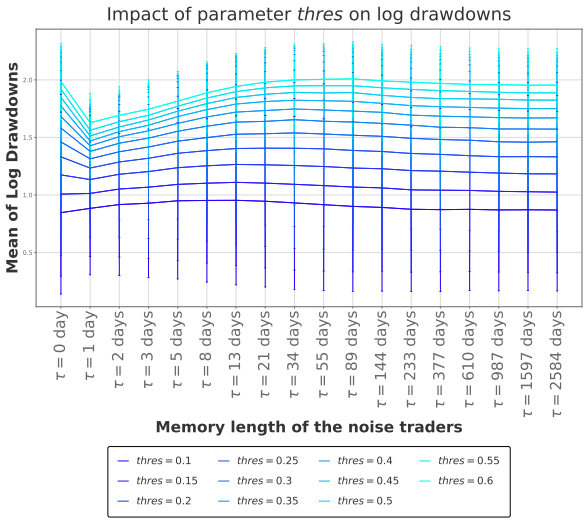


Figure 2.8: Evolution of the average drawdown w.r.t. the *thres* parameter for $s = 250$. A small *thres* is futile as it captures too many price movements and cannot pinpoint the bubbles. Starting from $thres = 0.2$ we see a non trivial trend forming, with the smallest drawdown attributed to a memory of 1 day. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

drawdown (1.4 ± 0.3) than the other memory lengths (Figure 2.10, leftmost column of Table 2.2). This is also the memory generating the most bubbles.

Looking at the number of years between bubbles (rightmost column of Table 2.2), we see that we have much less bubbles for all τ than what Westphal and Sornette (2019) reported for the same ABM but with a third agent acting as a Dragon Rider (DR) (2.5 years in average between bubbles in their simulations). Also, the average drawdowns in our simulation without the DR agent are up to 3 times higher than what Westphal and Sornette (2019). As mentioned above, the higher number of bubbles and the lower average drawdown in Westphal and Sornette (2019) may be linked to their metric choice (Equation 2.1). A comparison with the results obtained in the present work can therefore not be made.

The noise traders with no memory, $\tau = 0$, have shown to behave as if they were embodied by a very long memory. Analyzing Equation 1.18 for an infinite memory $\theta \to 1$ shows that the recent price return has no impact on the momentum and the momentum behaves as a constant (after a large enough number of time steps). It follows that the switching probabilities (Equation 1.22) do

not depend on the previous value of the momentum in the limit $\theta \to 1$. This is in accordance with the implementation of the noise traders with no memory (Table 2.1) where the momentum is set to $H_t = 0$. Therefore, it is expected for the noise traders with no memory to have the social coupling strength of noise traders with a very long memory, up to a constant equal to the expected value of the returns, the growth rate of the dividends r_d (Equation 1.2). Indeed, the expected value of the price returns is constant (Equation 1.14) and should, reasonably, be equal to the long term growth rate. See Westphal and Sornette (2019) for the verification that the long term growth rate of the price is equal to the growth rate of the dividends. Plugging Equation 1.18 in Equation 1.22 yields

$$\begin{aligned} p_t^{\pm} = &\quad \frac{p^{\pm}}{2}(1 \mp \kappa_t(s_t + \theta H_{t-1} + (1-\theta)R_t)) \ , \\ \overset{\theta \to 1}{=} &\quad \frac{p^{\pm}}{2}(1 \mp \kappa_t(s_t + H_{t-1})) \ , \\ \approx &\quad \frac{p^{\pm}}{2}(1 \mp \kappa_t(s_t + r_d)) \ . \end{aligned} \tag{2.4}$$

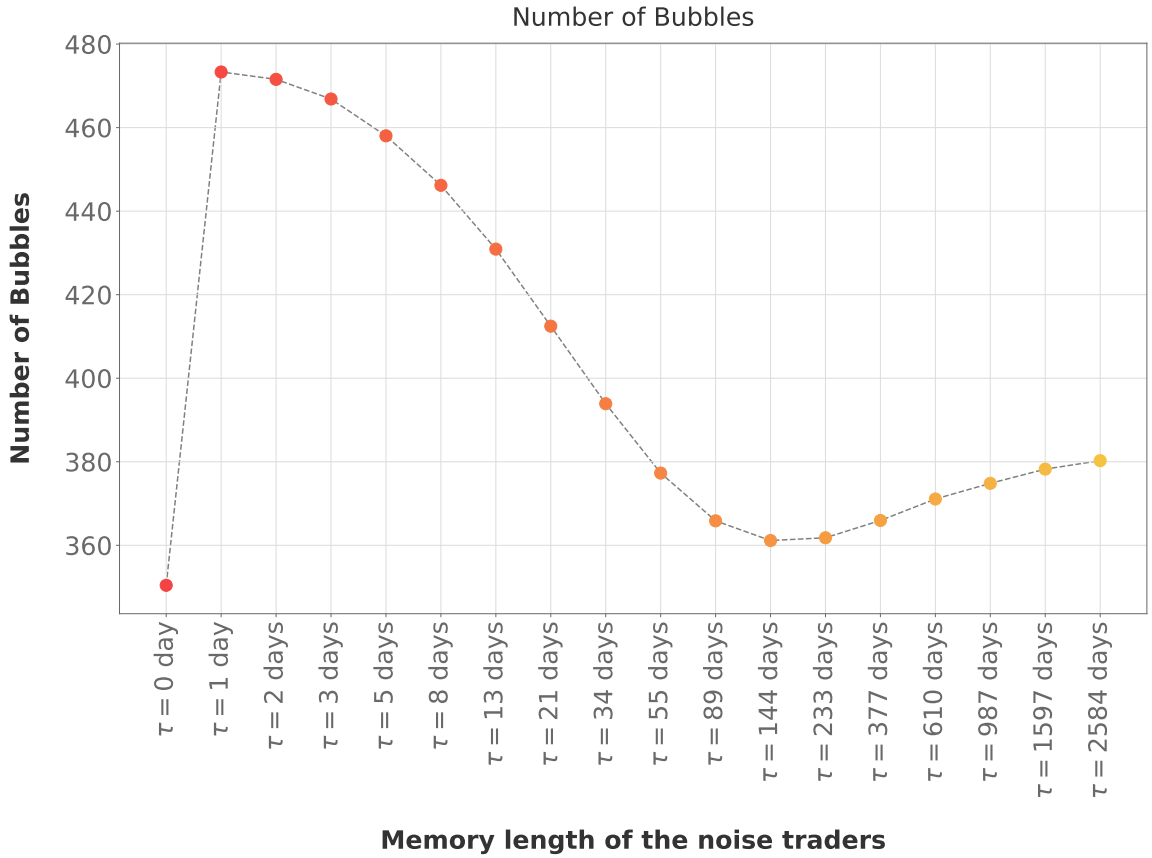


Figure 2.9: Average number of bubbles created for each memory parameter listed in Table 2.1 over 1000 simulations of 495000 time steps. The metric applied to detect the peaks is given in Equation 2.2. The errors (not visible at this scale) are computed as the standard error of the mean as we are in a counting process. Noise traders with a 1 day memory produce more bubbles than all other memories. As it can be seen, the shorter the memory (except for no memory at all) the more bubbles are created. Numerical values as well as the average number of years between bubbles are given in Table 2.2. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

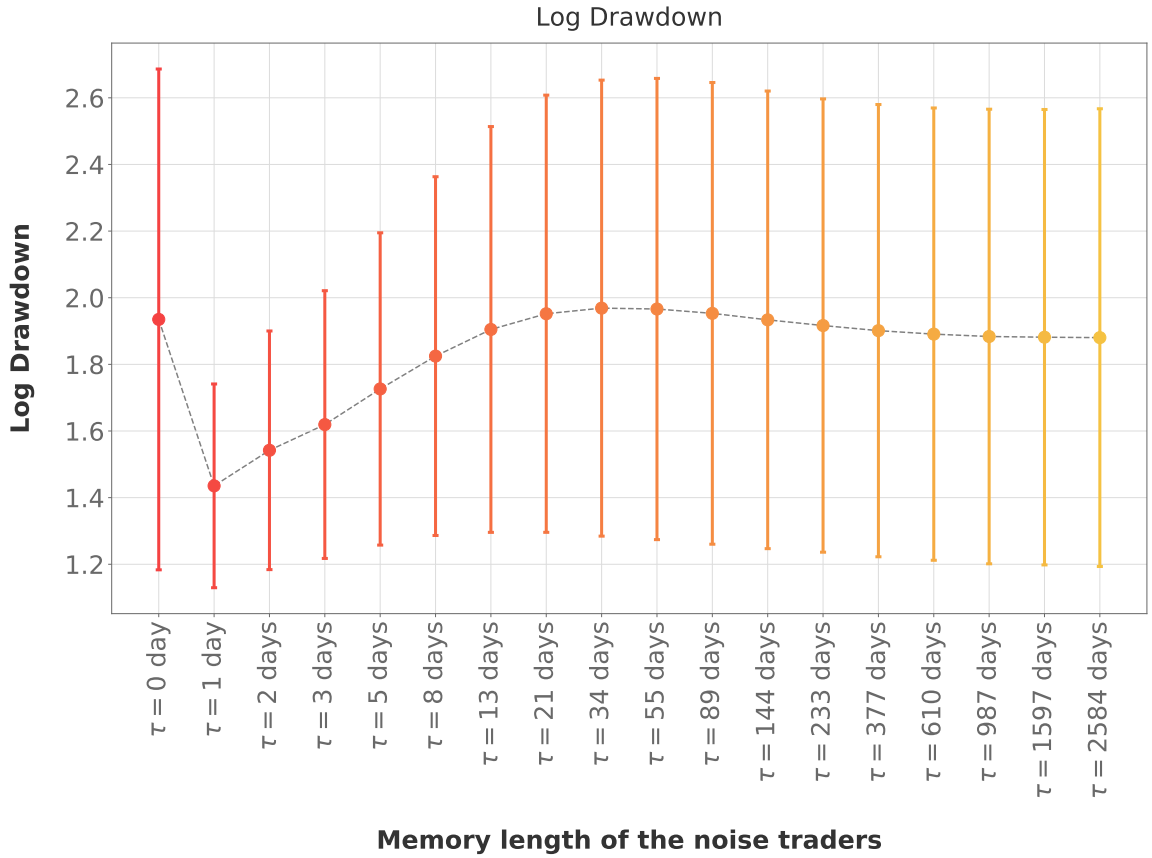


Figure 2.10: Average drawdown of bubbles created for each memory parameter listed in Table 2.1 over 1000 simulations of 495000 time steps. The metric applied to detect the peaks is given in Equation 2.2. The drawdowns are computed as in Equation 2.3. The noise traders with a 1 day memory create the bubbles with the smallest drawdown but also the most bubbles (Figure 2.9). The errors are computed as the standard deviation of the distribution of the log drawdowns. The values of the drawdowns are not statistically different. However, the computation of the drawdowns for different value of s (Figure 2.6) displayed a systematic clear and non-overlapping trend. The standard deviation may thus be too pessimistic. The values plotted are listed in Table 2.2. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

τ	**Drawdown** $s = 250$, $thres = 0.4$	**Number of bubbles** $s = 250$, $thres = 0.4$	**Number of years between bubbles** $s = 250$, $thres = 0.4$
0	1.9 ± 0.8	350.4 ± 0.6	5.650 ± 0.010
1	1.4 ± 0.3	473.3 ± 0.7	4.183 ± 0.006
2	1.5 ± 0.4	471.6 ± 0.7	4.199 ± 0.006
3	1.6 ± 0.4	466.8 ± 0.7	4.241 ± 0.006
5	1.7 ± 0.5	458.0 ± 0.7	4.323 ± 0.006
8	1.8 ± 0.5	446.2 ± 0.7	4.438 ± 0.007
13	1.9 ± 0.6	430.9 ± 0.7	4.595 ± 0.007
21	2.0 ± 0.7	412.5 ± 0.6	4.800 ± 0.007
34	2.0 ± 0.7	393.9 ± 0.6	5.027 ± 0.008
55	2.0 ± 0.7	377.3 ± 0.6	5.248 ± 0.009
89	2.0 ± 0.7	365.9 ± 0.6	5.412 ± 0.009
144	1.9 ± 0.7	361.1 ± 0.6	5.483 ± 0.009
233	1.9 ± 0.7	361.8 ± 0.6	5.473 ± 0.009
377	1.9 ± 0.7	365.9 ± 0.6	5.411 ± 0.009
610	1.9 ± 0.7	371.1 ± 0.6	5.336 ± 0.009
987	1.9 ± 0.7	374.8 ± 0.6	5.282 ± 0.009
1597	1.9 ± 0.7	378.2 ± 0.6	5.235 ± 0.009
2584	1.9 ± 0.7	380.3 ± 0.6	5.207 ± 0.008

Table 2.2: Numerical values for the results shown in Figure 2.9 and Figure 2.10. The simulations for each τ are run over 495000 time steps for each of the 1000 seeds . This represents 1980 trading years for each seed of each τ. The error on the drawdown is taken as the standard deviation, whereas the error for the number of bubbles is the standard error of the mean. The noise traders with a memory of 1 day create the most bubbles but with the lowest average drawdown. The values of the drawdown, plotted in Figure 2.10, are not statistically different. However, the computation of the drawdowns for different value of s (Figure 2.6) displayed a systematic clear and non-overlapping trend. The standard deviation may thus be too pessimistic. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

Chapter 3

Time Series Analysis

As mentioned, volatility clustering is the tendency of large absolute returns to be followed by large absolute returns, and small absolute returns followed by small absolute returns.

It is known that multiple time scales are needed for the emergence of volatility clustering (LeBaron (2006)). The present ABM fulfills this criterion. The fundamentalist bases his investment decision on his long term view of the market and is not influenced by short term market fluctuations. The noise traders has a memory of length τ, which is varied in the present work between 0 day and 2584 days. Furthermore, the social imitation strength governed by the Ornstein-Uhlenbeck process has an expected reverting time from the critical to the subcritical regime of 21 days.

A classical approach to quantify volatility clustering is through time series analysis. Cont (2007) demonstrates that volatility clustering translates to signed returns r_t showing no autocorrelation at all lags while the powers of the returns $|r_t|^d$ show a slowly decaying autocorrelation function. Ding et al. (1993) argue this behavior can characterize a "long memory" and is the strongest for d close to 1.

The present chapter is structured as follows: section 3.1 analyzes the autocorrelations and the partial autocorrelations of the signed returns generated by the ABM presented in chapter 1. The section 3.2 studies the autocorrelation of the absolute returns. As LeBaron (2000) pinpoints, the challenge of complex agent-based models is to clarify which aspect of the model is responsible for the stylized facts observed. In this regard, the simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length.

The analysis presented in the following sections rely on key concepts of time series analysis. Appendix A presents an introduction to the key definitions and equations used in this chapter. The shortcomings and pitfalls of the considered approximation methods are highlighted.

3.1 Autocorrelation of the signed returns

Real markets exhibit significant autocorrelations for very short time scales only (Malkiel and Fama (1970), Fama (1991), Pagan (1996), Cont et al. (1997)). The correlation period represents the delay needed for the market to react to the information (Cont (2001)), or that arbitraging the remaining trend is not cost efficient.

In the present artificial market no fees are taken into considerations, but the data generated and analyzed is on the finest information time scale as no exchange of information or opinion is possible between two consecutive returns. A delay in the arbitrage is thus compatible to some extent. However, as will be shown, significant coefficients are found up to lag 100. The first negative lags can partially be explained by noise traders switching too promptly from a bullish to a bearish behavior (and vice-versa). Nonetheless, it has to be raised that the current artificial market does not show typical market behavior for the autocorrelation of signed returns at small lags and could benefit from an additional agent arbitraging the autocorrelations. Such an agent could use a Kalman-Levy filter to fit an ARMA process to account for the remaining nonconstant conditional mean and constant conditional variance* as a trend following strategy. In the current setting, returns exhibit strong non stationarity which might imply an underlying bias in the results presented in the current chapter.

The coefficients of the autocorrelation function (ACF)† are presented in Figure 3.1. The autocorrelation coefficient for a specific lag and τ is the average coefficient for this lag taken over 1000 simulations of 494999 returns generated for this memory parameter‡. The red horizontal dashed lines represent the 95% confidence level to accept null hypothesis that the coefficients are i.i.d., i.e. coefficients within the dashed lines can be considered to be 0. The interval is given by $\pm 1.96/\sqrt{494999} = \pm 0.0027$ (see section A.2). The errors bars (hardly visible at this scale) represent the standard deviation of the coefficients. The coefficients of the partial autocorrelation function (PACF) are displayed in Figure 3.2. Their values and standard deviation are obtained similarly to the ACF coefficients. The ACF shows an exponential decay while the PACF has a clear cut-off between lag 2 and 7 depending on the memory of the noise traders. The cut-off values, lowest coefficient and the lag of the lowest coefficient for the ACF and PACF are given in Table C.3. As expected§, longer memories are linked to an autoregressive process of higher lags. Based on the insights from section A.7 these out of sample observations support an AR(p) model, though an ARMA(p, q) could be more parsimonious.

*To account for a remaining nonconstant conditional variance (known as conditional heteroskedasticity), a GARCH model would be preferred.

†The ACF coefficients are computed with a Fast Fourier Transform convolution (order $n \log n$), instead of the brute force method (of order n^2), speeding the process by a factor 37750.

‡As for section 2.5 and chapter 4, each simulation generates a price path of 500000 time steps of which we remove a burn in period of 5000 returns. This yields 494999 returns.

§See section A.2 and section A.6 for the intuition behind it. Also, this can be traced back to the price momentum part of the social interaction of the noise traders (Equation 1.18). A higher τ leads to older price returns taken into consideration.

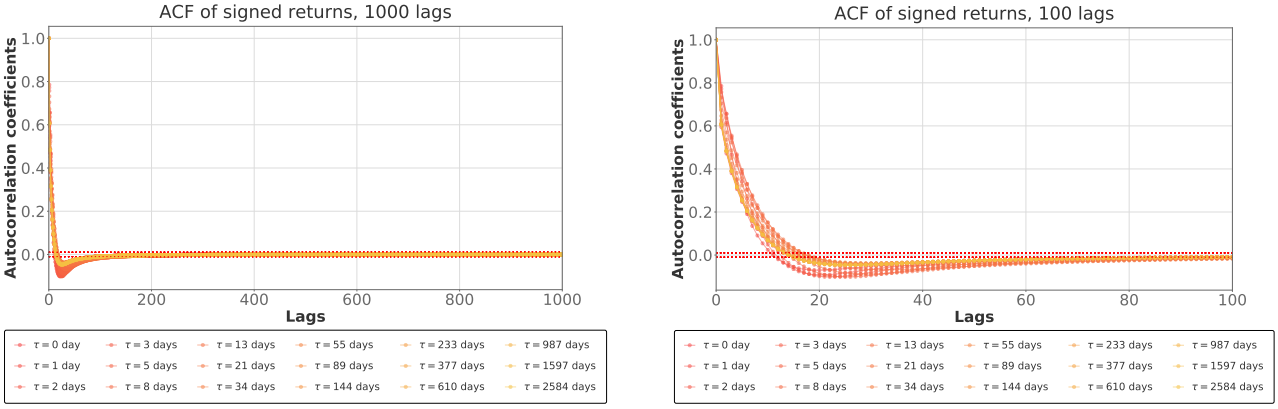


Figure 3.1: ACF coefficients of signed returns up to lag 1000 (left) with a zoom on the first 100 lags (right), for each memory τ. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. The red horizontal dashed lines represent the 95% confidence level to accept null hypothesis that the coefficients are i.i.d. Significant coefficients are found up to lag 100, indicating a strong non stationarity in the returns generated. The cut-off lag (first lag to cross the i.i.d. null hypothesis), lowest coefficient and the lag of the lowest coefficient are given for each τ in Table C.3.

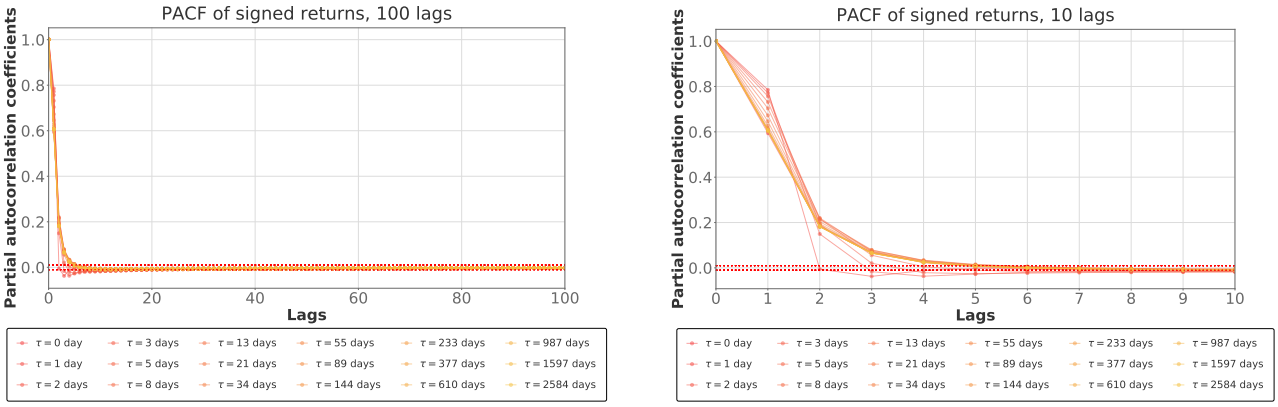


Figure 3.2: PACF coefficients of signed returns up to lag 100 (left) with a zoom on the first 10 lags (right). The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. The red horizontal dashed lines represent the 95% confidence level to accept null hypothesis that the coefficients are i.i.d. The cut-off lag (first lag to cross the i.i.d. null hypothesis), lowest coefficient and the lag of the lowest coefficient are given for each τ in Table C.3

.

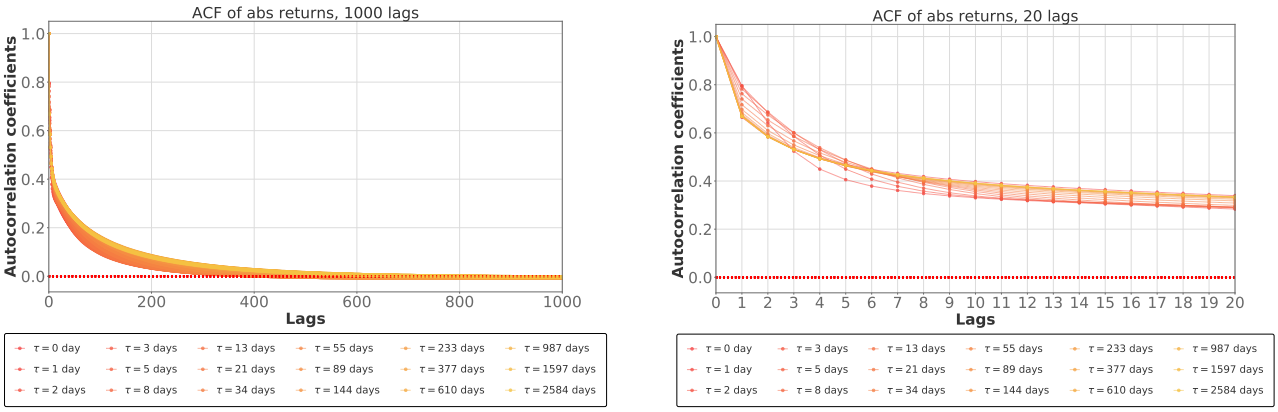


Figure 3.3: ACF coefficients of the absolute returns up to lag 1000 (left) with a zoom on the first 20 lags (right). The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. The coefficients and the standard errors are computed similarly to the signed returns (section 3.1). The red horizontal dashed lines represent the 95% confidence level to accept null hypothesis that the coefficients are i.i.d. A slower decay indicates stronger volatility clustering.

3.2 Autocorrelation of the absolute returns

The slow decay of the autocorrelation coefficients of the absolute returns is an evidence of volatility clustering (Cont (2007)). The slower the decay, the more volatility clustering is present (Ding et al. (1993)). Figure 3.3 displays the decay for the different memory parameters. The analysis of the impact of the memory parameter on this decay can offer an alternative metric of volatility clustering to the one presented in chapter 4.

It is known from Equation A.34 that the sum of the coefficients of a GARCH(p, q) fit on the signed returns quantifies the decay of the autocorrelation of absolute returns. A GARCH(1,1) model is fitted on every simulation. The sum of the two GARCH coefficients is averaged over the 1000 simulations. The error is taken as the standard deviation of the distribution of the sums. The results displayed in Figure 3.4 do not allow any clear interpretation. Note that the mean of the sum is below but close to the critical value 1.0 (Equation A.32) for all memory lengths[¶], indicating a variance marginal distribution close to infinite. Mikosch et al. (2000) argue at length that in such a case the sample autocorrelations are deceptive estimators for the signed and absolute returns. The ACF of the GARCH(1,1) is an unreliable estimator (tardy convergence), becomes senseless, and one should deflect drawing conclusions from the ACF of the absolute returns (see also Resnick (1998) and Cont (2007)). The non-stationarity of the returns (Figure 3.1) may also play a part (see Equation A.25 and more generally section A.8 where ε_t is always defined as WN(0,1)).

Another approach to quantify the power decay is its decay exponent β. The objective is to find a subset of lags for which the power decay exponent is constant and representative. Let (x_τ, y_τ) be a

[¶]The confidence bands of some memories overlap with the critical value.

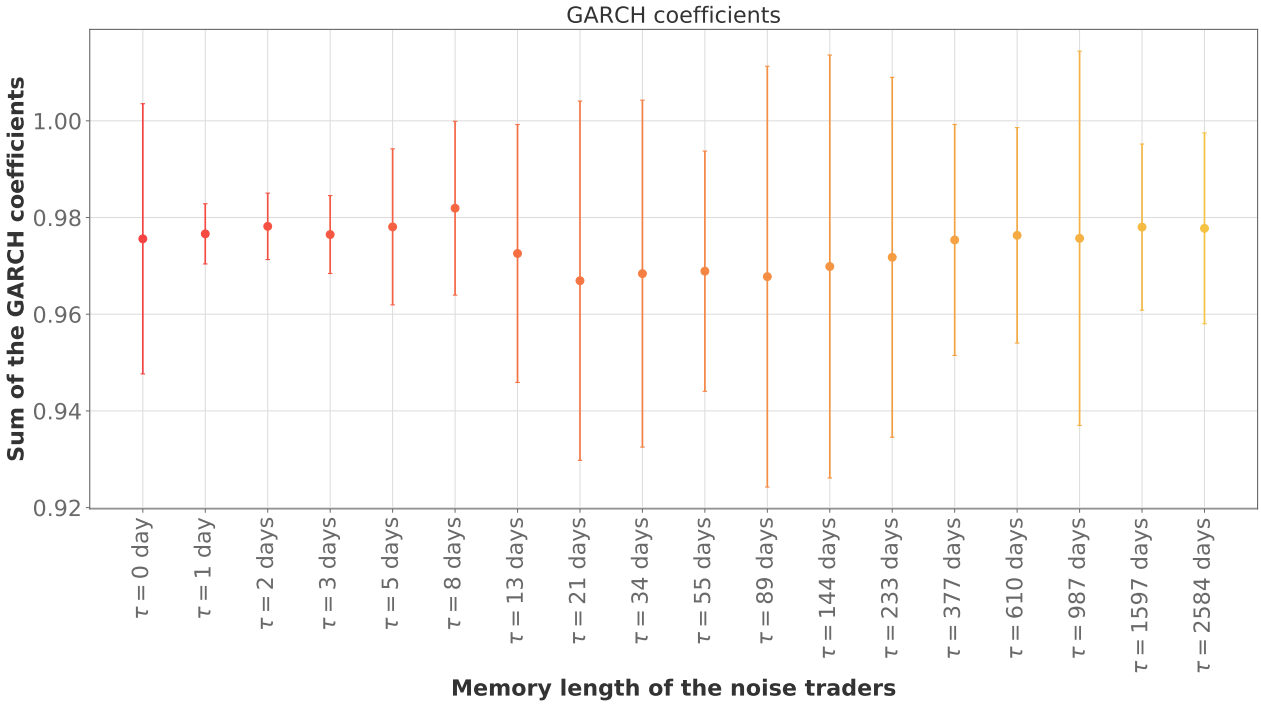


Figure 3.4: Average sum of the two GARCH(1,1) coefficients for each memory parameter. The error bars are the standard deviations of the distributions of the sums. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. All values are close to the critical point 1.0, with some error bars including it. No clear trend can be deduced and these results are not investigated further.

point with y_τ representing the average coefficient for the lag x_τ, computed over 1000 simulations. A null hypothesis for each τ is generated by the least squares linear fit of the points $(x_\tau, \log(y_\tau))$. Each point (x_τ, y_τ) is assigned a p-value w.r.t. the corresponding null hypothesis $(x_\tau, y_{\tau,pred})$. To test the quality of the power law exponent β (slope of the linear fit), Fisher's combined probability test is used (Kost and McDermott (2002), DeGroot and Schervish (2012)). The test is based on the fact that a sum of k variables with a χ^2 distribution is χ^2_ν with $\nu = 2k$ degrees of freedom. First, note that the exponential distribution has the cumulative distribution function (cdf)

$$F(x, \lambda) = (1 - e^{-\lambda x}) \cdot H(x) \,, \tag{3.1}$$

where $H(x)$ is the Heaviside step function. The quantile function is then

$$x = F^{-1}(p, \lambda)^{-1} = \frac{-\log(1-p)}{\lambda} = \frac{-\log p}{\lambda} \,, \quad 0 \leq p < 1 \,. \tag{3.2}$$

The cdf of the χ^2_ν distribution for $\nu = 2$ is the cdf of the exponential distribution with rate $\lambda = \frac{1}{2}$

$$F(x, \nu = 2) = \frac{\int_0^{\frac{x}{2}} t^{\frac{2}{2}-1} e^{-t} dt}{(\frac{2}{2}-1)!} = 1 - e^{-\frac{x}{2}}. \tag{3.3}$$

To combine the product of p-values, the sum is taken and multiplied by 2 to get the $\frac{1}{2}$ rate parameter. The test statistics is then

$$\chi^2_{2k} \sim -2 \sum_{i=1}^{k} \log p_i \ , \tag{3.4}$$

where k indicates the number of lags fitted and p_i the p-value for the null hypothesis at lag i. From the null hypothesis, a large combined p-value is interpreted as a confirmation of a significant fit.

To pinpoint a constant and representative exponent β, the combined p-value for different ranges are probed. The ranges have a starting lag s_τ and end lag f_τ. The ranges are then shrunken by keeping f_τ fixed and increasing s_τ up to $f_\tau - m$, where m is the minimal amount of lag to be fitted. In Figure 3.5 (zoom of Figure 3.3 on a semilogy scale), the lags up to lag 20 display a faster decay. They are disregarded and the initial lag set at $s_{\mathrm{H}i} = 20$ (black dashed vertical line in Figure 5.7) To determine the end lag f_τ is more difficult as some memories τ seem to cross the i.i.d. null hypothesis (red horizontal dashed line at 0.0027 as computed in section 3.1) at much lower lags than others. Evidently, f_τ depends on τ. In a first approach, f_τ is naively defined as the lag preceding the first lag for which the ACF coefficient is below the i.i.d. null hypothesis for a given τ.

Imposing a minimum of $m = 200$ lags to be fitted[||], no ranges, for any τ, yields a combined p-value above 0[**]. This confirms the visual intuition that the i.i.d. null hypothesis is not the cut-off corresponding to the end of the constant decay regime. The cut-offs 0.01, 0.02, 0.03, 0.04 for $m = 200$, $m = 100$ did not yield at least one range for every τ with a positive combined p-value. Shifting f_τ for all the mentioned cut-offs to the 50th and 100th lag preceding the first lag for which the ACF coefficient is below the cut-off turned out to be unsuccessful as well (for $m = 200$ and 100). The cut-off set at 0.01 (black dashed-line in Figure 3.5) and $m = 50$ were the least restrictive parameters for which all 18 memories had at least one range yielding a non zero combined p-value. The lower value of m may raise doubts on the intrinsic value of the results. It is stressed that m only influences the size of the final ranges tested. To support this, the evolution of β will be analyzed.

Equipped with at least one range for each memory parameter displaying a significant fit of the power decay, the robustness of these ranges are analyzed. The evolution of the combined p-value displays an unexpected percolating behavior for all memories (Figure 3.6, left) with an irreversible and sharp transition from a combined p-value of 0 to a combined p-value of 1. The percolation thresholds p^c_τ (vertical dashed lines) are the mean lag between the highest lag with a combined p-value below 0.01 and the lowest lag with a combined p-value above 0.99. Note that for all values

[||] This means that the last range to be fitted is $[f_\tau - 200, f_\tau]$.

[**] Concretely, the decay of the memory parameter $\tau = 0$ (1 day) is above the i.i.d. null hypothesis up to lag 490. For this parameter, a linear fit is computed on the ranges $[i, 490]$ for $i \in [20, 490 - 200]$, and the combined p-value is computed for each range.

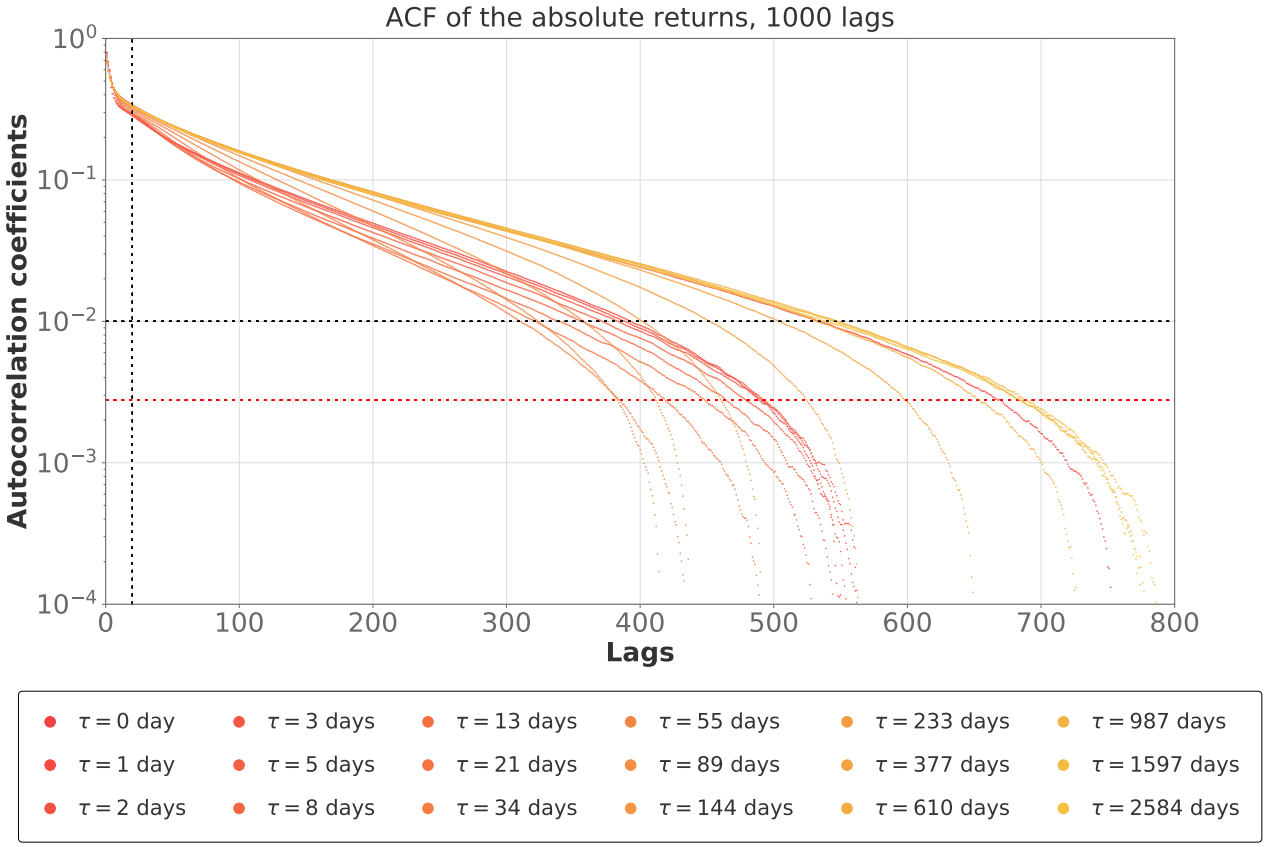


Figure 3.5: ACF coefficients of the absolute returns for all memory parameters up to lag 800 (zoom of Figure 3.3 on a semilogy scale). The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. The red horizontal dashed line represents the upper band of the 95% confidence level to accept the null hypothesis that the coefficients are i.i.d. (see section A.2). The black vertical dashed line is at lag 20, the start lag s_τ used to pinpoint a range of lags yielding a positive combined p-value. The black horizontal dashed line corresponds to ACF coefficients with value 0.01. This cut-off is the fix end lag f_τ of the ranges used to pinpoint a positive combined p-value. A slower decay of the ACF of the absolute returns implies stronger volatility clustering.

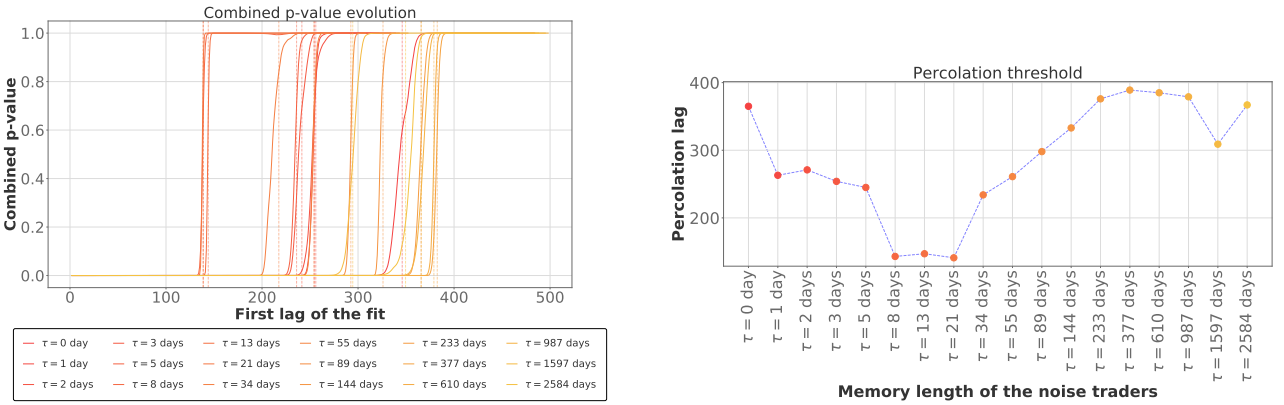


Figure 3.6: The left figure shows the evolution of the combined p-value of the linear fit on the range $[\log(y_\tau(x_{s_\tau})), \log(y_\tau(x_{f_\tau}))]$ of the points $(x_\tau, \log(y_\tau))$ displayed in Figure 3.5, for the starting lags $s_\tau \in [20, f_\tau - m]$ given on the x-axis of the present plot, the final lag f_τ corresponding to the lag preceding the first lag for which the ACF coefficient is below 0.01 (black dashed line in Figure 3.5) for a given τ, and minimal amount of lags fitted $m = 50$. The evolution of the combined p-value for each τ displays an unexpected percolating behavior with an irreversible and sharp transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds p^c_τ (vertical dashed lines). The value p^c_τ is the mean lag between the highest lag s_τ for which the combined p-value is below 0.01 and the lowest lag s_τ with combined p-value of at least 0.99. The plot on the right displays the percolation thresholds. The values of p^c_τ are listed in Table 3.1. The data on which the ACF of absolute returns are computed in Figure 3.5 are simulations run varying only the memory parameter τ of the noise traders for the values listed in Table 1.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length.

of f_τ and m analyzed, the memories for which at least one range yielded a non-zero combined p-value systematically displayed the same behavior[††]. The percolation thresholds are the tipping points at which the decays turn to a constant power decay. The percolation thresholds appear to be an intrinsic property of the decays analyzed.

The values of p^c_τ and f_τ are listed in Table 3.1. It shows that for $\tau \in \{33, 55, 89, 144, 233\}$, the maximal range with a statistically significant fit consists of less than 100 lags. This sustains why the 18 memories all showed a positive combined p-value only for $m = 50$. The small number of lags for these memories motivates us to disregard the power decay exponent associated with them. For the exponent to be representative, it must embody a legitimate part of the decay.

The evolution of the decay exponents β for all ranges analyzed is displayed in Figure 3.7 (left). The numerical values of β are given in Table 3.1. The decay exponents show a downward trend for

[††]The cut-off at 0.01 and $m = 50$ are the least restrictive parameters yielding at least one range with a combined p-value above 0 for *all* memory parameters, but some of the other choices cut-offs (0.02, 0.03, 0.04 and the shifts of 50 and 100 lags for each of them) and m (200, 100, 50) had some memories, not all, that had at least a range with a statistically significant fit. In each case, a similar percolating behavior was displayed.

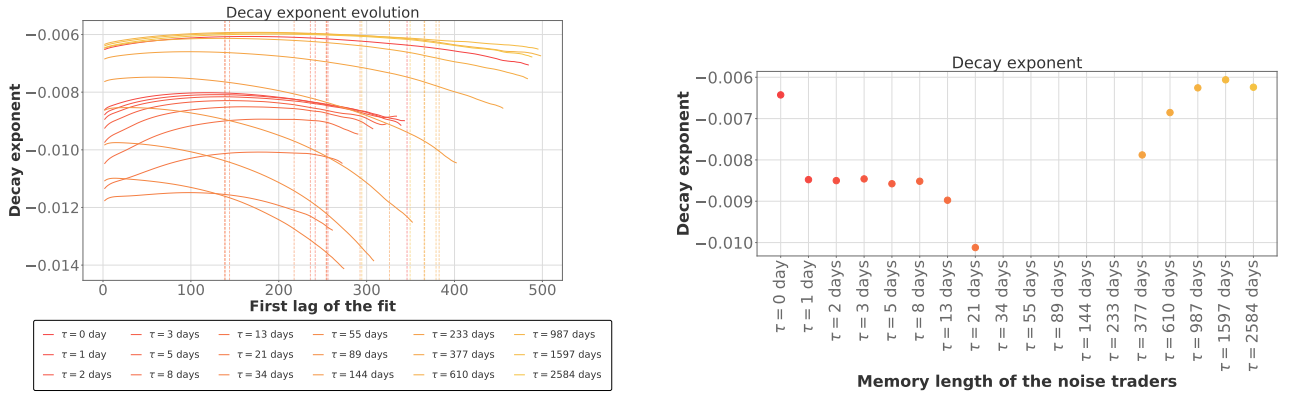


Figure 3.7: The left figure shows the values of the decay exponent β of the linear fit on the range $[\log(y_\tau(x_{s_\tau})), \log(y_\tau(x_{f_\tau}))]$ of the points $(x_\tau, \log(y_\tau))$ displayed in Figure 3.5, for the starting lags $s_\tau \in [20, f_\tau - m]$ given on the x-axis of the plot, the final lag f_τ corresponding to the lag preceding the first lag for which the ACF coefficient is below 0.01 (black dashed line in Figure 3.5) for a given τ, and $m = 50$. The evolution of the combined p-value for each τ displays an unexpected percolating behavior with an irreversible and sharp transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds p^c_τ (vertical dashed lines). The value p^c_τ is the mean lag between the highest lag s_τ for which the combined p-value is below 0.01 and the lowest lag s_τ with combined p-value of at least 0.99. The plot on the right displays the decay exponent computed on the range $[p^c_\tau + 1, f_\tau]$. The values of β yielding a non representative behavior of the overall decay due to an insufficient number of lags fitted are not displayed ($\tau \in \{34, 55, 89, 144, 233\}$). The values of β are listed in Table 3.1. The data on which the ACF of absolute returns are computed in Figure 3.5 are simulations run varying only the memory parameter τ of the noise traders for the values listed in Table 1.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length.

higher starting lags s_τ which is explained by increasingly shorter ranges of points fitted and an acceleration of the decay for lags approaching the i.i.d. null hypothesis, as seen in Figure 3.5. The memories $\tau \in \{34, 55, 89, 144, 233\}$ days display the strongest acceleration. This confirms that, even though these memories have a range of the decay with a positive combined p-value, these ranges are too short (Table 3.1) to be significant and are non representative of the overall decay of the ACF of absolute returns by being in an accelerating phase of the decay[‡‡]. For all other memories the decay exponent is stable. To quantify volatility clustering, β is computed for the remaining memories on the range $[p^c_\tau + 1, f_\tau]$, the largest range displaying a positive combined p-value, and is shown in Figure 3.7 (right). Cont (2001) cites $[0.2, 0.4]$ for the exponent of the decay of the ACF on absolute return in real markets. The results obtained show much slower decay, hence more volatility clustering.

Based on the strong non stationarity of the autocorrelation coefficicients of the signed returns, the results obtained through the GARCH approach and their implications (Resnick (1998), Mikosch

[‡‡]Note that, as mentioned, higher cut-offs were also tested and did not show ranges of a significant fit for these memories.

et al. (2000), Cont (2007)), the difficulty to find a clear power decay of the coefficients of the absolute returns, the results presented in Figure 3.7 (right) as representative of the volatility clustering must be taken with caution. The evidences suggests that most memories have an oscillating behavior around a constant power decay up to their percolation threshold (Figure 3.6, with the decay exponent staying constant in Figure 3.7 for $\tau \notin \{34, 55, 89, 144, 233\}$). It might be argued that the percolation threshold, in this artificial market, embodies more significantly the volatility clustering (compare Figure 3.6 (right) with Figure 4.8) than the decay exponent β. Under these considerations, short strictly positive memories display a faster decay and thus less volatility clustering than longer memories. The memory producing the least amount of volatility clustering is found at $\tau = 21$. For $\tau \geqslant 377$, the longer the memories the more volatility clustering is present. The noise traders with $\tau = 0$ behave as expected as if they had an infinite memory (Equation 2.4).

τ	**Percolation lag** p_τ^c	**End lag** f_τ	**Number of lags fitted**	**Decay exponent** β
0	346	534	188	-0.006
1	254	393	139	-0.008
2	256	389	133	-0.008
3	241	384	143	-0.008
5	236	372	136	-0.009
8	139	357	218	-0.009
13	144	340	196	-0.009
21	138	322	184	-0.010
34	217	311	94	-0.012*
55	254	324	70	-0.014*
89	292	358	66	-0.013*
144	326	402	76	-0.012*
233	365	452	87	-0.010*
377	382	505	123	-0.008
610	379	533	154	-0.007
987	366	548	182	-0.006
1597	294	545	251	-0.006
2584	349	538	189	-0.006

Table 3.1: The range $[\log(y_\tau(x_{s_\tau})), \log(y_\tau(x_{f_\tau}))]$ of the points $(x_\tau, \log(y_\tau))$ displayed in Figure 3.5 are fitted for varying starting lag s_τ. The evolution of the combined p-value for each τ displays an unexpected percolating behavior with an irreversible and sharp transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds p_τ^c (vertical dashed lines). The percolation thresholds p_τ^c are the mean lag between the highest starting lag for which the range yields a combined p-value below 0.01 and the lowest starting lag for which the range yields a combined p-value above 0.99. See Figure 3.6. The end lag f_τ represents the lag preceding the first lag for which the ACF coefficients of the absolute returns are below the cut-off 0.01. The number of lag fitted is the difference between the end lag and the percolation lag. The decay exponent β is computed on the ranges $[p_\tau^c + 1, f_\tau]$. The asterisk * shows the values of β yielding a non representative behavior of the overall decay due to an insufficient number of lags fitted. The data on which the ACF of absolute returns are computed in Figure 3.5 are simulations run varying only the memory parameter τ of the noise traders for the values listed in Table 1.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length.

Chapter 4

Impact of Noise Traders' Memory on Volatility Clustering

In chapter 3, the impact of each parameter τ listed in Table 2.1 on volatility clustering was quantified through the decay of the autocorrelation coefficients of the absolute returns. This approach proved difficult, and did not yield a significant result for 5 of the 18 memories. The present chapter offers a more innovative approach to quantify volatility clustering based on the moments of the frequency distribution of the large absolute returns.

The first section studies the probability of each memory to create large absolute returns. The following section presents an novel approach to quantify volatility clustering. Finally, the endogeneity of the market is analyzed through a non parametric approximation of the branching ratio for self-exciting Hawkes point processes in one dimension.

4.1 Distribution of returns

In this section, the unconditional probability for a market with the noise traders memory length τ to produce large absolute returns is studied. Then, the conditional probability of generating a large absolute returns r_t above v given that the previous return r_{t-1} is a large absolute return above v is presented. The thresholds $v = 1\sigma$ and $v = 2\sigma$ are used. An example of the subset of returns above each threshold is depicted in Figure 4.1.

4.1.1 Unconditional probability

In a Gaussian distribution, the unconditional probability that the absolute value of the next step is outside 1σ is $1 - P(\mu - 1\sigma \leq X \leq \mu + 1\sigma) \approx 1 - 0.6827 \approx 0.3137$. For 2σ, we have $1 - P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) \approx 1 - 0.9545 \approx 0.0455$. To familiarize ourselves with the distribution of the returns in the current artificial market, the ratio of returns larger in absolute value than a threshold v for all memory lengths depicted in Table 2.1 is computed. This ratio is the unconditional probability to find a return with absolute value above v. It is referred to it as p_u^v.

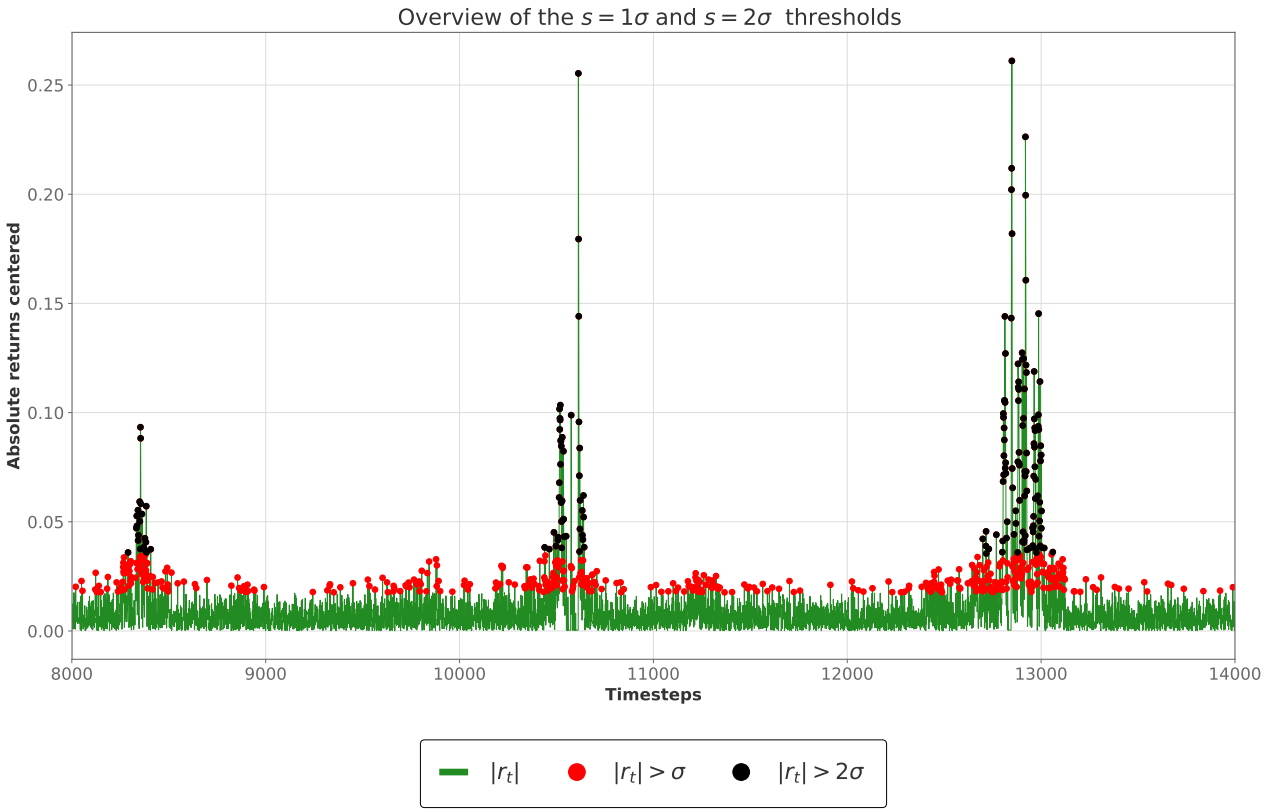


Figure 4.1: Overview of the 1σ and 2σ thresholds for a simulation with the noise traders memory fixed at 89 days. The 1σ threshold is represented by the black and red dots. This threshold includes many returns that are not part of the largest clusters. The 2σ threshold is depicted by the black dots only.

For each τ, 1000 price paths of 494999 returns are generated*. The distributions of $p_u^{1\sigma}$ and $p_u^{2\sigma}$ for each τ are presented in Figure 4.2 where the dashed lines highlight the mean values of p_u^v. The numerical values of the means are given in Table 4.1. The memory parameter has an impact on $p_u^{1\sigma}$ and $p_u^{2\sigma}$, with longer memories having the most central values. Note that $p_u^{1\sigma} > p_u^{2\sigma}\ \forall\tau$.

As comparison, the NASDAQ composite Index has on average 28.4% of its return above 1σ and 10.1% above 2σ for the period 1973-01-02 to 2018-12-31. For both thresholds the results obtained are smaller than the index, indicating less volatility. Also, this indicates that the distributions of the returns for the NASDAQ has longer tails than the Gaussian distributions but shorter tails than the distribution of the returns generated by the ABM.

The threshold $v = 1\sigma$ yields values much smaller than for a Gaussian distribution while values for $v = 2\sigma$ are already much closer to the value found for a Gaussian distribution. This indicates the distribution of the returns has longer tails than the Gaussian distribution.

*As for section 2.5, each simulation generates a price path of 500000 time steps of which we remove a burn in period of 5000 returns. This yields 494999 returns.

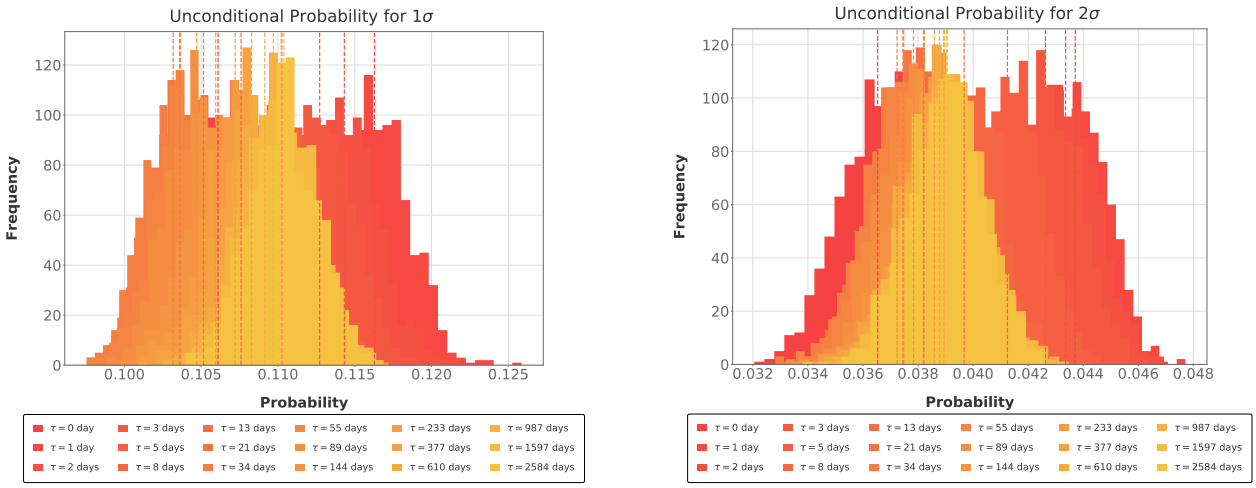


Figure 4.2: Distribution of the absolute returns above $v = 1\sigma$, $p_u^{1\sigma}$ (left) and above $v = 2\sigma$, $p_u^{2\sigma}$ (right). The computation is run over 494999 returns for each of the 1000 seeds for each τ and v. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. The dashed lines represent the means of the distributions. The numerical values of the means and their errors (the standard deviation of the corresponding distribution) can be found in Table 4.1 and are plotted in Figure B.2. The threshold $v = 1\sigma$ yields values much smaller than for a Gaussian distribution while $v = 2\sigma$ is already much closer. This may indicate the distribution of the returns has longer tails than the Gaussian distribution.

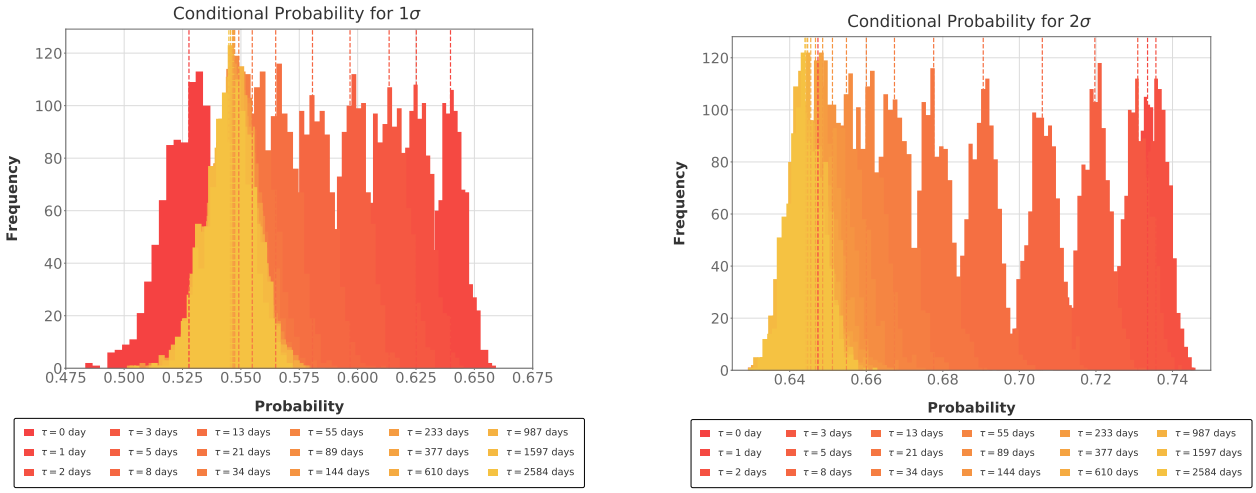
4.1.2 Conditional probability

We now compute the conditional probability p_c^v that, given the absolute return at time t is larger than v, the absolute return at time $t+1$ is also larger than v. Let A denote the set formed by all returns r_t for which $|r_t| > v$, B the set of all returns r_t for which $|r_t| > v$ and $|r_{t-1}| > v$, and C the number of uninterrupted sequences of returns $|r_t| > v$ [†]. For a simulation of length T, $\mathbb{E}[|A|] = T \cdot p_u^v$. The cardinality of C depicts the number of sequences of returns above v. It follows $|A| = |B| + |C|$ and

$$p_c^v = \Pr(|r_{t+1}| > v \mid |r_t| > v) = \frac{|B|}{|A|} = 1 - \frac{|C|}{|A|} \,. \tag{4.1}$$

The distributions of $p_c^{1\sigma}$ and $p_c^{2\sigma}$ for each τ are presented in Figure 4.3a and Figure 4.3b where the dashed lines highlight the mean value of each distribution. The numerical values of the means are listed in Table 4.1 and plotted in Figure B.3. As in the previous section, each mean is computed over 494999 returns for each of the 1000 seeds for each τ and v.

[†]As an example, consider the sequence 1111 00 11 0 1 00 111, where 1 represents a return with absolute value above v. Then $|A| = 10$, $|B| = 6$, $|C| = 4$.



(a) Distribution of the conditional probability $p_c^{1\sigma}$

(b) Distribution of the conditional probability $p_c^{2\sigma}$

Figure 4.3: Distribution of the absolute returns above v when the preceding absolute return is also above v, for $v = 1\sigma$ (left), $v = 2\sigma$ (right). The computation is run over 494999 returns for each of the 1000 seeds for each τ and v. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. The dashed lines represent the mean of each τ. The dashed line for $\tau = 0$, $v = 2\sigma$ is the red vertical line at 0.647, see Figure B.3 for a plot of the means and their errors (the standard deviation of the corresponding distribution), and Table 4.1 for the numerical values. Shorter memories tend to increase the probability to find two high absolute returns in a row.

Figure 4.3 shows that lower values of τ yield higher conditional probabilities, except for the case of no memory that behaves as for noise traders with an infinite memory (Equation 2.4). Comparing Figure 4.3 and Figure 4.2, or more straightforwardly Figure B.3 and Figure B.2, the means of $p_c^{2\sigma}$ are larger than $p_c^{1\sigma}$ for all memories, whereas $p_u^{2\sigma}$ are smaller than $p_u^{1\sigma}$. This behavior can be quantified by

$$q_v = \frac{p_c^v}{p_u^v}\,, \tag{4.2}$$

the ratio between the conditional and the unconditional probability for a threshold v. The larger q_v, the further away the conditional probability is from the unconditional probability. A higher q_v represents a stronger absolute probability of a second high absolute return in a row. In other words, for two memories with the same unconditional probability p_u^v, the memory with the highest q_v has the highest probability to create a second absolute return above v. In this sense, q_v depicts the strength of the market to stay in a volatility cluster of size two[‡]. The values for $q_{1\sigma}$ and $q_{2\sigma}$ are given

[‡]For the analysis of a more persistent behavior, the dependency of the endogeneity of the market on the memory parameter will be analyzed in section 4.3 through a branching ratio approximation of the self-exciting Hawkes process.

in Table 4.1 and plotted in Figure 4.4.

For $q_{1\sigma}$, the ratio follows again the trend depicted in Figure 2.9 with the highest values for the shortest strictly positive memories. The $q_{2\sigma}$ case is more interesting: the memory lengths between 1 and 13 days show an opposite behavior to Figure 2.9. For a given return absolute above v, the memories that have the highest probabilities to keep the market in a regime of large absolute returns of at least two days are 13 and 22 days.

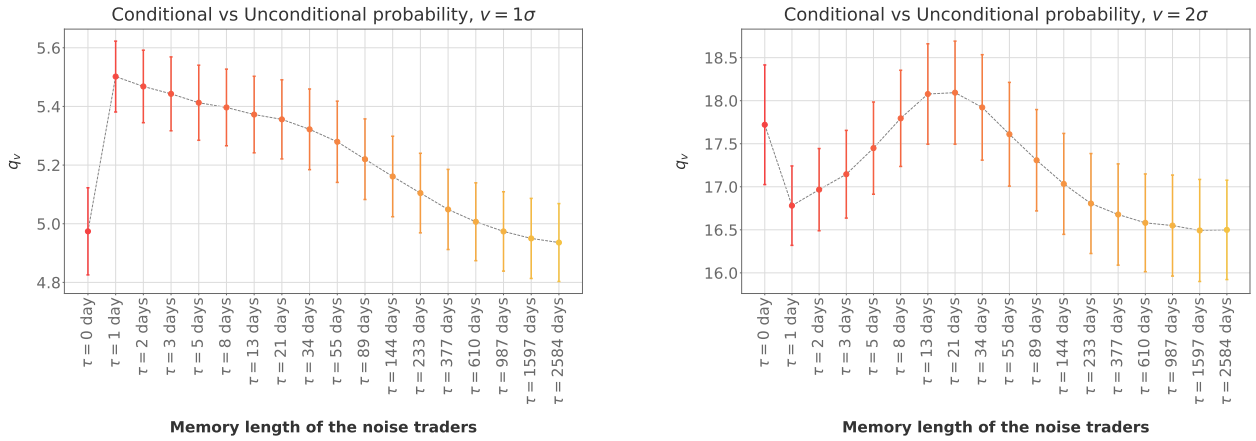


Figure 4.4: Ratio of the conditional probability p_c^v of the absolute return at time $t+1$ to be above v when the absolute return at time t is also above v, over the unconditional probability p_u^v to have an absolute return above v, for $v = 1\sigma$ (left) and $v = 2\sigma$ (right). The values plotted for $v = 1\sigma$ are listed in the third column of Table 4.1, and for $v = 2\sigma$ in the last column of Table 4.1. The error bars are the error propagation of the errors computed for p_c^v and p_u^v, both of which were the standard deviation of the corresponding distributions for each τ and v. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length.

The NASDAQ composite Index has an average of 42.4% of its return above 1σ when the preceding return is also above 1σ, and 29% for the 2σ threshold, for the period 1973-01-02 to 2018-12-31. We find $q_{1\sigma} = 1.49$ and $q_{1\sigma} = 2.87$. The values of q_v are much lower than the values found in our simulations. Combined with the values $p_u^{1\sigma}$, $p_u^{2\sigma}$ of the index mentioned in subsection 4.1.1 we conclude that the simulated market with only a noise trader and a fundamentalist agent creates less volatility than the NASDAQ index, while generating stronger clustering regimes of at least 2 days.

τ	$p_u^{1\sigma}$	$p_c^{1\sigma}$	$q_{1\sigma} = \frac{p_c^{1\sigma}}{p_u^{1\sigma}}$	$p_u^{2\sigma}$	$p_c^{2\sigma}$	$q_{2\sigma} = \frac{p_c^{2\sigma}}{p_u^{2\sigma}}$
0	0.106 ± 0.002	0.528 ± 0.012	5.0 ± 0.1	0.037 ± 0.001	0.647 ± 0.005	17.7 ± 0.7
1	0.116 ± 0.002	0.640 ± 0.006	5.5 ± 0.1	0.044 ± 0.001	0.733 ± 0.003	16.8 ± 0.5
2	0.114 ± 0.002	0.625 ± 0.007	5.5 ± 0.1	0.043 ± 0.001	0.736 ± 0.004	17.0 ± 0.5
3	0.113 ± 0.002	0.613 ± 0.008	5.4 ± 0.1	0.043 ± 0.001	0.731 ± 0.004	17.1 ± 0.5
5	0.110 ± 0.002	0.597 ± 0.008	5.4 ± 0.1	0.041 ± 0.001	0.720 ± 0.004	17.4 ± 0.5
8	0.108 ± 0.002	0.581 ± 0.008	5.4 ± 0.1	0.040 ± 0.001	0.706 ± 0.004	17.8 ± 0.6
13	0.105 ± 0.002	0.565 ± 0.009	5.4 ± 0.1	0.038 ± 0.001	0.691 ± 0.004	18.1 ± 0.6
21	0.104 ± 0.002	0.555 ± 0.009	5.4 ± 0.1	0.037 ± 0.001	0.678 ± 0.004	18.1 ± 0.6
34	0.103 ± 0.002	0.549 ± 0.010	5.3 ± 0.1	0.037 ± 0.001	0.667 ± 0.005	17.9 ± 0.6
55	0.104 ± 0.002	0.547 ± 0.010	5.3 ± 0.1	0.037 ± 0.001	0.660 ± 0.005	17.6 ± 0.6
89	0.105 ± 0.002	0.547 ± 0.010	5.2 ± 0.1	0.038 ± 0.001	0.655 ± 0.004	17.3 ± 0.6
144	0.106 ± 0.002	0.547 ± 0.010	5.2 ± 0.1	0.038 ± 0.001	0.651 ± 0.004	17.0 ± 0.6
233	0.107 ± 0.002	0.547 ± 0.010	5.1 ± 0.1	0.039 ± 0.001	0.649 ± 0.005	16.8 ± 0.6
377	0.108 ± 0.002	0.547 ± 0.010	5.0 ± 0.1	0.039 ± 0.001	0.647 ± 0.005	16.7 ± 0.6
610	0.109 ± 0.002	0.546 ± 0.010	5.0 ± 0.1	0.039 ± 0.001	0.646 ± 0.005	16.6 ± 0.6
987	0.110 ± 0.002	0.546 ± 0.010	5.0 ± 0.1	0.039 ± 0.001	0.645 ± 0.005	16.5 ± 0.6
1597	0.110 ± 0.002	0.546 ± 0.010	4.9 ± 0.1	0.039 ± 0.001	0.644 ± 0.005	16.5 ± 0.6
2584	0.110 ± 0.002	0.545 ± 0.010	4.9 ± 0.1	0.039 ± 0.001	0.644 ± 0.005	16.5 ± 0.6
Nasdaq	0.284	0.424	1.5	0.101	0.290	2.9

Table 4.1: Unconditional probability p_u^v and conditional probability p_c^v to have the absolute return of the next time step higher than $v = 1\sigma$ (columns 2 to 4) or than $v = 2\sigma$ (3 rightmost columns). The column τ has the number of days corresponding to the value of θ (Table 2.1). The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. The values of $p_u^{1\sigma}$ represent the vertical dashed lines in Figure 4.2 (left) and are plotted in Figure B.2 (left). The values of $p_c^{1\sigma}$ represent the vertical the dashed lines in Figure 4.3 (left) and are plotted in Figure B.3 (left). The ratios $q_{1\sigma}$ are plotted in Figure 4.4 (left). For $v = 2\sigma$, the values of $p_u^{2\sigma}$ represent the vertical dashed lines in Figure 4.2 (right) and are plotted in Figure B.2 (right). The values of $p_c^{2\sigma}$ represent the vertical the dashed lines in Figure 4.3 (right) and are plotted in Figure B.3 (right). The ratios $q_{2\sigma}$ are plotted in Figure 4.4 (right). For each τ, the values p_u^v and p_c^v are computed over 1000 simulations of 494999 returns for each value of τ and v. The errors for p_u^v and p_c^v correspond to the standard deviation of the corresponding distribution. The errors for q_v are the propagation of the errors of p_u^v and p_c^v.

4.2 Clustering index

4.2.1 Frequency distribution

We propose here a metric to quantify volatility clustering, developed independently from Tseng and Li (2012).

A window of length m is applied on the first $\tilde{T} := T - (T \mod m)$ price returns of a simulation j of length T. There are $L := \left\lfloor \frac{T}{m} \right\rfloor = \frac{\tilde{T}}{m}$ non-overlapping [§] windows of length m. In each window, the number of daily returns above a threshold v is given by a discrete random variable $X : \Omega \to \Omega$, with domain and image $\Omega = \{0, 1, 2, \cdots, m\} \subset \mathbb{N}$. Recording the L realizations of X yields the probability distribution of X whose moments quantify the volatility clustering. Indeed, if the price returns were generated by a Gaussian distribution, we would expect $\mathbb{E}(X) \approx 0.3137 \cdot m$ (for $v = 1\sigma$), i.e. in average one third of the m days in each window would have a return above 1σ. When volatility clustering is present however, most windows have a much smaller number of returns above 1σ, but some windows have almost all of their returns above 1σ. Thus, the more volatility clustering is present, the lower the mean of the distribution and the higher its variance[¶].

Let us take a closer look at how to compute the moments of the distribution. Let $\vec{h} = (0\ 1\ 2\ \cdots\ m)^T$ with $h = \{0, 1, 2, \cdots, m\}$ the set including all the components of $\vec{h}$. In each simulation j, each window i yields one realization of X, an integer $x_{j,i} \in h \subset \mathbb{N}$. The results from all the windows of the simulation j are gathered in the simulation-sample vector $\vec{x}_j$ of dimension $(L \times 1)$, $x_{j,i} \in \vec{x}_j$. For each $\vec{x}_j$, we build a new vector $\vec{w}_j$ of dimension $(m + 1 \times 1)$ with elements $w_{j,i} = \sum_{k=1}^{L} [x_{j,k} = (i-1)], \forall i \in [1, m+1]$, where the first square brackets $[...]$ in the sum are the Iverson brackets with $[P] = 1$ if P is true, 0 otherwise. We see that $w_{j,i}$ contains the number of windows that have h_i days above v in the simulation j. The vector $\vec{w}_j$ captures the frequency distribution of the simulation j. Note that $\sum_{i=1}^{m+1} w_{j,i} = L \ \forall j$.

There are two ways to approach the computations of the moments of the frequency distribution and their errors for N independent simulations.

Method 1. Appending simulations

We see the N simulations as forming one sample of size $\tilde{T} \cdot N$.

Definition 4.1 *The sample mean of the combined N simulations is*

$$\overline{\mu} = \frac{1}{L \cdot N} \sum_{j=1}^{N} \sum_{i=1}^{L} x_{j,i} \ . \tag{4.3}$$

[§] Using $T + 1 - m$ overlapping windows does not yield independent observations.

[¶] With higher volatility clustering comes longer regimes of low volatility, hence the frequency of windows with *no* absolute returns higher than v is higher which skews the mean of the distribution towards $m = 0$. At the same time and for symmetric reasons, the number of windows of size m with close to m absolute returns above v is higher, enlarging the variance of the frequency distribution.

We can also compute $\overline{\mu}$ with the frequency weights. Let $\vec{h}, \vec{w}_j$ be defined as above, $\vec{W} = \sum_{j=i}^{N} \vec{w}_j$ with $\sum_{i=1}^{n+1} W_i = L \cdot N$. Then

$$\overline{\mu} = \frac{\sum_{i=1}^{m+1} W_i \cdot h_i}{\sum_{i=1}^{n+1} W_i} \,. \tag{4.4}$$

Definition 4.2 *The sample variance of the combined N simulations is*

$$\sigma^2 = \frac{1}{L \cdot N} \sum_{j=1}^{N} \sum_{i=1}^{L} (x_{j,i} - \overline{\mu})^2 \,. \tag{4.5}$$

The sample mean, variance and standard deviation are random variables with their respective sampling distribution. We compute the expected value of the sample variance. For more clarity, the $L \cdot N$ realizations $x_{j,i}$ of the random variable X are represented by $X_1, \cdots, X_{L \cdot N}$.

$$\begin{aligned}
\mathbb{E}[\sigma^2] &= \mathbb{E}\left[\frac{1}{L \cdot N} \sum_{k=1}^{L \cdot N} \left(X_k - \frac{1}{L \cdot N} \sum_{l=1}^{L \cdot N} X_l \right)^2 \right] \\
&= \frac{1}{L \cdot N} \sum_{k=1}^{L \cdot N} \mathbb{E}\left[X_k^2 - \frac{2}{L \cdot N} \sum_{l=1}^{L \cdot N} X_l X_k + \frac{1}{(L \cdot N)^2} \sum_{l_1=1}^{L \cdot N} \sum_{l_2=1}^{L \cdot N} X_{l_1} X_{l_2} \right] \\
&= (\sigma^2 + \overline{\mu}^2) - \frac{2}{L \cdot N} \left((L \cdot N - 1)\overline{\mu}^2 + \sigma^2 + \overline{\mu}^2 \right) + \frac{L \cdot N (L \cdot N - 1)\overline{\mu}^2 + L \cdot N (\sigma^2 + \overline{\mu}^2)}{(L \cdot N)^2} \\
&= \sigma^2 \frac{L \cdot N - 1}{L \cdot N}
\end{aligned} \tag{4.6}$$

Definition 4.3 *The unbiased sample variance is defined as*

$$s^2 = \frac{L \cdot N}{L \cdot N - 1} \sigma^2 = \frac{1}{L \cdot N - 1} \sum_{j=1}^{N} \sum_{i=1}^{L} (x_{j,i} - \overline{\mu})^2 \,, \tag{4.7}$$

with $\mathbb{E}[s^2] = \sigma^2$.

Definition 4.4 *The sample standard deviation is defined as*

$$s = \sqrt{\frac{1}{L \cdot N - 1} \sum_{j=1}^{N} \sum_{i=1}^{L} (x_{j,i} - \overline{\mu})^2} \,. \tag{4.8}$$

As the square root in Equation 4.8 does not commute with the expectation operator, bias is reintroduced. For a non normal distribution, there exists no exact correction for this bias (Gurland and Tripathi (1971)). We will thus work with s as defined above.

The errors of these three moments are now derived. The measure of uncertainty of the sample mean is the standard deviation of the sampling distribution of the sample mean, which is referred to as the standard error.

Definition 4.5 *The standard error of the sample mean is*

$$se_{\overline{\mu}} = \frac{s}{\sqrt{L \cdot N}} , \tag{4.9}$$

where s is the sample standard deviation[11].

To compute the standard error of the unbiased sample variance, we first compute the variance of the sample variance. The idea is to see the following decomposition

$$\begin{aligned} \mathrm{Var}(s^2) &= \mathbb{E}[s^4] - \mathbb{E}[s^2] \\ &= \mathbb{E}[(s^2)^2] - \sigma^4 \\ &= \mathbb{E}[(\mathbb{E}[X^2] - \mathbb{E}[X]^2)^2] - \sigma^4 \\ &= \mathbb{E}\left[\left(\frac{1}{L \cdot N}\sum_{k=1}^{L \cdot N} X_k^2 - \left(\frac{1}{L \cdot N}\sum_{k=1}^{L \cdot N} X_k\right)^2\right)^2\right] - \sigma^4 . \end{aligned} \tag{4.10}$$

Definition 4.6 *The unbaised variance of the unbiased sample variance for any distribution is defined as*

$$Var(s^2) = \frac{1}{L \cdot N}\left(\mu_4 - \frac{L \cdot N - 3}{L \cdot N - 1}\sigma^4\right) , \tag{4.11}$$

where μ_4, the fourth central moment, has to be finite. The standard error of the unbiased sample variance is therefore

$$se_{s^2} = \sqrt{\frac{1}{L \cdot N}\left(\mu_4 - \frac{L \cdot N - 3}{L \cdot N - 1}\sigma^4\right)} . \tag{4.12}$$

Equation 4.12 agrees with (Rao et al., 1973, p. 438), Cho et al. (2005) and Cho (2018).

There does not exist any general equation for the standard error of the standard deviation. We can however use the large sample solution presented in (Rao et al., 1973, p. 386), Equation 6a-2-4 with $\tau = s^2$, $g(r) = \sqrt{r}$. Correcting for the missing absolute values in Rao's Equation 6a-2-4, we have

$$se(g(\tau)) \approx |g'(\tau)| \cdot se(\tau) \tag{4.13}$$

and

$$se_s \approx \frac{1}{2\sigma} \cdot se_{s^2} . \tag{4.14}$$

[11] Equation 4.9 can be easily derived by computing $\mathrm{Var}(\overline{\mu})$ and applying basic properties of the variance.

Method 2. Sampling distribution

Each simulation is considered as a sample. We compute the mean, variance and standard deviation of each simulation and form a sampling distribution for each statistic.

Definition 4.7 *The sample mean for simulation j is*

$$\overline{\mu_j} = \frac{1}{L}\sum_{i=1}^{L} x_{j,i} \,. \tag{4.15}$$

The unbiased sample variance for simulation j is

$$s_j^2 = \frac{1}{L-1}\sum_{i=1}^{L}(x_{j,i} - \overline{\mu_j})^2 \,, \tag{4.16}$$

with $\mathbb{E}[s_j^2] = \sigma_j^2$.

The sample standard deviation for simulation j is

$$s_j = \sqrt{\frac{1}{L-1}\sum_{i=1}^{L}(x_{j,i} - \overline{\mu_j})^2} \,. \tag{4.17}$$

As previously mentioned, the measure of uncertainty of the sample mean is referred to as the standard error.

Definition 4.8 *The mean of the sample means is defined in Equation 4.3 as* $\overline{\mu}$. *The standard error of Equation 4.15 is*

$$se(\overline{\mu_j}) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(\overline{\mu_j} - \overline{\mu})^2} \,. \tag{4.18}$$

The mean of the unbiased sample variance over the N simulations is defined as $\overline{s_j^2}$. *The standard error of Equation 4.16 is*

$$se(s_j^2) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(s_j^2 - \overline{s_j^2})^2} \,. \tag{4.19}$$

The mean of the sample standard deviation over the N simulations is defined as $\overline{s_j}$. *The standard error of Equation 4.17 is*

$$se(s_j) = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(s_j - \overline{s_j})^2} \,. \tag{4.20}$$

In summary, the first method supposes that the moments do not depend on the simulation seed. The second method takes into account that each simulation forms a specific sample. Let us take a numerical example to highlight the difference between the two methods: take $T = 494999$, a window of length $m = 200$ and the threshold $v = 1\sigma$. This yields $L = 2474$ realizations of the random variable per simulation. For $N = 1000$ simulations, both method 1 and 2 yield similar moments, as expected. However the standard errors computed from method 1 and method 2 differ significantly with method 2 having standard errors about 17 times as large as method 1. This artifact comes from the fact that the number of simulations has a large impact on the standard errors in method 2 (by construction). As the results should asymptotically not depend on the simulation seed, and because it is faster in the current artificial market to have longer simulations (larger T, hence larger L) than more simulations (larger m), method 1 will be used.

4.2.2 Moments of the frequency distribution

The frequency distribution for $m = 200$ days, $v = 1\sigma$, $N = 1000$ simulations is plotted in Figure 4.5 for each value of τ given in Table 2.1 and for Gaussian returns. For each point (x_τ, y_τ), y_τ represents how many of the $2474 \cdot 1000$ non-overlapping windows of size 200 had x_τ days above 1σ. The corresponding error bar is the square root of the count x_τ as we are in a counting process. The vertical dashed lines represent the sample mean given by Equation 4.3. The frequency distributions clearly depict the volatility clustering present in the time series. Most windows have very few days (out of 200) with a return above 1σ while a significant number of windows has more than 150 days above 1σ for each τ. Volatility clustering can be quantified through the standard deviation of the frequency distribution (Equation 4.8).

Figure 4.6 depicts the dependency of the standard deviation on small window sizes. For each window size, the standard deviations are scaled to the first value (memory of 0 day). The trend for $\tau \geq 1$ day is robust w.r.t. the window size m and the threshold v. The case $\tau = 0$ is more erratic. For $v = 1\sigma$, a change of regime occurs at $m = 150$, with smaller window sizes yielding less volatility clustering for this memory than all the other ones. For $m > 150$, this agent produces more volatility clustering than the next memory lengths. For $v = 2\sigma$, a similar phenomenon is observed, with the change of regime for $\tau = 0$ occurring at $m = 33$. For both thresholds, the trend stabilizes with higher window sizes. The standard deviation as a metric of the volatility clustering developed here does not yield an absolute value for the volatility clustering but allows for a robust comparison between all memories for sufficient large windows.

Figure 4.7 extends the snapshot of Figure 4.6 and shows the evolution of the standard deviation from a window of size $m = 5$ up to $m = 10000$. The simulations are run for every $m \in [5, 1000]$, then for every multiple of 1000 for $m \in [1000, 10000]$. The change of symmetry highlighted in Figure 4.6 (left) can be seen at $m = 150$ on Figure 4.7 (left) and $m = 33$ on Figure 4.7 (right). From this point, the standard deviations do not overlap each other for all m up to $m = 10000$. Figure 4.8 presents the values of the standard deviation for $m = 10000$. As expected from the construction of the metric, the values for $v = 2\sigma$ are lower than for $v = 1\sigma$ (Figure 4.2 showed that returns above 2σ are scarcer than above 1σ for all τ). The results demonstrate that for large enough window sizes, this metric of volatility clustering is robust for both v and all τ.

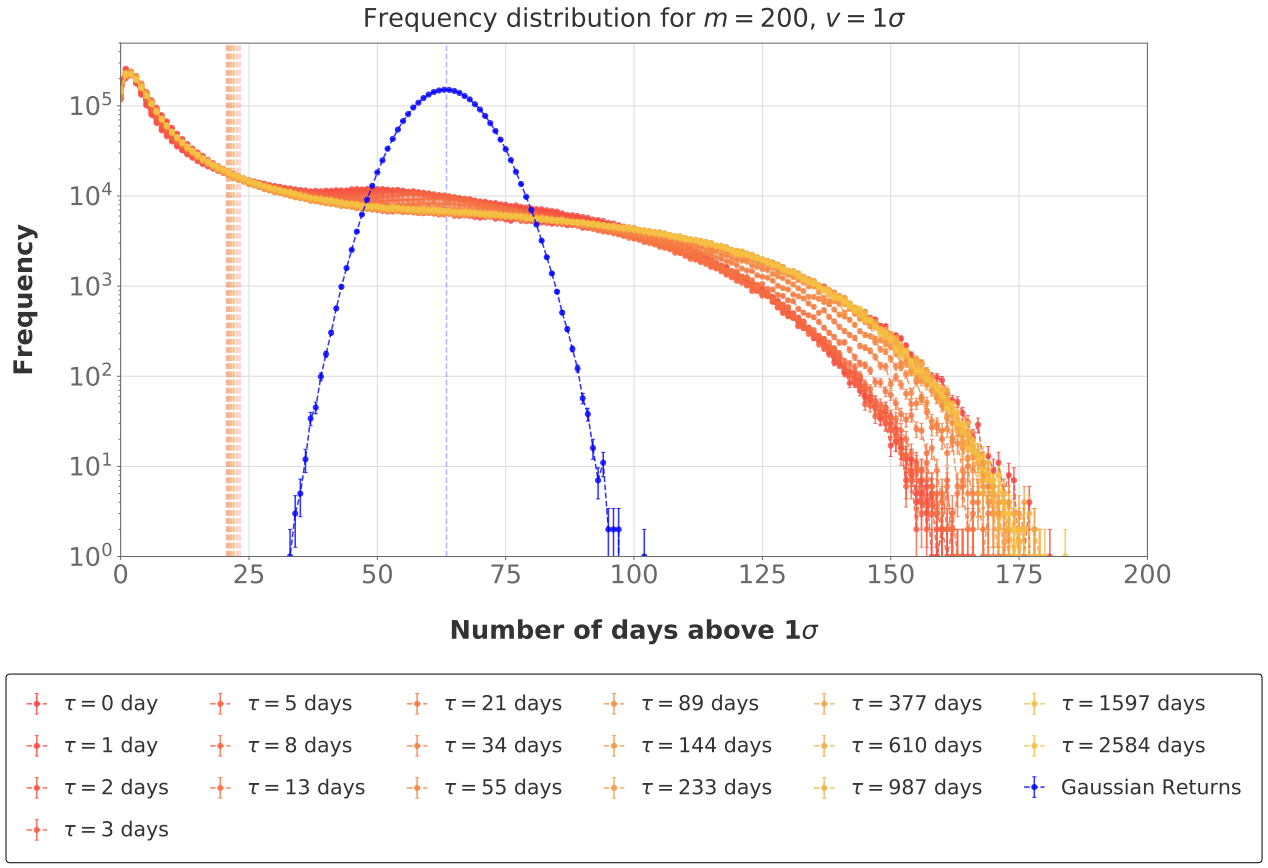


Figure 4.5: Frequency distribution of absolute returns above $v = 1\sigma$ in a window of size $m = 200$. For each value of τ, 1000 simulations of 494999 returns are generated. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. For comparison, the distribution obtained by running the analysis on Gaussian returns is shown in blue. For each point (x_τ, y_τ), y_τ represents how many of the $2474 \cdot 1000$ non-overlapping windows of size 200 had x_τ days above 1σ. The corresponding error bar is the square root of the count x_τ as we are in a counting process (vertical error bar at each point (x_τ, y_τ), more visible for large window sizes). The vertical dashed lines (left of $m = 25$ for all values of τ, in blue for the Gaussian distribution) represent the sample mean given by Equation 4.3. The frequency distributions clearly depict the volatility clustering present in the time series. Most windows have very few days (out of 200) with a return above 1σ while a significant number of windows has more than 150 days above 1σ for each τ. Volatility clustering can be quantified through the standard deviation of the frequency distribution (Equation 4.8).

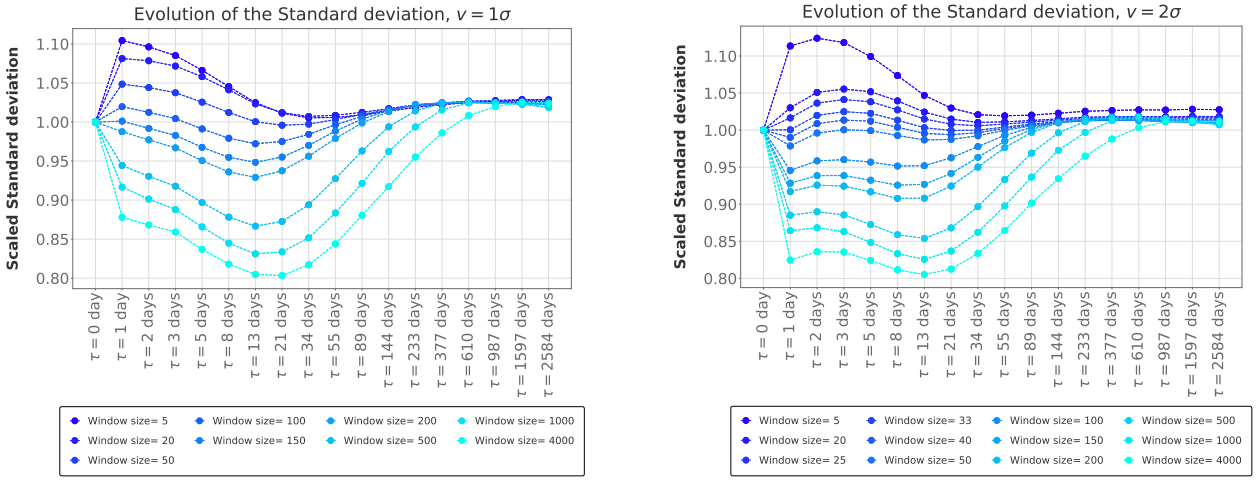


Figure 4.6: Dependency of the standard deviation (Equation 4.8) on small window sizes m for absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). For each m, the trend is scaled to its value at $\tau = 0$ for easier comparison. Both thresholds show a trend inversion for $\tau = 0$. The metric is robust for large enough window sizes. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. For each value of τ, 1000 simulations of 494999 returns are generated.

For returns above 1σ, the memories producing the least clustering are 13 and 22 days (Figure 4.8, left). These memory lengths also produce the least volatility clustering above 2σ (Figure 4.8, right), along with $\tau = 1$ day. For both thresholds, the longer the memory length, the more clustering is generated. Noise traders with no memory behave again as if they had an infinite memory.

The three moments of the frequency distribution depend on the window size m, the memory parameter τ and threshold v. However, it is found that the sample mean $\overline{\mu}$ (Equation 4.3) depends only on the window size and can be considered as a scaling function. Figure 4.9 shows the corresponding data collapse. The critical exponent is $a = 1.00 \pm 10^{-12}$, thus $\overline{\mu} = \exp^{a \log m + b} = m \cdot c$.

Analyzing the ratio of the mean $\overline{\mu}$ to the variance s^2 (Equation 4.7) quantifies the heterogeneity generated by each memory length. Indeed, if the data were produced by an homogeneous Poisson point process, its mean and variance would be equal. We propose to quantify the degree of reflexivity of the market through

$$R = 1 - \frac{\overline{\mu}}{s^2} \,. \tag{4.21}$$

As will be shown in the next section, a slight modification to Equation 4.21 allows to compute an approximation of the branching ratio.

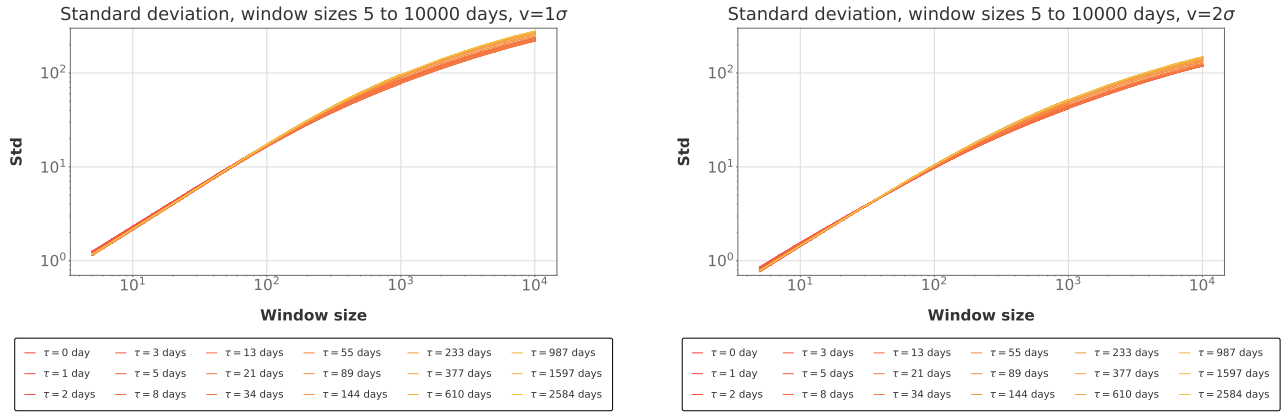


Figure 4.7: Evolution of the standard deviation over increasingly larger window sizes m for absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). The density of lines changing at $m = 1000$ is due to computing the standard deviation (and the error bars) for window sizes that are multiples of 1000 above $m = 1000$. From $m = 5$ to $m = 1000$, every window size is computed. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. For each value of τ, 1000 simulations of 494999 returns are generated.

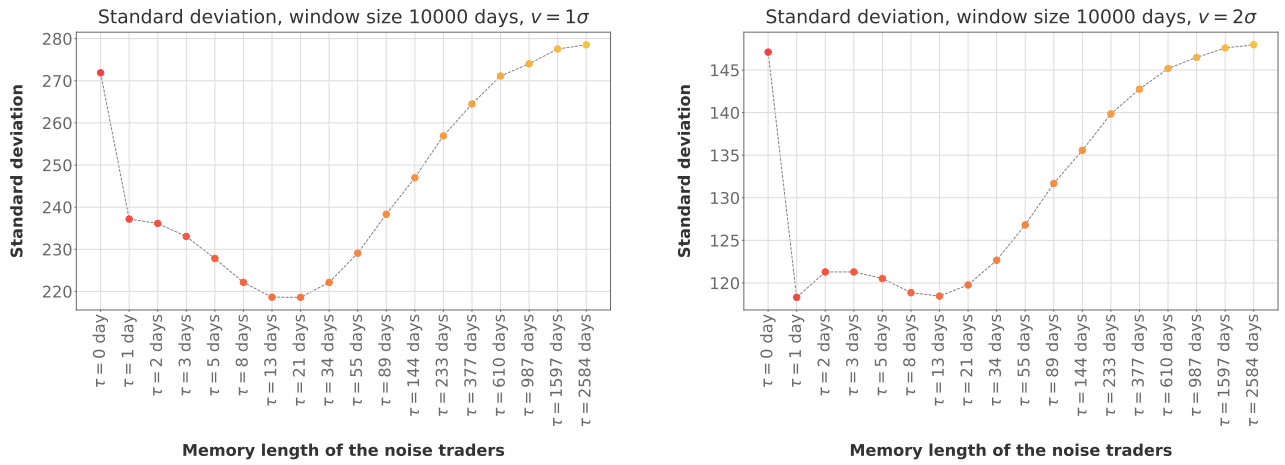


Figure 4.8: Volatility clustering measured through the sample standard deviation of the frequency distribution for absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). Longer memory lengths generate more clustering. Memories of 1, 2, 3, 5, 8 days are non robust w.r.t. v. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. For each value of τ, 1000 simulations of 494999 returns are generated.

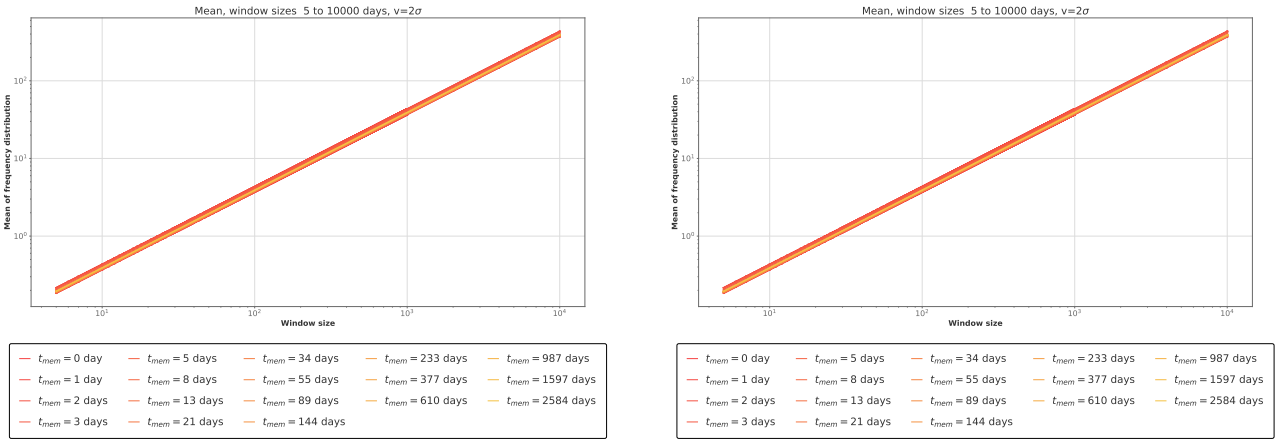


Figure 4.9: Evolution of the sample mean showing data collapse for all τ and for absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). The data collapsed the critical exponent $a = 1.00 \pm 10^{-12}$, thus $\overline{\mu} = \exp^{a \log m + b} = m \cdot c$. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. For each value of τ, 1000 simulations of 494999 returns are generated.

4.3 Non parametric approximation of the Hawkes process

The presence of volatility clustering implies a strong endogenous underlying process. Filimonov and Sornette (2012) were the first to propose a measure to quantify how much of the variation of the price change was due to an endogenous process. Their measure, to be detailed later, is based on a calibrated self-excited conditional Hawkes point process. The Hawkes process was introduced in Hawkes (1971a) and Hawkes (1971b), and has been used in contexts ranging from earthquake prediction (Hawkes (1973), Helmstetter and Sornette (2002)) to clustering of order arrivals in high frequency markets (Hewlett (2006)).

The Hawkes process is a self-exciting point process. A point process is a set of random points in a mathematical subspace. The simplest point process is the Poisson point process for which all events are independent from one another and are produced with constant event arrival rate λ (also called intensity). The Hawkes process extends the Poisson point process with a time dependent intensity $\lambda(t)$

$$\lambda(t) = \mu(t) + \sum_{t_i < t} h(t - t_i) \ . \tag{4.22}$$

The expected number of events in the time interval $[t, t + dt)$ is $\lambda(t)dt$. The function $h(t)$ is the memory kernel and embodies the dependence on past events, which captures the endogenous feedback of the process. The memory kernel $\mu(t)$ represents the background intensity, i.e. the exogeneity. Note the independence of the two factors.

Harris (2002) showed that the Hawkes process can be mapped to the branching ratio. In a branching process, each event is labeled either as an *immigrant* or *descendant*. The immigrants represent the exogenous events and occur with intensity μ. Each immigrant may generate descendants, which in turn may produce further generations of descendants**. By definition, the descendants have an endogenous origin. The average number of descendants per immigrant (endogenous events per exogenous event) is the branching ratio m. As for its nuclear physics counterpart, the branching ratio can be divided into 3 regimes, subcritical ($n < 1$), critical ($m = 1$) and explosive ($n > 1$). In one dimension, the average event rate of Equation 4.22 is (see among others Helmstetter and Sornette (2002), Bacry et al. (2015))

$$a\Lambda = \frac{\mu}{(1-n)} \, . \tag{4.23}$$

The branching ratio n can be computed by reverse engineering every point to its corresponding zeroth-generation immigrant, see for example Zhuang et al. (2002). The limitations of this approach are highlighted in Sornette and Utkin (2009). From Equation 4.22 and Equation 4.23 it can be seen that the L^1 norm of the memory kernel $h(t)$ satisfies $\|h(t)\| = \int_{\mathbb{R}} |h(t)| dt = n$. As the log-likelihood function for the Hawkes process is known in close form (Ogata (1978)), $h(t)$ can be predicted with a maximum likelihood estimation.

Using an exponential kernel, Filimonov and Sornette (2012) showed that the level of endogeneity on the E-mini S&P500 futures contracts (on a high frequency timescale) significantly increased from 1998 to 2010. Hardiman et al. (2013) criticized this approach noting that Bacry et al. (2012) found that the self-excitation regimes in these timescales showed a long range memory. Therefore, the Hawkes kernel should be calibrated with a power-law decay. In response, Filimonov and Sornette (2015) highlighted the pitfalls of a power-law approach, showing the impact of outliers, strong edge effect and the impact of the regularization of the power-law kernel. Wheatley et al. (2019) show the impact of the treatment of trends and external shocks (endo-exo problem) in the computation of the branching ratio. To bypass the selection of the kernel, Hardiman and Bouchaud (2014) presented an approximation of the branching ratio for self-exciting Hawkes point processes in one dimension depending only on the variance and mean of the event count in very large windows. They find the relation

$$n \approx 1 - \sqrt{\frac{\mathbb{E}[N_W]}{\mathrm{Var}[N_W]}} \, . \tag{4.24}$$

This result overlaps greatly with our approach derived in subsection 4.2.2, and in particular Equation 4.21. It is worth noting that the debate between Filimonov and Sornette and Hardiman and Bouchaud is based on high-frequency data, whereas our market has a clock of about one day per time-step (with the parameter choice given in Table 1.1). However, the measure proposed in Equation 4.24 depends only on the window size. Also, our data is as for high frequency data on the finest information time scale as no exchange of information or opinion is possible between two consecutive returns. Note that the estimator in Equation 4.24 is noisy and has little power.

**In Earthquake terminology, the immigrant can be seen as the main event and the descendants as aftershocks.

Figure 4.10 shows a smooth evolution of the branching ratio $n(\tau)$ over larger window sizes. Note that the Hawkes processes are calibrated on the absolute returns above $v = 1\sigma$ (left) and on the absolute returns above $v = 2\sigma$ (right). As in subsection 4.2.2, the simulations are run for every $m \in [5, 1000]$, then for every multiple of 1000 for $m \in [1000, 10000]$. The endogeneity grows steadily up to a window size of 1000 days after which the trend flattens and even slightly decreases. Window sizes larger than 10000 were not investigated in details, due to time and computation power constraints. The window size 100000 was computed for $v = 2\sigma$ and is shown in Figure B.1 (page 91). The values are slightly lower than for a window of size 10000 for all memory parameters but present the same trend. The branching ratio, approximated with Equation 4.24, shows a consistent and converging behavior at very large windows. The necessity for the approximation of the branching ratio to converge only for large window sizes may be tracked back to the necessity to cover all the correlation of the returns.

For both thresholds and all memory parameters, the branching ratio is high but sub critical. The values are presented in Figure 4.11 and given in Table 4.2. Consistently with the results shown in the preceding chapters, the noise traders with no memory seem behave as noise traders with a very long memory (see Equation 2.4). The shorter strictly positive memories create the least endogeneity. The longer memories tend to create the most endogenous markets, regardless of the select thresholds v.

The trend of the branching ratio is very similar to our metric for volatility clustering (Figure 4.8) for both v. Using the Fisher transformation, the correlation for $v = 1\sigma$ between the branching ratio and our metric of volatility clustering is $\rho = 0.981^{+0.012}_{-0.007}$ with a z-score of 9.0, statistically confirming the visual intuition and rejecting the null hypothesis at 9σ. For $v = 2\sigma$, we find $\rho = 0.96^{+0.02}_{-0.01}$ and a z-score of 7.8. These results can be attributed to the scaling behavior of the mean.

τ	**Branching ratio** $v = 1\sigma$	**Branching ratio** $v = 2\sigma$
0	0.8802 ± 0.0004	0.8701 ± 0.0004
1	0.8562 ± 0.0005	0.8233 ± 0.0006
2	0.8568 ± 0.0005	0.8283 ± 0.0005
3	0.8559 ± 0.0005	0.8298 ± 0.0005
5	0.8543 ± 0.0005	0.8315 ± 0.0005
8	0.8523 ± 0.0005	0.8324 ± 0.0005
13	0.8517 ± 0.0005	0.8350 ± 0.0005
21	0.8528 ± 0.0005	0.8384 ± 0.0005
34	0.8554 ± 0.0005	0.8427 ± 0.0005
55	0.8594 ± 0.0005	0.8473 ± 0.0005
89	0.8642 ± 0.0004	0.8523 ± 0.0005
144	0.8682 ± 0.0004	0.8558 ± 0.0005
233	0.8725 ± 0.0004	0.8595 ± 0.0004
377	0.8756 ± 0.0004	0.8620 ± 0.0004
610	0.8781 ± 0.0004	0.8641 ± 0.0004
987	0.8791 ± 0.0004	0.8653 ± 0.0004
1597	0.8804 ± 0.0004	0.8661 ± 0.0004
2584	0.8807 ± 0.0004	0.8665 ± 0.0004

Table 4.2: Asymptotic values for the branching ratio, approximated with Equation 4.24, for a window of size 10000 (yielding 49 windows in each of the 1000 simulations), calibrated on the absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). For both thresholds and all memory parameters, the branching ratio is high but sub critical. The values are plotted in Figure 4.11. Consistently with the results shown in the preceding chapters, the noise traders with no memory seem to behave as noise traders with a very long memory (see Equation 2.4). The shorter strictly positive memories create the least endogeneity. The longer memories tend to create the most endogenous markets, regardless of the select thresholds v. The trend of the branching ratio is very similar to our metric for volatility clustering (Figure 4.8) for both v. The trend for a window of size 100000 is shown in Figure B.1, page 91. The proximity of the results for both thresholds might tracked back to the similarity of the trend of the sample standard deviation for both thresholds (Figure 4.8). Also, the set of absolute returns above $v = 2\sigma$ is included in the set of absolute returns above $v = 1\sigma$. For the proportion of absolute returns above 1σ that are also above 2σ, see Figure 4.2 and Figure B.2. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. For each value of τ, 1000 simulations of 494999 returns are generated.

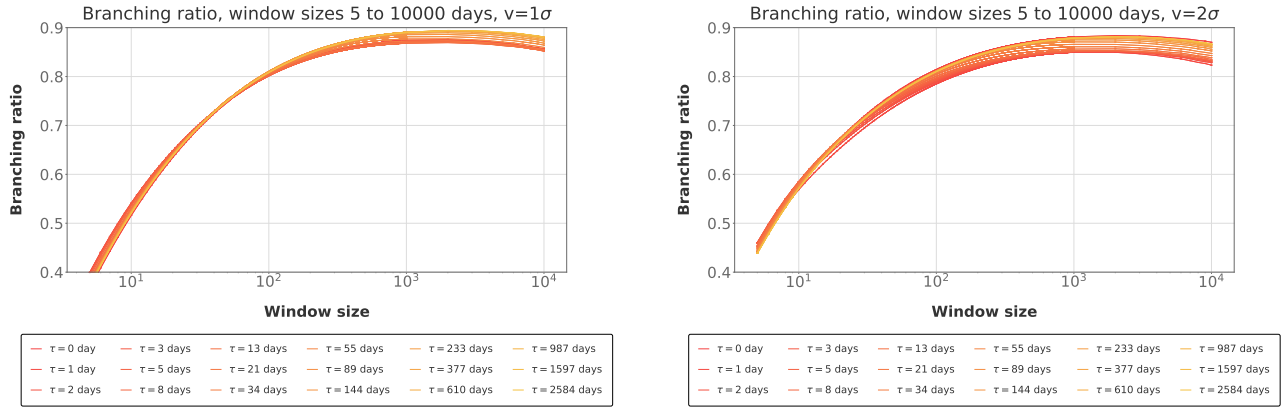


Figure 4.10: Evolution of the branching ratio (Equation 4.24) over very long windows. The Hawkes process is calibrated on absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). The trend stabilizes from window size 1000. A snapshot of the values for window size 10000 is presented in Figure 4.11. The necessity for the approximation of the branching ratio to converge only for large window sizes may be tracked back to the necessity to cover all the correlation of the returns. The proximity of the results for both thresholds might tracked back to the similarity of the trend of the sample standard deviation for both thresholds (Figure 4.8). Also, the set of absolute returns above $v = 2\sigma$ is included in the set of absolute returns above $v = 1\sigma$. For the proportion of absolute returns above 1σ that are also above 2σ, see Figure 4.2 and Figure B.2. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. For each value of τ, 1000 simulations of 494999 returns are generated.

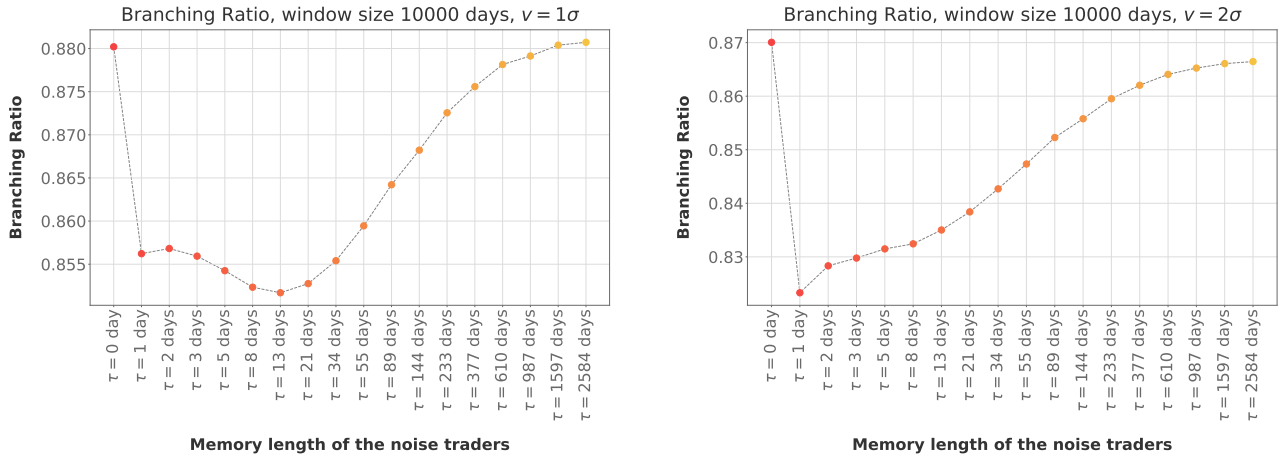


Figure 4.11: Asymptotic values for the branching ratio, approximated with Equation 4.24, for a window of size 10000 (yielding 49 windows in each of the 1000 simulations), for $v = 1\sigma$ (left) and $v = 2\sigma$ (right). For both thresholds and all memory parameters, the branching ratio is high but sub critical. The values are listed in Table 4.2. Consistently with the results shown in the preceding chapters, the noise traders with no memory behave as embodied by a infinite memory (see Equation 2.4). The shorter strictly positive memories create the least endogeneity. The longer memories tend to create the most endogenous markets, regardless of the select thresholds v. The trend of the branching ratio is very similar to our metric for volatility clustering (Figure 4.8) for both v. The trend for a window of size 100000 is shown in Figure B.1, page 91. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length. For each value of τ, 1000 simulations of 494999 returns are generated.

Chapter 5

Impact of Noise Traders with Heterogeneous Time Scales on the Frequency and Size of Bubbles, Volatility Clustering and Endogeneity

The time scales present in real markets have a strong impact on the volatility clustering. It is known that at least two different time scales are necessary to produce volatility persistence (LeBaron (2006))*. Also, Guillaume et al. (1997) show that high frequency markets are composed of heterogeneous agents whose different time horizons are linked to the volatility clustering. Guillaume and Pictet (1995) demonstrate that the presence of the interplay of many memories in real markets. Tumminello et al. (2012) establish that the behaviors of investors tend to cluster w.r.t. their memory.

All of these observations lead us to extend the properties of the noise trader agent of the ABM presented in chapter 1. In the preceding chapters, all noise traders shared the same memory. The noise trader agents are now split into N_G^n groups with different memory lengths. Within each group, all noise traders share the same memory length τ (Equation 1.19). All noise traders share the same coupling strength, regardless of their memory length.

This finer grid of memory lengths enhances the heterogeneity of time scales, permitting to probe its impact on the number of bubbles and their size (section 5.3), on the autocorrelations of the signed returns (subsection 5.4.1), on the decay of the autocorrelations coefficients of the absolute returns (subsection 5.4.2), on volatility clustering measured through the standard deviation of the frequency distribution and on the endogeneity of the heterogeneous markets (section 5.5). Initial results for markets where each group of noise traders sharing the same memory is endowed by its own coupling strength are presented in section 5.6.

5.1 Wealth dynamics

To study the impact of the heterogeneity of time scales on the number of bubbles and volatility clustering, six different markets are analyzed. The markets are referred to as H2, H3, H5, H9,

*The market defined in chapter 1 fulfilled this criterion with two different time scales, the fundamentalist basing their decision on the long term view of the market, and all noise traders sharing the same predefined memory length.

H17, H40. Each market Hi is composed of one fundamentalist agent and i groups of noise traders. All fundamentalists share the same utility function and expectations and are thus considered to be identical, as in chapter 1. Their behavior is represented by a single fundamentalist. The noise traders are split into N_G^n groups. Within each group, all noise traders share the same memory length τ (Equation 1.19). Noise traders cannot switch between groups and keep the memory assigned to them. The market Hi is constituted of noise traders with i different memories. The wealth of each group of noise traders is W^{n^i}. The sum of the initial wealths of all the noise traders groups is equal to the initial wealth of the fundamentalist. Note that all the noise traders share a common coupling strength κ_t (Equation 1.21), regardless of their memory length.

With N_G^n noise trader agents, the market clearing condition (Equation 1.31) evolves to

$$\sum_{i=1}^{N_G^n} \Delta D_{t\to t+1}^{n^i} + \Delta D_{t\to t+1}^{f} = 0\ . \tag{5.1}$$

The quadratic equation for the equilibrium market price P_{t+1}

$$a_{t+1}P_{t+1}^2 + b_{t+1}P_{t+1} + c_{t+1} = 0\ , \tag{5.2}$$

has the new parameters

$$a_{t+1} = \frac{1}{P_t}\Big[\sum_{i=1}^{N_G^n} \nu_t^{nf} x_t^{n^i}(x_{t+1}^n - 1) + x_t^f\left(x_{min}^f - 1\right)\Big] \tag{5.3}$$

$$b_{t+1} = x_t^f \frac{1}{\gamma\sigma_{R_{ex}}^2}\frac{d_{t+1}(1+r_d)}{P_t} + x_{min}^f\left[x_t^f\left(\frac{d_{t+1}}{P_t} - R_f\right) + R_f\right] + \sum_{i=1}^{N_G^n} \nu_t^{nf} x_{t+1}^{n^i}\left[x_t^n\left(\frac{d_{t+1}}{P_t} - R_f\right) + R_f\right] \tag{5.4}$$

$$c_{t+1} = \frac{d_{t+1}(1+r_d)}{\gamma\sigma_{R_{ex}}^2}\left[x_t^f\left(\frac{d_{t+1}}{P_t} - R_f\right) + R_f\right]. \tag{5.5}$$

5.2 Considered heterogeneities of noise traders

The distribution of the time scales present in the six markets considered is listed in Table 5.1. The market Hi is formed of $N_G^n = i$ noise trader groups and has i different memory lengths. In this regard, the market presented in chapter 1 and used in all the preceding chapters will be referred to as H1(τ), τ representing the common memory assigned to all the noise traders.

The number of days for the memories of the 17 groups of noise traders in the market H17 are chosen to follow the Fibonacci sequence, producing a converging scaling between two successive memory lengths. Also, these are the memories, with $\tau = 0$, analyzed in the preceding chapters. Then, the memories for the markets H9, H5, H3 and H2 are selected as increasingly smaller subsets of the market H17 such that the scaling between the memories within each market remains constant, and

the shortest and longest memories are shared for all markets. After running the analysis on these 5 markets, it became apparent that a market with a larger heterogeneity should be considered. In this regard, the market H40 was created with memories spanning the same range as the existing markets, but on a finer grid[†].

The average memory length of the different memories of the noise trader agents of market Hi, denoted $\overline{\tau(\mathrm{H}i)}$, differs greatly between the markets considered (last row of Table 5.1). The choice to construct markets with a constant scaling between each successive memory rather than with a constant market average memory was motivated by the fact that it is straightforward to compare the results obtained for each market Hi with a market formed as in chapter 1 with one fundamentalist and one noise trader agent with memory $\overline{\tau(\mathrm{H}i)}$. On the other hand, comparing markets with equal average memory but different intrinsic structures requires creating many different markets with different distributions of the memories for each level of heterogeneity[‡].

The market composed of one fundamentalist trader and of noise traders all sharing the same memory length $\overline{\tau(\mathrm{H}i)}$ is referred to as H1$(\overline{\tau(\mathrm{H}i)})$.

5.3 Impact of market heterogeneity on the frequency and size of bubbles

The impact of the heterogeneity of time scales of the noise traders on the frequency and size of bubbles is analyzed. The metric used is defined in Equation 2.2. The minimal number of days between two peaks is constant and set to $k = 250$ days. The parameters $s = 250$ and $thres = 0.4$ of the metric are as in section 2.5.

As mentioned, for the analysis to be robust each level of heterogeneity should be analyzed with different underlying average memory and different distributions of these memories. However, by comparing the results for Hi with a market representing the average memory of Hi, the results strongly gain significance.

Figure 5.1 (left) shows the average number of bubbles for each level of heterogeneity given in Table 5.1. The values are calculated over 1000 simulations of 494999 returns generated for the corresponding market. The error bars are the standard deviation of the distribution of the number of bubbles. Figure 5.1 (right) displays the average number of bubbles for the six markets H1$(\overline{\tau(\mathrm{H}i)})$. The results obtained in section 2.5 for long memories agree with the trend observed in Figure 5.1 (right)[§].

[†]Creating a sixth market as an extension of market H17 with the memories scaling as the Fibonacci sequence would rapidly yield very long memories. When an approach for a sixth market with a larger number of agents was first considered, the simulations were run on simulations of 15000 time steps (20000 minus a burn in period of 5000), making the impact of very large memories statistically insignificant. Even though all noise traders trade at every time step, regardless of their memory, the noise traders with memory $2584 \cdot 1.6^3 = 10584$ would only yield a behavior that could be attributed to his memory for the last 4416 time steps, i.e. $1/3$ of the price path.

[‡]Constructing different markets with a common shortest and longest memory, a constant average memory and a constant scaling within each market is not feasible (with integer memory lengths). Say we start with a market of heterogeneity 2, with the memories 1 day and 2584 days. The next market (H3) would have to have memories of 1, 1293, 2584. After a few rounds it becomes obvious that the high memories cannot be compensated with many small memories while keeping a scale factor between the memories.

[§]See the values of $\overline{\tau(\mathrm{H}i)}$ in the last row of Table 5.1

Heterogeneity	**H2**	**H3**	**H5**	**H9**	**H17**	**H40**
Number of noise trader groups	2	3	5	9	17	40
Scaling ratio between 2 successive memories	2584	≈ 50	≈ 6.85	≈ 2.64	≈ 1.618	≈ 1.2
Memory length τ	1	1	1	1	1	1
					2	2
				3	3	3
						4
						4
					5	5
						6
						7
			8	8	8	
						9
						11
					13	13
						15
						19
				21	21	
						22
						27
						32
					34	
						39
						46
		55	55	55	55	55
						67
						80
					89	
						96
						115
						138
				144	144	
						166
						199
					233	
						238
						286
						343
			377	377	377	
						412
						495
						593
					610	
						712
						855
				987	987	
						1025
						1231
						1477
					1597	
						1772
						2126
	2584	2584	2584	2584	2584	2584
Average Memory $\overline{\tau(\mathrm{H}i)}$	1293	880	605	464	398	383

Table 5.1: Each column Hi lists the i memories of the i noise trader groups forming the market Hi. The memories within each market have a constant scaling between them, given in the second row. The average of all the memories for the market Hi is referred to as $\overline{\tau(\mathrm{H}i)}$ and is given in the last row.

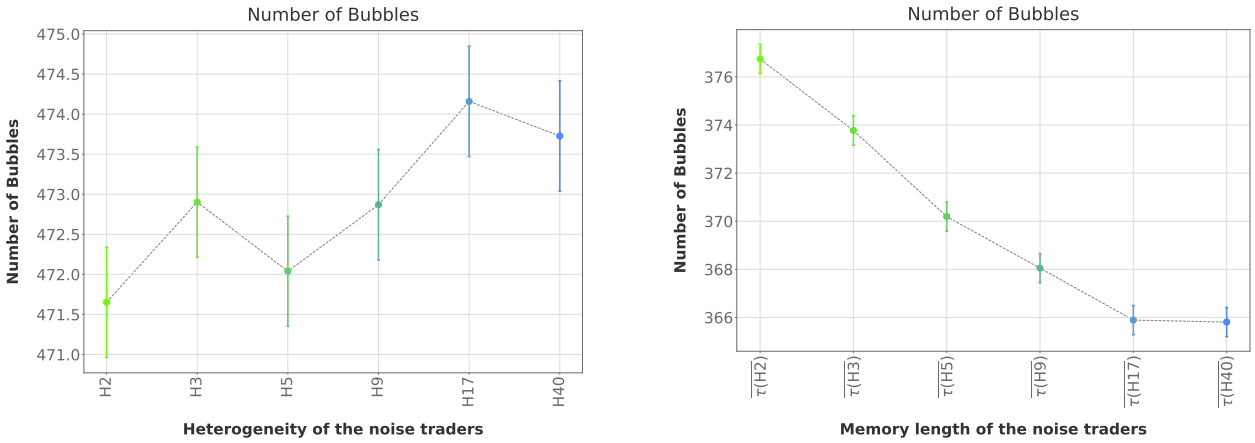


Figure 5.1: Left: number of bubbles for markets Hi with one fundamentalist agent and i group of noise traders with i different memories. Right: number of bubbles for markets H1($\overline{\tau(\text{H}i)}$) composed of one fundamentalist and one noise trader agent (as in the ABM presented in chapter 1 and analyzed in chapter 2) with memory $\overline{\tau(\text{H}i)}$, the average memory of each market Hi (right). The ratio between both results is presented in Figure 5.3 (left). For both cases (the 6 markets Hi and the 6 markets H1($\overline{\tau(\text{H}i)}$)), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

Figure 5.2 shows the average log drawdown computed with Equation 2.3 for each market Hi (left) and for a market with the corresponding average memory (right). No clear trend appears for both analyses. A similar conclusion was drawn for the longer memories in Figure 2.10.

To grasp the intrinsic impact of the heterogeneity of the market Hi, the average number of bubbles and log drawdown for each Hi is divided by the values obtained for the corresponding market H1($\overline{\tau(\text{H}i)}$). The values are listed in Table 5.2 and shown in Figure 5.3. The results for the number of bubbles show that a higher heterogeneity of memories in the market increases the number of bubbles created up to a heterogeneity of 17 time scales for the noise traders. For the market with two groups of noise traders (H2), the number of bubble is $25.2 \pm 0.2\%$ higher than for the market H1($\overline{\tau(\text{H2})}$) (a market with one group of noise traders with the memory set to the average of the memories present in H2). This increase of heterogeneity (doubling the number of memories of the noise traders from 1 to 2) has a higher impact on the number of bubbles than all the other increases of heterogeneity considered. For example, a heterogeneity of 5 memories (H5) generates $27.5 \pm 0.2\%$ more bubbles than the market H1($\overline{\tau(\text{H5})}$). This represents however only $\approx$ 1% more bubbles than the market H2. Figure 5.3 (right) shows that the average drawdowns, though statistically indifferent between heterogeneities, is on average 20% lower than when noise traders all share the same memory.

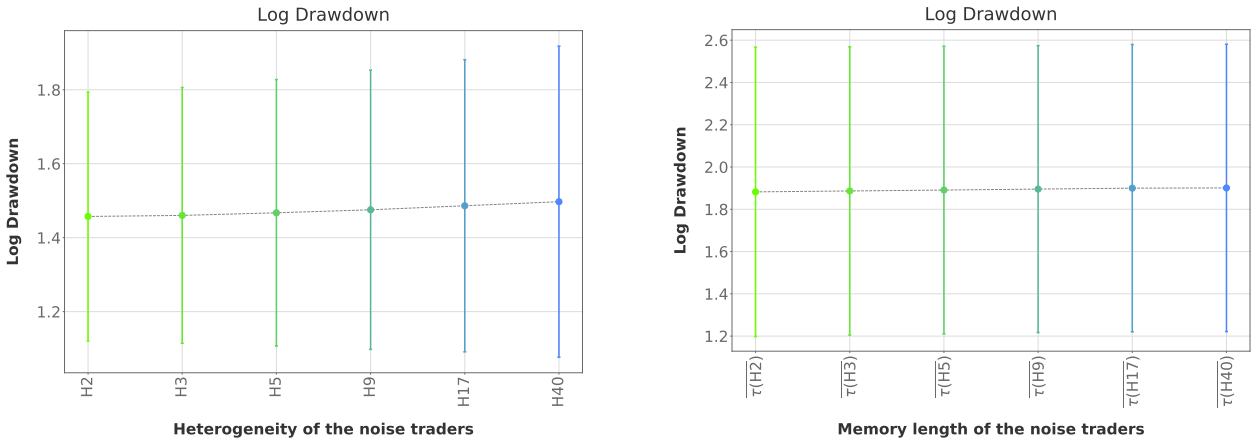


Figure 5.2: Left: average log drawdown for markets Hi with one fundamentalist agent and i group of noise traders with i different memories. Right: average log drawdown for markets H1$(\overline{\tau(\mathrm{H}i)})$ composed of one fundamentalist and one noise trader agent (as in the ABM presented in chapter 1 and analyzed in chapter 2) with memory $\overline{\tau(\mathrm{H}i)}$, the average memory of each market Hi. The ratio between both results is presented in Figure 5.3 (right). For both cases (the 6 markets Hi and the 6 markets H1$(\overline{\tau(\mathrm{H}i)})$), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

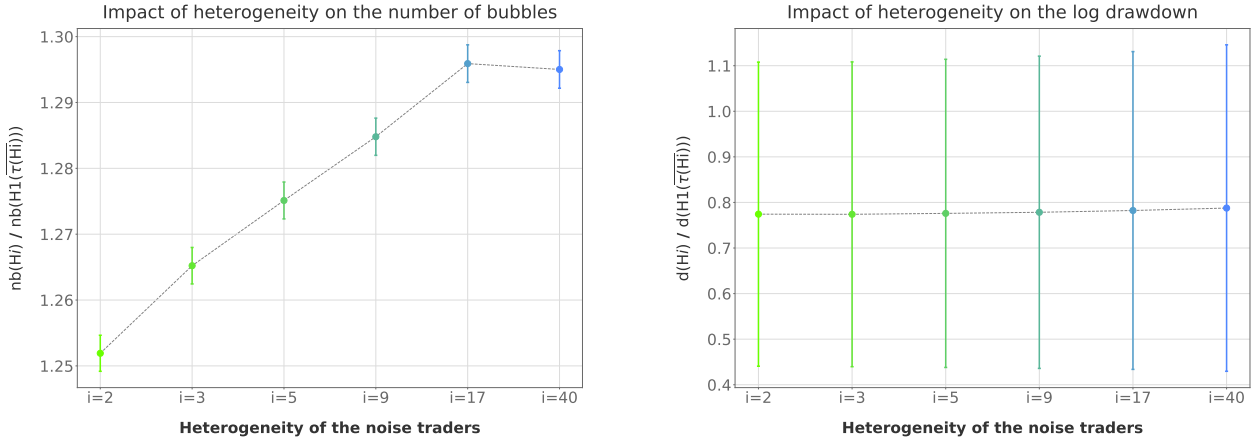


Figure 5.3: Ratio of the number of bubbles (left) and of the log drawdowns (right) obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with the results obtained for markets H1$(\overline{\tau(\mathrm{H}i)})$ with all the noise traders sharing the same memory with their memory set to the average memory $\overline{\tau(\mathrm{H}i)}$ of the heterogeneous market Hi. See the text and Table 5.2 for an analysis of the results. For both cases (the 6 markets Hi and the 6 markets H1$(\overline{\tau(\mathrm{H}i)})$), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

	$i=2$	$i=3$	$i=5$	$i=9$	$i=17$	$i=40$
$\frac{\mathrm{Nb}(\mathrm{H}i)}{\mathrm{Nb}(\mathrm{H1}(\overline{\tau(\mathrm{H}i)}))}$	1.252 ± 0.002	1.265 ± 0.002	1.275 ± 0.002	1.284 ± 0.002	1.295 ± 0.002	1.295 ± 0.002
$\frac{\mathrm{d}(\mathrm{H}i)}{\mathrm{d}(\mathrm{H1}(\overline{\tau(\mathrm{H}i)}))}$	0.7 ± 0.3	0.7 ± 0.3	0.7 ± 0.3	0.7 ± 0.3	0.7 ± 0.3	0.7 ± 0.3

Table 5.2: Ratio of the number of bubbles (Nb) and of the log drawdowns (d) obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with the results obtained for markets H1$(\overline{\tau(\mathrm{H}i)})$ with all the noise traders sharing the same memory with their memory set to the average memory $\overline{\tau(\mathrm{H}i)}$ of the heterogeneous market Hi. The values listed here are plotted in Figure 5.3. Doubling the number of different memories for the noise traders agent from one to two increases the number of bubbles by $25.2 \pm 0.2\%$. This increase of heterogeneity has a higher impact on the number of bubbles than all the other increases considered. For example, a heterogeneity of 5 memories generates $\approx 1\%$ more bubbles than a market with only two memories. Heterogeneous markets tend to lower the size of the bubbles by 30%, though the results are not significant. For both cases (the 6 markets Hi and the 6 markets H1$(\overline{\tau(\mathrm{H}i)})$), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

5.4 Time series analysis

5.4.1 Autocorrelation of the signed returns

The autocorrelations functions of the signed returns generated by the six markets H2, H3, H5, H9, H17 and H40 are analyzed.

In section 3.1, all the memories displayed significant autocorrelation coefficients up to lag 100. Markets with a higher heterogeneity of time scales also produce significant autocorrelation coefficients up to lag 100 (Figure 5.4). The partial autocorrelation coefficients (PACF) of the signed returns are pictured in Figure 5.5.

Table C.1 (page 101) presents a closer look at the behavior of the coefficients near the crossing of the i.i.d. null hypothesis (cut-off lag). All markets Hi display the same cut-off lag. Also, their lowest coefficient is consistently at lag 20 with a value almost identical between the different heterogeneities. A similar behavior is found for the markets H1$(\overline{\tau(\mathrm{H}i)})$, see Figure B.5 and Table C.2. Comparing Table C.1 with Table C.2 shows that the PACF coefficients have a faster decay in markets Hi than in markets H1$(\overline{\tau(\mathrm{H}i)})$. Even though the PACF coefficients reach lower values in the markets Hi than in markets H1$(\overline{\tau(\mathrm{H}i)})$, these minima occur much earlier (at lag 3 for Hi, between lag 11 and 12 for H1$(\overline{\tau(\mathrm{H}i)})$). This out of sample analysis shows that market with a higher heterogeneity of time scales may correspond to an ARMA(p,q) dynamic with lower p, q coefficients than H1 markets.

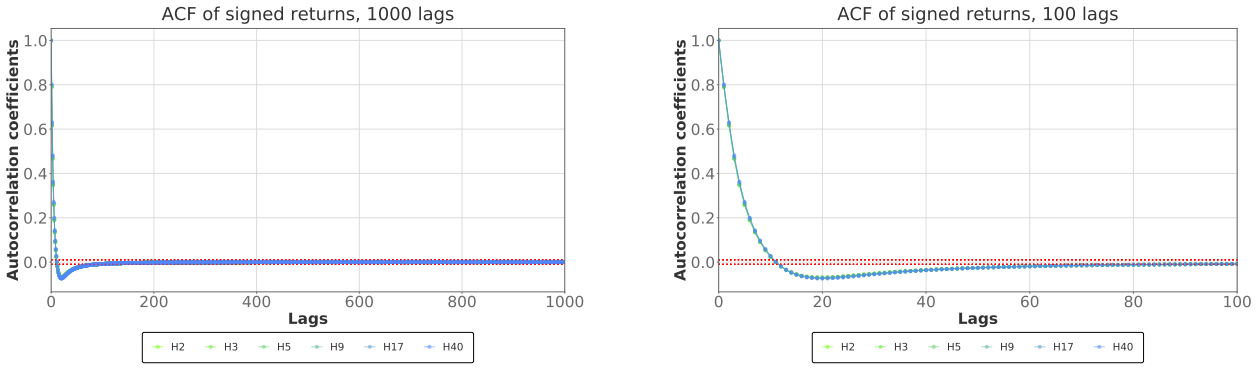


Figure 5.4: ACF coefficients of signed returns up to lag 1000 (left) with a zoom on the first 100 lags (right), for each market Hi listed in Table 5.1. The coefficients displayed correspond to the average coefficient over 1000 simulations of 494999 returns generated for each Hi. The errors, not clearly visible at this scale, are the standard error of the distribution of the 1000 coefficients for each lag. The red horizontal dashed lines represent the 95% confidence level to accept null hypothesis that the coefficients are i.i.d. The first 100 lags are significantly non vanishing, indicating a strong non stationarity in the returns generated. The first lag for which the ACF coefficient is zero for each Hi and the lag with the lowest coefficients and its value are given in Table C.1, along with an analysis of the difference with the results obtained for markets H1$(\overline{\tau(\mathrm{H}i)})$. For both cases (the 6 markets Hi and the 6 markets H1$(\overline{\tau(\mathrm{H}i)})$), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

5.4.2 Autocorrelation of the absolute returns

The ACF of the absolute returns for each market Hi are displayed in Figure 5.6 and Figure 5.7. Figure B.4 shows the decay of the absolute returns for the markets H1$(\overline{\tau(\mathrm{H}i)})$. The markets H$i$ have lower cut-off lags and thus a faster decay than H1$(\overline{\tau(\mathrm{H}i)})$ for all heterogeneities analyzed. As discussed in chapter 3, this is a mark that less volatility clustering is present in the markets Hi than in their H1 counterparts. To quantify this decrease of volatility clustering, the decays exponent β are computed as in section 3.2[¶].

To pinpoint a constant and representative decay exponent β, the combined p-value for different ranges are probed. The ranges have a starting lag $s_{\mathrm{H}i}$ and end lag $f_{\mathrm{H}i}$. The ranges are then shrunk by keeping $f_{\mathrm{H}i}$ fixed and increasing $s_{\mathrm{H}i}$ up to $f_{\mathrm{H}i} - m$, where m is the minimal amount of lag to be fitted. As in section 3.2 (Figure 3.5), the lags up to lag 20 display a faster decay. They are disregarded and the initial lag set at $s_{\mathrm{H}i} = 20$ (black dashed vertical line in Figure 5.7). At first, $f_{\mathrm{H}i}$ is naively defined as the lag preceding the first lag for which the ACF coefficient is below the i.i.d. null hypothesis for a given Hi (red horizontal dashed line in Figure 5.7). For $m = 200$ and $m = 100$, for any market Hi, no range yields a combined p-value above 0. Setting the cut-off at 0.01 (black

[¶]The results of the GARCH analysis, as described in section 3.2, on the markets Hi and H1$(\overline{\tau(\mathrm{H}i)})$ are shown in Figure B.6. As for the results presented in section 3.2 (Figure 3.4), no clear trend can be deduced and the results are not investigated further.

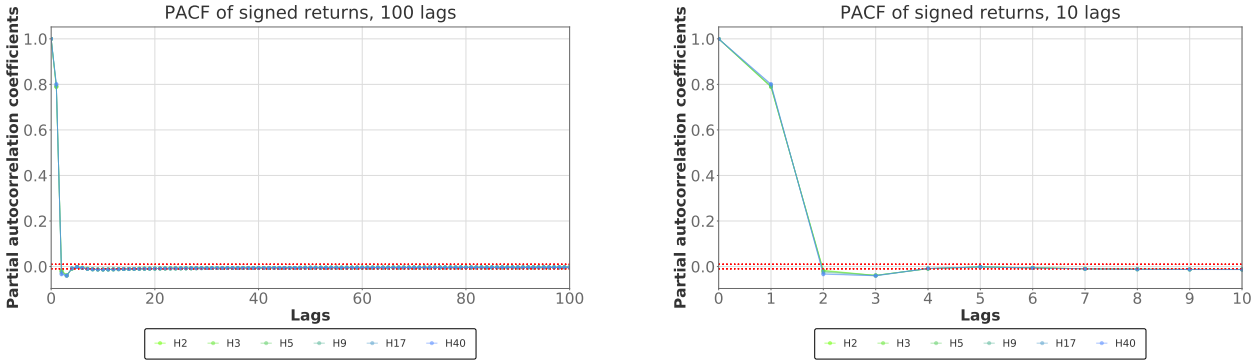


Figure 5.5: PACF coefficients of signed returns up to lag 100 (left) with a zoom on the first 10 lags (right), for each market Hi listed in Table 5.1. The coefficients displayed correspond to the average coefficient over 1000 simulation of 494999 returns generated for each Hi. The errors, not clearly visible at this scale, are the standard error of the distribution of the 1000 coefficients for each lag. The red horizontal dashed lines represent the 95% confidence level to accept null hypothesis that the coefficients are i.i.d. The first lag for which the PACF coefficient is zero for each Hi and the lag with the lowest coefficients and its value are given in Table C.1, along with an analysis of the difference with the results obtained for markets H1$(\overline{\tau(\mathrm{H}i)})$. For both cases (the 6 markets Hi and the 6 markets H1$(\overline{\tau(\mathrm{H}i)})$), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

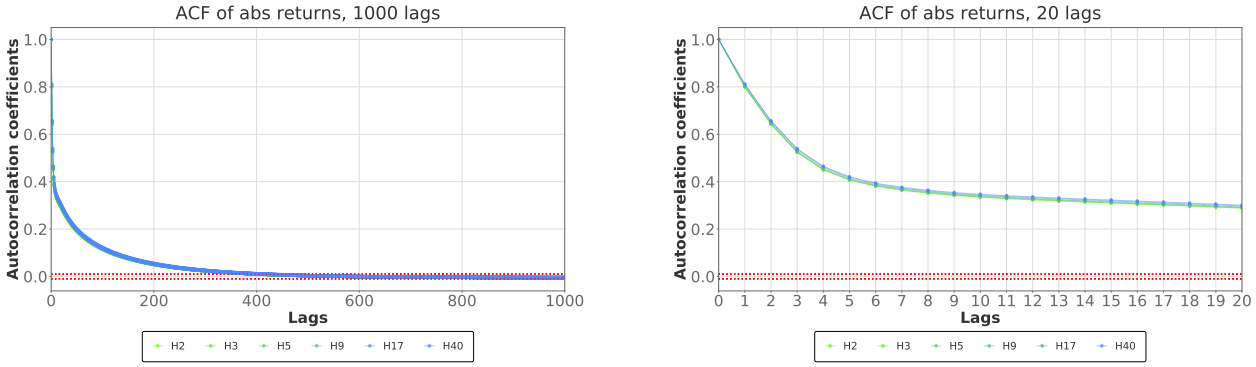


Figure 5.6: ACF coefficients of absolute returns up to lag 1000 (left) with a zoom on the first 20 lags (right), for each market Hi listed in Table 5.1. The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. The coefficients displayed correspond to the average coefficient over 1000 simulation of 494999 returns generated for each Hi. The errors, not clearly visible at this scale, are the standard error of the distribution of the 1000 coefficients for each lag. The red horizontal dashed lines represent the 95% confidence level to accept null hypothesis that the coefficients are i.i.d.. The decay of the first 800 lags are shown in a semilogy scale in Figure 5.7.

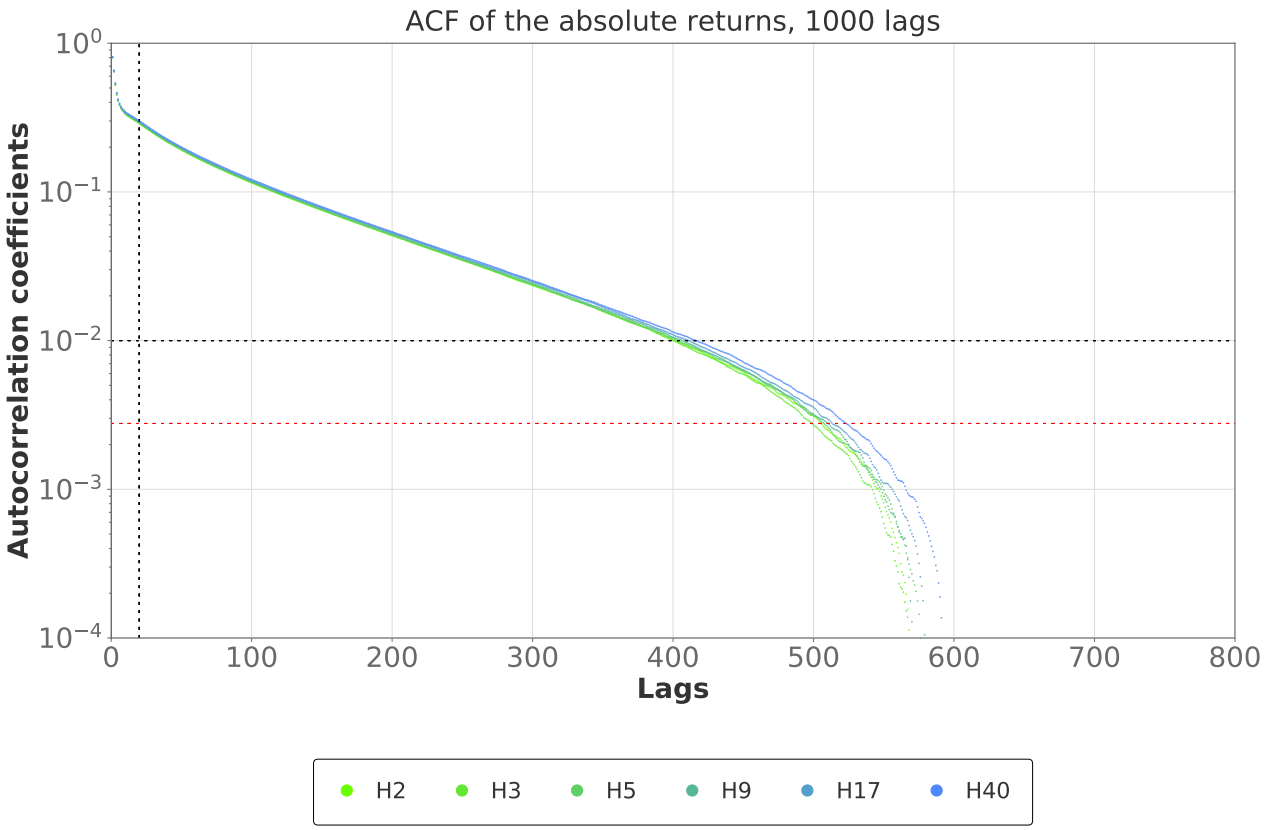


Figure 5.7: ACF coefficients of the absolute returns for all memory parameters up to lag 800 (zoom of Figure 5.6 (left) on a semilogy scale). The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. The red horizontal dashed line represents the upper band of the 95% confidence level to accept the null hypothesis that the coefficients are i.i.d. (see section A.2). The black vertical dashed line is at lag 20, the start lag $s_{\mathrm{H}i}$ used to pinpoint a range of lags yielding a positive combined p-value for the decay exponent β. The black horizontal dashed line corresponds to ACF coefficients with value 0.01. This cut-off is the fix end lag $f_{\mathrm{H}i}$ of the ranges used to pinpoint a positive combined p-value. A slower decay of the ACF of the absolute returns implies stronger volatility clustering.

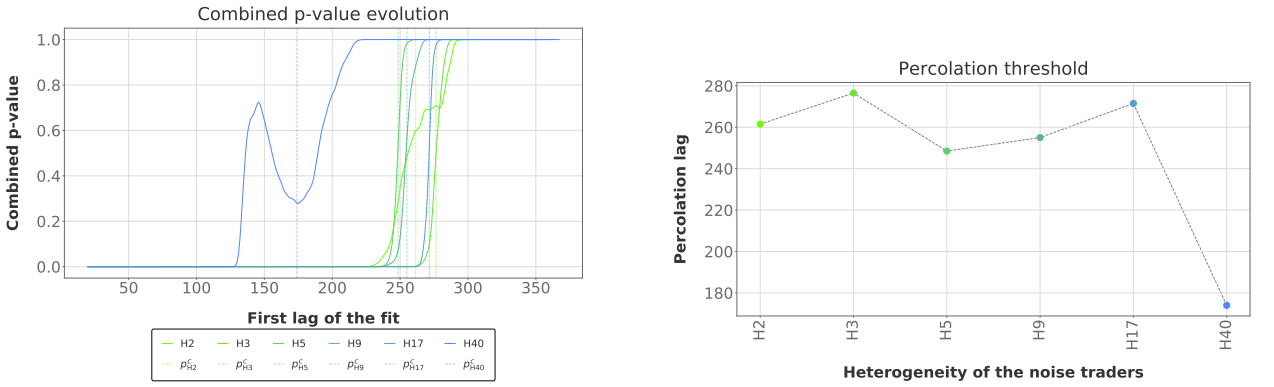


Figure 5.8: The left figure shows the evolution of the combined p-value of the linear fit on the range $[\log(y_{\mathrm{H}i}(x_{s_{\mathrm{H}i}})), \log(y_{\mathrm{H}i}(x_{f_{\mathrm{H}i}}))]$ of the points $(x_{\mathrm{H}i}, \log(y_{\mathrm{H}i}))$ displayed in Figure 5.7, for the starting lags $s_{\mathrm{H}i} \in [20, f_{\mathrm{H}i} - m]$ given on the x-axis of the present plot, the final lag $f_{\mathrm{H}i}$ corresponding to the lag preceding the first lag for which the ACF coefficient is below 0.01 (black dashed line in Figure 5.7) for a given Hi, and minimal amount of lags fitted $m = 50$. The evolution of the combined p-value for each Hi displays an unexpected percolating behavior with a transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds $p^c_{\mathrm{H}i}$ (vertical dashed lines). The value $p^c_{\mathrm{H}i}$ is the mean lag between the highest lag $s_{\mathrm{H}i}$ for which the combined p-value is below 0.01 and the lowest lag $s_{\mathrm{H}i}$ with combined p-value of at least 0.99. The plot on the right displays the values of the percolation thresholds, with the values listed in Table 5.3. The simulations in Figure 5.7 are run for the values listed in Table 2.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing.

horizontal dashed line in Figure 5.7) and $m = 50$ yields at least one range with a positive combined p-value for each market Hi. The evolution of the combined p-value is shown in Figure 5.8 (left). The markets H2 and H40 do not display a percolation behavior as clear as the other 4 markets H3, H5, H9 and H17 and as the evolution observed in Figure 3.6 for H1 markets with memories given in Table 2.1. The percolation threshold $p^c_{\mathrm{H}i}$ (displayed in Figure 5.8 (right), and as dashed lines in Figure 5.8 (left)) is the mean lag between the highest lag $s_{\mathrm{H}i}$ for which the combined p-value is below 0.01 and the lowest lag $s_{\mathrm{H}i}$ with combined p-value of at least 0.99.

Figure 5.9 (left) displays the evolution of the decay exponents β for all ranges analyzed. The numerical values of β are given in Table 5.3. The decay exponents show a downward trend for higher starting lags $s_{\mathrm{H}i}$ which is explained by increasingly shorter ranges of points fitted and an acceleration of the decay for lags approaching the i.i.d. null hypothesis, as seen in Figure 5.7.

To quantify volatility clustering, β is computed for the six markets on the range $[p^c_{\mathrm{H}i} + 1, f_{\mathrm{H}i}]$, the largest range displaying a positive combined p-value, and is shown in Figure 5.9 (right). As mentioned in chapter 3, Cont (2001) cites $[0.2, 0.4]$ for the exponent of the decay of the ACF on absolute return in real markets. The results obtained here show a much slower decay[||], hence more volatility clustering.

[||] As was the case for markets with all noise traders sharing the same memory, see Figure 3.7.

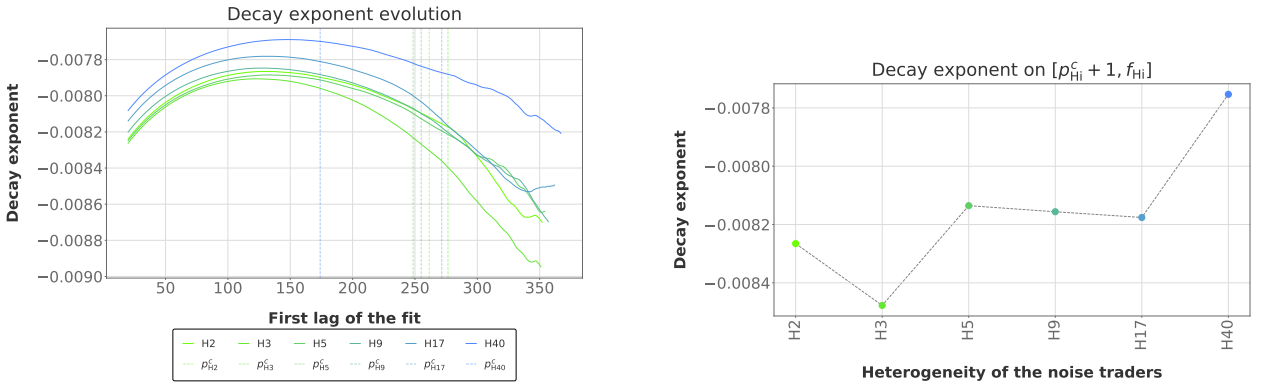


Figure 5.9: The left figure shows the values of the decay exponent β of the linear fit on the range $[\log(y_{\mathrm{H}i}(x_{s_{\mathrm{H}i}})), \log(y_{\mathrm{H}i}(x_{f_{\mathrm{H}i}}))]$ of the points $(x_{\mathrm{H}i}, \log(y_{\mathrm{H}i}))$ displayed in Figure 5.7, for the starting lags $s_{\mathrm{H}i} \in [20, f_{\mathrm{H}i} - m]$ given on the x-axis of the plot, the final lag $f_{\mathrm{H}i}$ corresponding to the lag preceding the first lag for which the ACF coefficient is below 0.01 (black dashed line in Figure 3.5) for a given Hi, and $m = 50$. The evolution of the combined p-value for each Hi displays an unexpected percolating behavior with a transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds $p^c_{\mathrm{H}i}$ (vertical dashed lines). The value $p^c_{\mathrm{H}i}$ is the mean lag between the highest lag $s_{\mathrm{H}i}$ for which the combined p-value is below 0.01 and the lowest lag $s_{\mathrm{H}i}$ with combined p-value of at least 0.99. The plot on the right displays the decay exponent computed on the range $[p^c_{\mathrm{H}i} + 1, f_{\mathrm{H}i}]$. Higher heterogeneities are linked to a slower decay, hence produce more volatility clustering. The values of the percolation thresholds and of the decay exponents are listed in Table 5.3. The simulations in Figure 5.7 are run for the values listed in Table 2.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing.

The results of the analysis of the combined p-value and of the decay β for the markets H1$(\overline{\tau(\mathrm{H}i)})$ are shown in Figure B.7 and Figure B.8, with values given in Table C.4. The percolation is more precisely defined than for the markets Hi. The decay exponent has a cleaner evolution with values of β that do not overlap each other. The number of lags for which the combined p-value is found to be at least 0.99 is large enough for the decay exponents to be representative. The results concord with what was observed in Figure 3.7 for memories above 377 (all $\overline{\tau(\mathrm{H}i)}$) are above 383) with shorter memories linked to a faster decay.

To decouple the effect of the underlying memory of each market Hi, the ratio of the decays of the markets Hi (Figure 5.9 (right), Table 5.3) with the decays of the markets H1$(\overline{\tau(\mathrm{H}i)})$ (Figure B.8 (right), Table C.4) is computed and shown in Figure 5.10**. The values are listed in Table 5.4. Using the decay of the autocorrelation coefficients of the absolute returns as a metric of volatility clustering, the market H2 produces 36% less volatility clustering than the market H1$(\overline{\tau(\mathrm{H2})})$ composed of one fundamentalist and all the noise traders sharing the same memory. The result for a hierarchy of 3 memories is similar (35%). Then, the higher the heterogeneity of the noise traders' memories, the

**Note that both decay exponents being negative, their ratio is positive. A higher ratio translates to a heterogeneous market having a faster decay, thus displaying less volatility clustering.

Hi	**Percolation lag** $p^c_{\mathrm{H}i}$	**End lag** $f_{\mathrm{H}i}$	**Number of lags fitted**	**Decay exponent** β
H2	261	402	141	-0.00812
H3	276	401	125	-0.00839
H5	248	404	156	-0.00810
H9	255	407	152	-0.00810
H17	271	412	141	-0.00813
H40	174	417	243	-0.00770

Table 5.3: The range $[\log(y_{\mathrm{H}i}(x_{s_{\mathrm{H}i}})), \log(y_{\mathrm{H}i}(x_{f_{\mathrm{H}i}}))]$ of the points $(x_{\mathrm{H}i}, \log(y_{\mathrm{H}i}))$ displayed in Figure 5.7 is fitted for varying starting lags $s_{\mathrm{H}i}$. The simulations in Figure 5.7 are run for the values listed in Table 2.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. The evolution of the combined p-value for each Hi displays an unexpected percolating behavior with a transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds $p^c_{\mathrm{H}i}$ (vertical dashed lines in Figure 5.8). The percolation thresholds $p^c_{\mathrm{H}i}$ are the mean lag between the highest starting lag for which the range yields a combined p-value below 0.01 and the lowest starting lag for which the range yields a combined p-value above 0.99 (Figure 5.8). The end lag $f_{\mathrm{H}i}$ represents the lag preceding the first lag for which the ACF coefficients of the absolute returns are below the cut-off 0.01. The number of lags fitted is the difference between the end lag and the percolation lag. The decay exponent β is computed on the ranges $[p^c_{\mathrm{H}i} + 1, f_{\mathrm{H}i}]$.

closer the amount of volatility is to a market with all the noise traders sharing the average memory of the heterogeneous market. Note that as no error is computed on the decay β, no confidence level can be given for these results. A crude approach to evaluate the error on the coefficients β would be to investigate the distributions of the slope for ranges near the one on which β is computed. Note that β is computed on the longest range yielding a combined p-value of at least 0.99 for all the sub ranges[††], which subsets the relevant slopes for the distribution to the ones with a higher starting lag. The evolution of the slope for the markets Hi (Figure 5.9) indicates that some values of β start weaving for starting lags s_τ just above the percolation thresholds. This behavior may weaken the values of the decay presented.

Contrary to what was observed in section 3.2, the percolation lags (Figure 5.8) do not seem to embody information about the volatility clustering. However, the present computation of the decays shows to be more relevant and stable than the case presented in section 3.2. By endowing the noise traders with two or three time scales rather than a single one, the volatility clustering is found to be 35% lower.

[††]The start lag of the range on which β is computed is the first lag yielding a combined p-value of at least 0.99, which usually corresponds to the first lag after the percolation thresholds, if the percolation is not erratic. As it has been repetitively shown, all ranges defined with a higher start lag consistently display a combined p-value of at least 0.99.

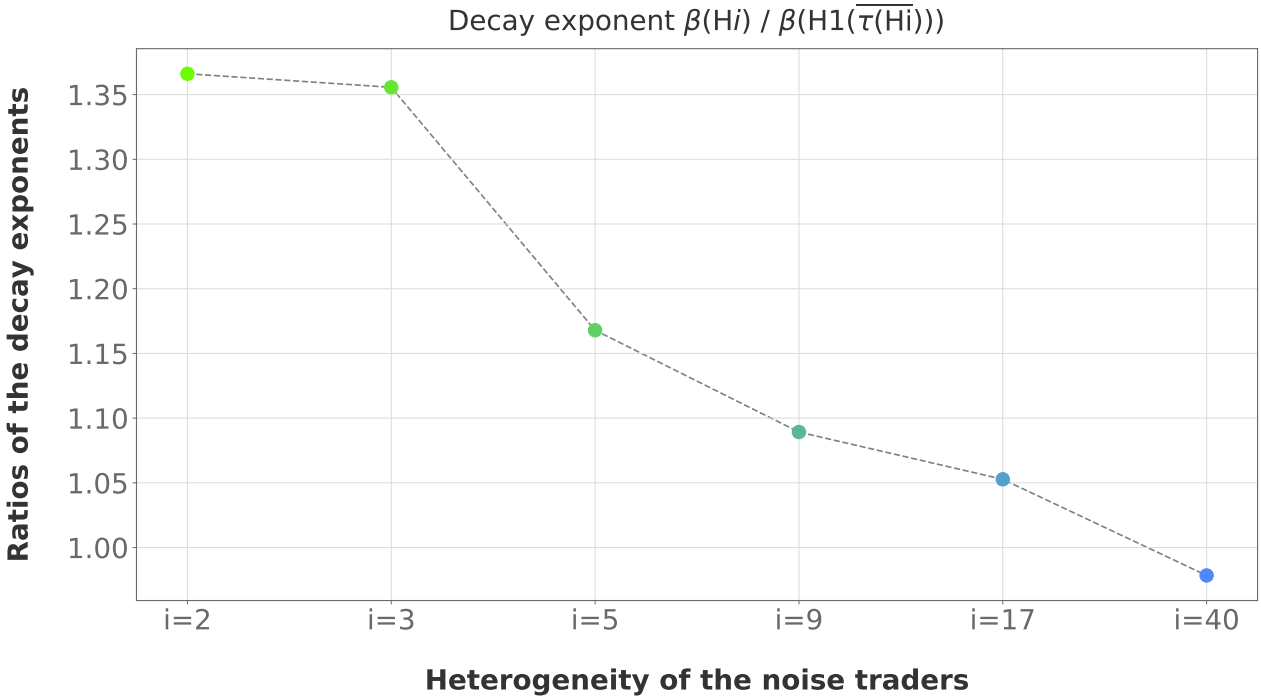


Figure 5.10: Ratio of the decay exponent obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with the results obtained for the markets H1$(\overline{\tau(\text{H}i)})$ with all the noise traders sharing the same memory with their memory set to the average memory $\overline{\tau(\text{H}i)}$ of the heterogeneous market Hi. The values plotted here are listed in Table 5.4. The market H2 produces 36% less volatility clustering than the market H1$(\overline{\tau(\text{H2})})$ composed of one fundamentalist and all the noise traders sharing the same memory $\overline{\tau(\text{H2})}$. The result for a hierarchy of 3 memories is similar (35%). Then, the higher the heterogeneity of the noise traders' memories, the closer the amount of volatility is to a market with all the noise traders sharing the average memory of the heterogeneous market. The details of the computation of the decays and a possible approach to compute error terms are given in the text. For both cases (the 6 markets Hi and the 6 markets H1$(\overline{\tau(\text{H}i)})$), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

	$i=2$	$i=3$	$i=5$	$i=9$	$i=17$	$i=40$
Ratio of the decay exponents $\frac{\beta(\mathrm{H}i)}{\beta(\mathrm{H1}(\overline{\tau(\mathrm{H}i)}))}$	1.36	1.35	1.16	1.08	1.05	0.97

Table 5.4: Ratio of the decay exponent obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with the results obtained for the markets H1$(\overline{\tau(\mathrm{H}i)})$ with all the noise traders sharing the same memory with their memory set to the average memory $\overline{\tau(\mathrm{H}i)}$ of the heterogeneous market Hi. The values listed here are plotted in Figure 5.10. See Figure 5.10 and the text for an analysis of the results.

5.5 Frequency distribution and branching ratio

The metric proposed in section 4.2 to quantify the volatility clustering through the sample standard deviation of the frequency distribution of absolute returns above a threshold v is applied on the returns generated by the six markets Hi presented in Table 5.1.

The evolution of the standard deviation w.r.t. the window size are shown in Figure 5.11 for both thresholds v. The evolution for the different heterogeneities is very similar throughout all window sizes. The values for the window size 10000 are shown in Figure 5.12. The results for the markets H1$(\overline{\tau(\mathrm{H}i)})$ with all noise traders sharing the same memory $\overline{\tau(\mathrm{H}i)}$ are presented in Figure B.9 and Figure B.10. For $v = 1\sigma$, Figure B.10 (left) shows more significant numbers than for the heterogeneous markets (Figure 5.12, left). The $v = 1\sigma$ threshold does not seem the most adequate to differentiate the behaviors of the Hi markets. The difference between both plots of Figure 5.12 shows that all markets Hi have a similar volatility clustering for $v = 1\sigma$, but differ for $v = 2\sigma$ where larger heterogeneities generate more clustering. The H1 markets do have significant volatility clustering for returns above $v = 1\sigma$, even though the biggest clusters are found for $v = 2\sigma$, see Figure B.10, Figure 4.1 and Figure 4.8. The choice of a window of size $m = 10000$ is arbitrary, but relies on the observations that window sizes up to $m = 150$ did not show robustness for all memories in subsection 4.2.2. Selecting the same window sizes for both thresholds allow for a robust comparison.

As in section 5.3 and section 5.4, to decouple the effect of the underlying memory of each market Hi, the ratio of the standard deviations of the markets Hi with the the standard deviations of the markets H1$(\overline{\tau(\mathrm{H}i)})$ is computed and shown in Figure 5.13. The values for both thresholds are given in Table 5.5. A very similar behavior to Figure 5.10[‡‡] is found: introducing a small heterogeneity reduces the volatility clustering in comparison to a market where all noise traders share the heterogeneous market average memory length. Larger heterogeneities bring the amount of volatility clustering of the markets Hi closer to H1$(\overline{\tau(\mathrm{H}i)})$.

In section 4.3, the endogeneity of the market was analyzed through a non parametric approximation of the branching ratio for self-exciting Hawkes point processes in one dimension. (Equation 4.24).

[‡‡]Remember that a higher ratio translates to a heterogeneous market having a faster decay, thus displaying less volatility clustering.

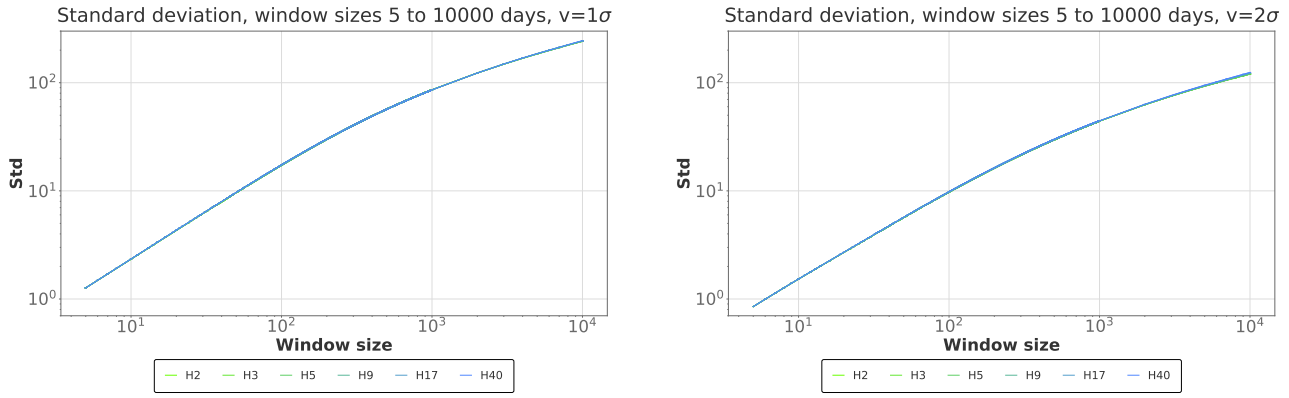


Figure 5.11: Evolution of the standard deviation for the six markets Hi for the threshold $v = 1\sigma$ (left) and $v = 2\sigma$ (right) over increasingly larger window sizes m. The density of lines changing at $m = 1000$ is due to computing the standard deviation (and the error bars) for window sizes that are multiples of 1000 above $m = 1000$. From $m = 5$ to $m = 1000$, every window size is computed. The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing.

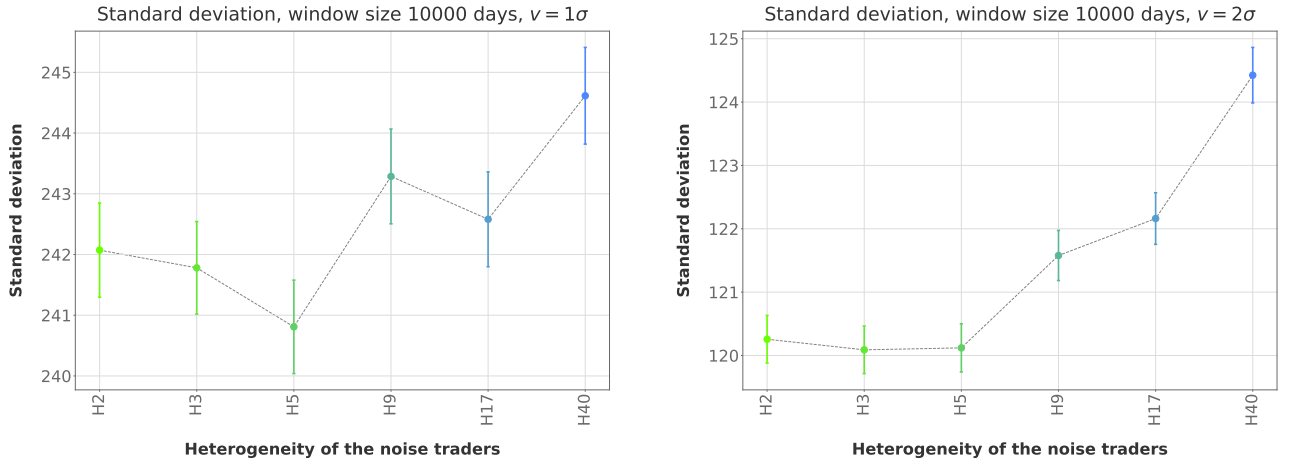


Figure 5.12: Volatility clustering measured through the sample standard deviation of the frequency distribution. All markets Hi have a similar volatility clustering for $v = 1\sigma$, but differ for $v = 2\sigma$ where larger heterogeneities generate more clustering. The H1 markets do have significant volatility clustering for returns above $v = 1\sigma$, even though the biggest clusters are found for $v = 2\sigma$, see Figure B.10, Figure 4.1 and Figure 4.8. The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing.

	$i=2$	$i=3$	$i=5$	$i=9$	$i=17$	$i=40$
$\frac{\text{std}(\text{H}i)}{\text{std}(\text{H1}(\overline{\tau(\text{H}i)}))}$, $v=1\sigma$	0.877 ± 0.003	0.882 ± 0.003	0.895 ± 0.004	0.909 ± 0.004	0.918 ± 0.004	0.923 ± 0.004
$\frac{\text{std}(\text{H}i)}{\text{std}(\text{H1}(\overline{\tau(\text{H}i)}))}$, $v=2\sigma$	0.816 ± 0.003	0.820 ± 0.003	0.829 ± 0.003	0.843 ± 0.003	0.854 ± 0.003	0.870 ± 0.004

Table 5.5: Ratio of the standard deviation of the frequency distribution as defined in Equation 4.8 obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with the results obtained for the markets H1$(\overline{\tau(\text{H}i)})$ with all the noise traders sharing the same memory with their memory set to the average memory $\overline{\tau(\text{H}i)}$ of the heterogeneous market Hi. The values listed are plotted in Figure 5.13. See Figure 5.13 and the text for an analysis of the results. For both cases (the 6 markets Hi and the 6 markets H1$(\overline{\tau(\text{H}i)})$), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

The same analysis is carried out here on the markets Hi. The results are presented in Figure 5.14 and Figure 5.15, with numerical values listed in Table 5.6.

Figure 5.14 shows a smooth evolution of the branching ratio $n(\text{H}i)$ over larger window sizes. As in subsection 4.2.2 and section 4.3, the simulations are run for every $m \in [5, 1000]$, then for every multiple of 1000 for $m \in [1000, 10000]$. The endogeneity grows steadily up to a window size of 1000 days after which the trend flattens and even slightly decreases. Window sizes larger than 10000 were not investigated in details, due to time and computation power constraints. The window size 100000 was computed for $v = 2\sigma$ and is shown in Figure B.13. The values are slightly lower than for a window of size 10000 for all markets Hi but present the same trend. The branching ratio, approximated with Equation 4.24, shows a consistent and converging behavior at very large windows. The necessity for the approximation of the branching ratio to converge only for large window sizes may be tracked back to the necessity to cover all the correlation of the returns.

The results for the markets H1$(\overline{\tau(\text{H}i)})$ with all noise traders sharing the same memory $\overline{\tau(\text{H}i)}$ are presented in Figure B.11 and Figure B.12. In agreement with the results obtained in Figure 4.11, larger memories show higher branching ratio (see Table 5.1 for the values of $\overline{\tau(\text{H}i)}$).

To decouple the effect of the underlying memory of each market Hi, the ratio of the branching ratios of the markets Hi (Figure 5.15) with the the branching ratios of the markets H1$(\overline{\tau(\text{H}i)})$ (Figure B.12) is computed and shown in Figure 5.16. The values are given in Table 5.7. The impact of an increase of heterogeneity in the memories of the noise traders is similar to what was observed for volatility clustering (Figure 5.10, Figure 5.13)[§§]. Using the branching ratios as a metric of market

[§§]The similarity to the results for the volatility clustering computed with the frequency distribution may be attributed to the data collapsed of the mean observed in Figure 4.9.

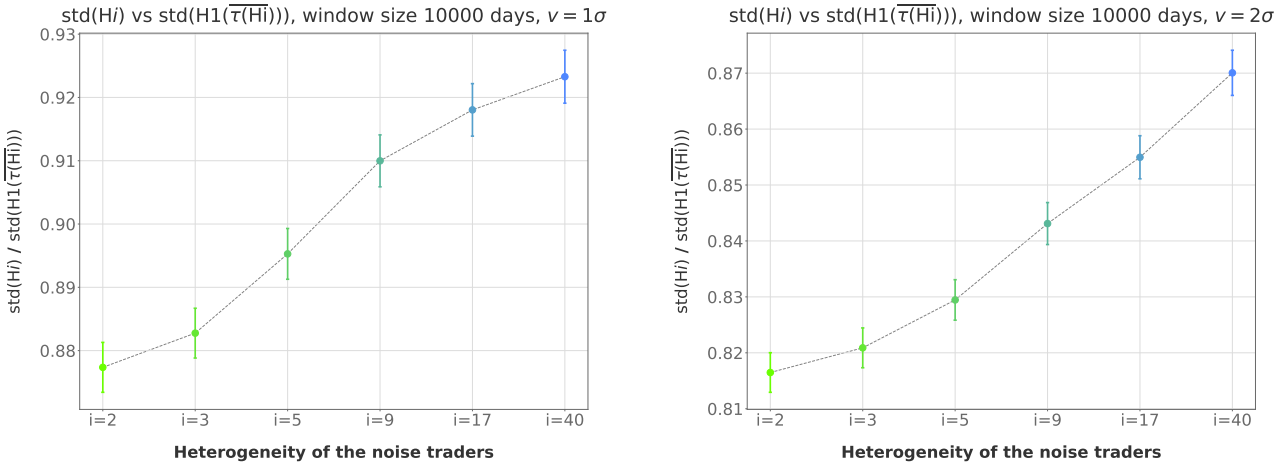


Figure 5.13: Ratio of the standard deviation of the frequency distribution as defined in Equation 4.8 obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with the results obtained for the markets H1($\overline{\tau(\text{H}i)}$) with all the noise traders sharing the same memory with their memory set to the average memory $\overline{\tau(\text{H}i)}$ of the heterogeneous market Hi. Using the standard deviation of the frequency distribution of returns above 2σ as a metric of volatility clustering, the market H2 produces $18.4 \pm 0.3\%$ less volatility clustering than the market H1($\overline{\tau(\text{H2})}$) composed of one fundamentalist and all the noise traders sharing the same memory $\overline{\tau(\text{H2})}$. The result for a hierarchy of 3 memories is similar ($18.0 \pm 0.3\%$ less volatility clustering). Then, the higher the heterogeneity of the noise traders' memories, the closer the amount of volatility is to a market with all the noise traders sharing the average memory of the heterogeneous market. The values plotted are listed in Table 5.5. For both cases (the 6 markets Hi and the 6 markets H1($\overline{\tau(\text{H}i)}$)), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

endogeneity, a market with two groups of noise traders produces $4.54 \pm 0.07\%$ less endogeneity than the market H1($\overline{\tau(\text{H2})}$) composed of one fundamentalist and all the noise traders sharing the same memory $\overline{\tau(\text{H2})}$ (for $v = 2\sigma$). The result for a hierarchy of 3 memories is similar ($4.46 \pm 0.03\%$ less endogeneity). Then, the higher the heterogeneity of the noise traders' memories, the closer the amount of endogeneity is to a market with all the noise traders sharing the average memory of the heterogeneous market.

Quantifying volatility clustering of heterogeneous markets through the standard deviation of the frequency distributions yields a smaller difference w.r.t. the markets H1($\overline{\tau(\text{H}i)}$) than when quantifying the volatility clustering through the decay of the ACF of absolute returns. However, the trend for different Hi is similar. It is also observed that the higher the heterogeneity of memories, the closer the volatility and endogeneity are to markets with a single memory for all the noise traders. This behavior can me summarized by

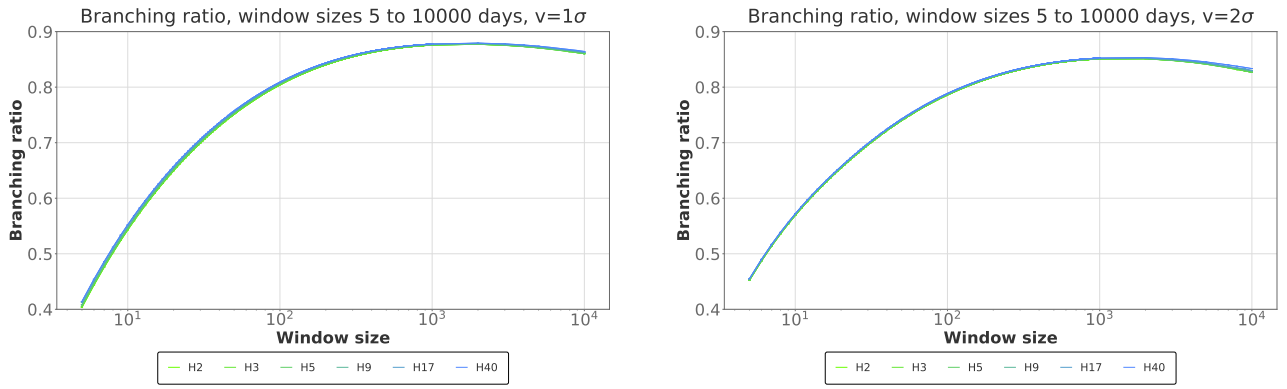


Figure 5.14: Evolution of the branching ratio over very long windows. The Hawkes process is calibrated on the absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). The trend stabilizes from window size 1000. The necessity for the approximation of the branching ratio to converge only for large window sizes may be tracked back to the necessity to cover all the correlation of the returns. A snapshot of the values for window size $m = 10000$ is presented in Figure 5.15. The simulations for each market Hi (formed of i groups of noise traders with i different memories, see Table 5.1) are run for the values listed in Table 2.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

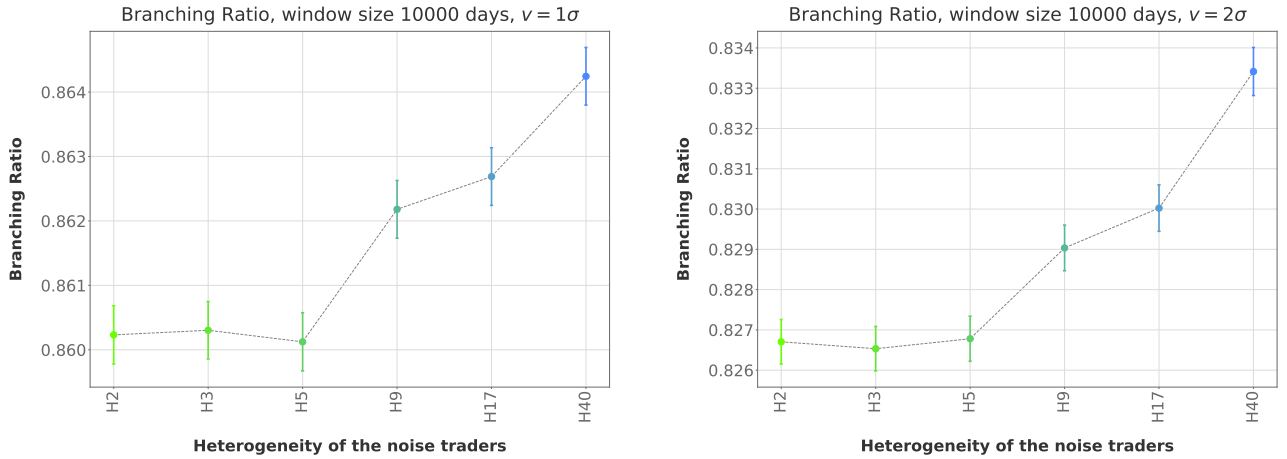


Figure 5.15: Asymptotic values for the branching ratio of the six markets Hi computed through a non parametric approximation of the self-exciting Hawkes point processes (Equation 4.24) in one dimension on a window of size $m = 10000$. The Hawkes process is calibrated on the absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). The values plotted are listed in Table 5.6. The values for $m = 100000$ are shown in Figure B.13. The simulations for each market Hi (formed of i groups of noise traders with i different memories, see Table 5.1) are run for the values listed in Table 2.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

Hi	**Branching ratio** $v = 1\sigma$	**Branching ratio** $v = 2\sigma$
H2	0.8602 ± 0.0005	0.8267 ± 0.0006
H3	0.8603 ± 0.0004	0.8265 ± 0.0006
H5	0.8601 ± 0.0005	0.8268 ± 0.0006
H9	0.8622 ± 0.0004	0.8290 ± 0.0006
H17	0.8627 ± 0.0004	0.8300 ± 0.0006
H40	0.8642 ± 0.0004	0.8334 ± 0.0006

Table 5.6: Branching ratios for each market Hi. The Hawkes process used to approximate the branching ratio is calibrated on the absolute returns above $v = 1\sigma$ and $v = 2\sigma$. The branching ratios listed are computed with a window of size 10000, yielding 49 windows in each of the 1000 simulations. The mean and variance of the frequency distribution (used to compute the approximation of the branching ratio, see Equation 4.24) is computed as described in the Method 1, subsection 4.2.1. The numerical values are shown in Figure 5.15.

$$\frac{\Lambda(\mathrm{H}i)}{\Lambda(\mathrm{H}1(\overline{\tau(\mathrm{H}i)}))} \longrightarrow 1 \qquad \text{for} \quad i \longrightarrow 40 \tag{5.6}$$

where Λ represents the function quantifying the volatility clustering or the market endogeneity. Larger heterogeneities should be considered to probe if this behavior is asymptotic, or holds only locally up to $i = 40$ heterogeneities. On the scales considered, one might argue that a noise traders agent with a single memory has the same impact as a large heterogeneity of memories. Hence, if it can be argued that the speculative traders in the real financial markets are formed by a large amount of different look back windows, their overall behavior can be modeled through a single agent endowed with the average of all the look back windows present. Note that the current ABM has the fundamentalist and noise traders trading at every time step. The model should thus first be extended to allow noise traders with longer memories (or larger wealth) to trade less frequently to mimic strategies with longer holding periods.

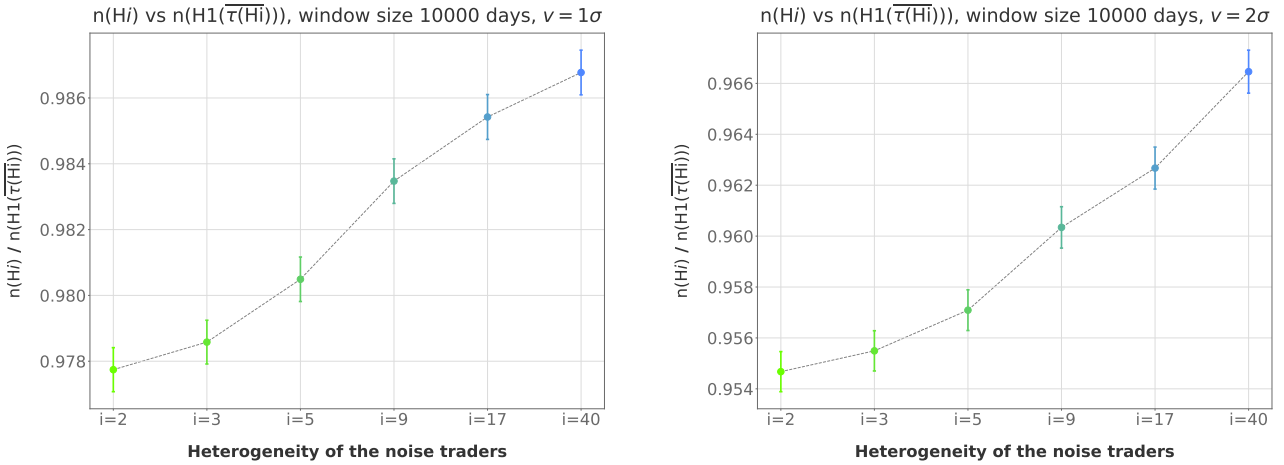


Figure 5.16: Ratio of the branching ratios (as defined in Equation 4.24) obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with the results obtained for the markets H1($\overline{\tau(\text{H}i)}$) with all the noise traders sharing the same memory with their memory set to the average memory $\overline{\tau(\text{H}i)}$ of the heterogeneous market Hi. The values plotted here are listed in Table 5.7. Using the branching ratios as a metric of market endogeneity, a market with two groups of noise traders produces $4.54 \pm 0.07\%$ less endogeneity (for $v = 2\sigma$) than the market H1($\overline{\tau(\text{H2})}$) composed of one fundamentalist and all the noise traders sharing the same memory $\overline{\tau(\text{H2})}$. The result for a hierarchy of 3 memories is similar ($4.46 \pm 0.03\%$ less endogeneity). Then, the higher the heterogeneity of the noise traders' memories, the closer the amount of endogeneity is to a market with all the noise traders sharing the average memory of the heterogeneous market. For both cases (the 6 markets Hi and the 6 markets H1($\overline{\tau(\text{H}i)}$)), The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

	$i = 2$	$i = 3$	$i = 5$	$i = 9$	$i = 17$	$i = 40$
$\frac{\text{n}(\text{H}i)}{\text{n}(\text{H1}(\overline{\tau(\text{H}i)}))}$, $v = 1\sigma$	0.9777 $\pm$ 0.0006	0.9785 $\pm$ 0.0006	0.9804 $\pm$ 0.0006	0.9834 $\pm$ 0.0006	0.9854 $\pm$ 0.0006	0.9867 $\pm$ 0.0006
$\frac{\text{n}(\text{H}i)}{\text{n}(\text{H1}(\overline{\tau(\text{H}i)}))}$, $v = 2\sigma$	0.9546 $\pm$ 0.0007	0.9554 $\pm$ 0.0007	0.9570 $\pm$ 0.0007	0.9603 $\pm$ 0.0007	0.9626 $\pm$ 0.0007	0.9664 $\pm$ 0.0007

Table 5.7: Ratio of the branching ratios (as defined in Equation 4.24) obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with the results obtained for the markets H1($\overline{\tau(\text{H}i)}$) with all the noise traders sharing the same memory with their memory set to the average memory $\overline{\tau(\text{H}i)}$ of the heterogeneous market Hi. The values listed here are plotted in Figure 5.16.

5.6 Initial results on the heterogeneity of the coupling strength

The extension of the ABM presented in this chapter endowed all the noise traders with a common coupling strength κ_t (Equation 1.21, also referred to as strength of the social interaction, social coupling or herding propensity), regardless of their memory length.

Now, a further extension of the ABM is made, where each group of noise traders sharing the same memory is endowed with its own social coupling, thus extending the heterogeneity to a heterogeneity of time scales and social coupling. The social coupling can be understood as how easy each noise traders is influenced by the behavior of the other noise traders from the same group.

It then becomes possible for a group of noise traders to enter a local organized phase, independently of the opinion[¶¶] of the other noise trader groups on the market. All N_G^n groups may also simultaneously be in N_G^n organized phases, with every group being in either a positive or negative feedback mechanism. This state can be best visualized as a Weiss domain[***] in a ferromagnet. With a strong external magnetic field (exogenous shock), the orientations of the N_G^n groups would all align.

The impact on the number of bubbles and their drawdowns are presented in Figure 5.17 with values listed in Table 5.8. When each group, endowed with a view of the market independent from the other groups' views, enters a positive or negative feedback regime, its impact on the price path may be counterbalanced by other groups of noise traders with a different memory. Only when a significant ratio of the groups share the same view can this view move the market significantly. This can be seen with the market H40 only showing a fraction ($0.073 \pm 0.002\%$) of the number of bubbles generated in H1($\overline{\tau(\text{H40})}$). Comparing Table 5.2 with Table 5.8 shows that for a memory heterogeneity of 2, 3 and 5, endowing each memory with its own opinion increases the number of bubbles.

The analysis of the impact of this further extension of heterogeneities on volatility clustering and market endogeneity have be run and will be shown in an updated version of the current work.

[¶¶] Note that opinion does not refer to the opinion index but how easy this group is convinced by others.
[***] Of which the quantum extension is modeled by the Ising model.

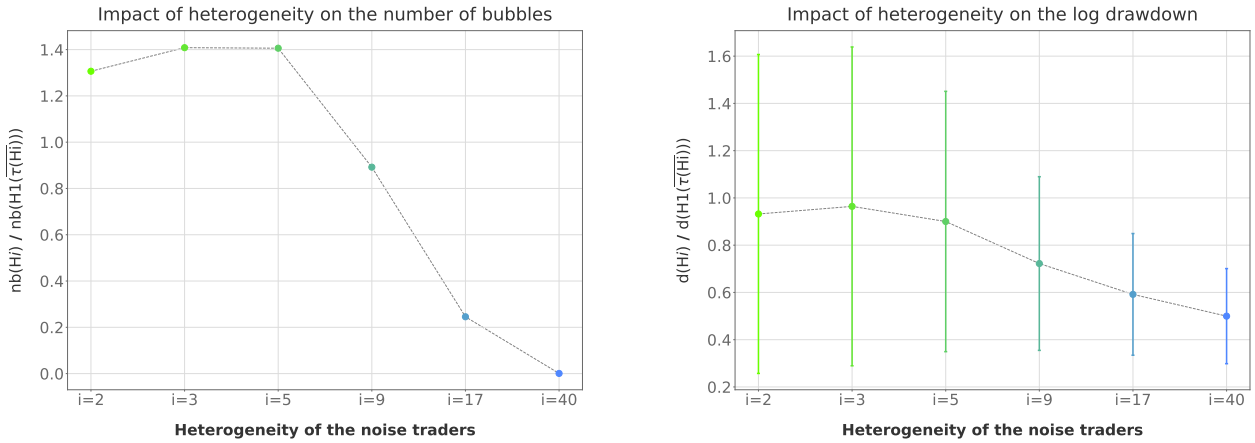


Figure 5.17: Ratio of the number of bubbles (left) and of the log drawdowns (right) obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with each group of noise traders endowed with its own social coupling strength κ (Equation 1.21) with the results obtained for the markets H1$(\overline{\tau(\mathrm{H}i)})$ with all the noise traders sharing the same memory and social coupling strength with their memory set to the average memory $\overline{\tau(\mathrm{H}i)}$ of the heterogeneous market Hi. See the text and Table 5.8 for an analysis of the results. The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

	$i=2$	$i=3$	$i=5$	$i=9$	$i=17$	$i=40$
$\frac{\mathrm{Nb}(\mathrm{H}i)}{\mathrm{Nb}(\mathrm{H1}(\overline{\tau(\mathrm{H}i)}))}$	1.306 ± 0.002	1.408 ± 0.003	1.408 ± 0.003	0.891 ± 0.002	0.245 ± 0.001	0.00073 ± 0.00002
$\frac{\mathrm{d}(\mathrm{H}i)}{\mathrm{d}(\mathrm{H1}(\overline{\tau(\mathrm{H}i)}))}$	0.9 ± 0.6	0.9 ± 0.6	0.9 ± 0.5	0.7 ± 0.3	0.6 ± 0.2	0.5 ± 0.2

Table 5.8: Ratio of the number of bubbles (Nb) and of the log drawdowns (d) obtained for the markets Hi (formed of i groups of noise traders with i different memories, see Table 5.1) with each group of noise traders endowed with its own social coupling strength κ (Equation 1.21) with the results obtained for the markets H1$(\overline{\tau(\mathrm{H}i)})$ with all the noise traders sharing the same memory and social coupling strength with their memory set to the average memory $\overline{\tau(\mathrm{H}i)}$ of the heterogeneous market Hi. The values listed here are plotted in Figure 5.17. Comparing with the results obtained for markets Hi composed of i different groups of noise traders with different memories but a shared coupling strength (Table 5.2) shows that for a memory heterogeneity of 2, 3 and 5, endowing each memory with its own opinion increases the number of bubbles. A market with large numbers of independent opinions generates only a fraction of the bubbles seen in markets with lower heterogeneity of opinions (see the two rightmost columns). The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. For each market, 1000 simulations of 494999 returns are generated.

Conclusion

The aim of this work was to analyze the impact of the memory of the noise traders on the frequency and size of bubbles, on volatility clustering and market endogeneity within the agent-based model (ABM) introduced by Kaizoji et al. (2015). The initial markets H1(τ) are formed of fundamentalist traders establishing their positions based on a value analysis of the market, and of noise traders herding and acting with a common memory τ on price momentum.

A metric to pinpoint the peaks of bubbles out of sample was introduced. The robustness of the parameters of the metric were probed. The metric was then applied on price returns generated by 18 markets H1(τ), each assigning a different memory τ to the noise traders. It was found that markets with the noise traders having no memory produced the least amount of bubbles, but with the highest average drawdown. Assigning a 1 day memory to the noise traders reverses the situation entirely with this memory producing the largest amount of bubbles with the lowest average drawdown. Increasing the memory further reduced the amount of bubbles created and asymmetrically increases their drawdowns. It was shown that a memory of 0 day corresponds to noise traders with an infinite memory.

Quantifying volatility clustering for the 18 different memories through the decay of the autocorrelations coefficients (ACF) of the absolute returns was proved delicate, and did not yield a significant result for 5 of the 18 memories. The short strictly positive memories displayed a faster decay and thus less volatility clustering than longer memories. The memory producing the least amount of volatility clustering was found at $\tau = 21$. For $\tau \geqslant 377$, the longer the memories the more volatility clustering was present. The noise traders with $\tau = 0$ displayed similar results as very long memories. A sharp percolation behavior was found for the combined p-value of the fit of the ACF decay, with longer memories having a percolation threshold at later lag. The memory to percolate the fastest was found for $\tau = 21$. Interestingly, the percolation thresholds showed a similar behavior as the value of the decays.

A more innovative approach to quantify volatility clustering based on the moments of the frequency distribution of the large absolute returns was then thoroughly presented. On the analyzed data, the metric was found robust for all the metric parameters for window sizes longer than 150 returns. This allowed to compare the volatility clustering between memories. The memories producing the least volatility clustering were found for 1, 13 and 21 days. For memories longer than 21 days, the

longer the memory length, the more clustering is generated. Noise traders with no memory showed again the same behavior as very large memories. The results obtained through the moments of the frequency distributions delivered similar results to the memories that were significant with the ACF approach. The mean of the frequency distribution showed data collapse with a critical exponent of 1.00 ± 10^{-12}.

Using the moments of the frequency distribution, the endogeneity of the market was analyzed through a non parametric approximation of the branching ratio for self-exciting Hawkes point processes in one dimension. The branching ratio was showed to be robust for very large windows only. For all memory parameters, the branching ratio was found high but sub critical. The shorter strictly positive memories created the least endogeneity with a branching ratio of $n = 0.8233 \pm 0.0006$ for $\tau = 1$. The longer memories tended to create the most endogenous markets. Consistently with the results shown in the preceding approaches, the noise traders with no memory behaved as embodied by a very long memory. These noise traders created the most endogenous markets with an average branching ratio of $n = 0.8701 \pm 0.0004$.

The ABM was extended to understand how a heterogeneity of time scales of the noise traders impacted the frequency and size of bubbles, the volatility clustering and the endogeneity. The noise traders were split in groups, with all the noise traders in each group sharing the same memory length. Six different markets Hi were generated with respective heterogeneity of $i \in \{2, 3, 5, 9, 17, 40\}$ noise traders time scales. By comparing the results for each market Hi with a market H1($\overline{\tau(\text{H}i)}$) where all the noise traders shared the average memory length $\overline{\tau(\text{H}i)}$ of the different memories of the noise traders in the heterogeneous market, a meaningful analysis of the impact of the heterogeneity on the market was made.

For the market with two groups of noise traders (H2), the number of bubble was found to be $25.2 \pm 0.2\%$ higher than for the market H1($\overline{\tau(\text{H2})}$) (a market with one group of noise traders with the memory set to the average of the memories present in H2). This increase of heterogeneity (doubling the number of memories of the noise traders from 1 to 2) had a higher impact on the number of bubbles than all the other increases of heterogeneity considered. For example, a heterogeneity of 5 memories (H5) generated $27.5 \pm 0.2\%$ more bubbles than the market H1($\overline{\tau(\text{H5})}$). This represents however only $\approx 1\%$ more bubbles than the market H2. Increasing the level of heterogeneity reduced the bubbles average drawdown by 30% for all heterogeneities considered.

An out of sample ACF and PACF analysis showed that a market with a higher heterogeneity of time scales may correspond to an ARMA(p,q) dynamic with lower p, q coefficients than H1 markets. Ftting the decay of the ACF of the absolute returns showed that the market H2 produced 36% less volatility clustering than the market H1($\overline{\tau(\text{H2})}$) composed of one fundamentalist and all the noise traders sharing the same memory $\overline{\tau(\text{H2})}$. The result for a hierarchy of 3 memories was found to be similar (35%). Also, the higher the heterogeneity of the noise traders' memories, the closer the amount of volatility was to a market with all the noise traders sharing the average memory of the heterogeneous market.

Quantifying the volatility clustering of the heterogeneous markets Hi with the moments of the frequency distribution showed that the market H2 produced $18.4 \pm 0.3\%$ less volatility clustering than the market H1($\overline{\tau(\text{H2})}$). The result for a hierarchy of 3 memories was found to be similar

($18.0 \pm 0.3\%$ less volatility clustering). Then, the higher the heterogeneity of the noise traders' memories, the closer the amount of volatility was to a market with all the noise traders sharing the average memory of the heterogeneous market. These results were found to be coherent with the ones found by fitting the decay of the ACF of absolute returns. Both methods showed the same trend, though did not return the same absolute value.

Using the branching ratios as a metric of market endogeneity, a market with two groups of noise traders was found to produce $4.54 \pm 0.07\%$ less endogeneity than the market H1($\overline{\tau(\text{H2})}$). The result for a hierarchy of 3 memories were similar ($4.46 \pm 0.03\%$ less endogeneity). Then, the higher the heterogeneity of the noise traders' memories, the closer the amount of endogeneity was to a market with all the noise traders sharing the average memory of the heterogeneous market.

To sum up, markets where all noise traders shared the same memory saw the most bubbles for a memory of 1 day. This memory also produced the bubbles with the smallest drawdown. The most volatility clustering in markets with a single memory length for all the noise traders was found for the largest memory lengths and no memory at all. A memory length of 1 day produces the least volatility clustering. This memory length also generates the market with the least endogeneity (branching ratio of $n = 0.8233 \pm 0.0006$). As for the amount of volatility clustering, the longest memory and no memory showed to produce the highest values with $n = 0.8701 \pm 0.0004$ when noise traders have no memory. Then we showed that increasing the heterogeneity of time scales from a single memory shared for all the noise traders to 2 memories increased the number of bubbles by $25.2 \pm 0.2\%$ and decreased their drawdown by 30%, by controlling the result w.r.t. to the average memory length of the heterogeneous market. This increase of heterogeneity also decreased the volatility clustering by 36% when computed through the decay of the ACF on absolute returns and $18.4 \pm 0.03\%$ when computed through the moment of the frequency distribution of absolute large returns. The endogeneity of the market was reduced by $4.54 \pm 0.07\%$ by endowing the noise traders with two memory lengths. It is observed that the higher the heterogeneity of memories, the closer the volatility and endogeneity are to markets with a single memory for all the noise traders.

The analysis of the impact of endowing each group of noise traders with its own coupling strength was run and will be shown in an updated version of the current work. In future works, implementing an agent arbitraging the significant autocorrelations found up to lag 100 could show useful to the robustness of the time series approach to quantify volatility clustering. This would bring the current ABM closer to the ACF behaviors observed in real markets. Moreover, different heterogeneities and intrinsic structure of time scales should be systematically analyzed. Also, an analysis of the evolution of the wealth of each group of noise traders may reveal interesting, to observe if some memories dominate and remove other memories from the market. Also, endowing each memory of a heterogeneous market with a corresponding trading frequency may bring the current ABM closer to the trading dynamics observed in real markets. Thanks to the robustness of the code developed for the present work, the impact of any parameter of the ABM (or of future extensions of the ABM) on frequency and size of bubbles, both approaches to quantify volatility clustering, and market endogeneity can be effortlessly run.

Appendix A

Key Concepts in Time Series Analysis

The analysis presented in chapter 3 rely on key concepts of time series analysis. This appendix presents an introduction to the key definitions and equations used in this chapter 3. The shortcomings and pitfalls of the approximation methods used in chapter 3 are highlighted.

This section is partly based on Dettling (2016) and Ruppert (2011).

A.1 Stationarity

A process is strictly stationary if all its moments are constant over time.

Let X_t be a stochastic process and let $F_X(x_{t_1+\tau}, ..., x_{t_n+\tau})$ be the cumulative distribution function of the unconditional joint distribution of X_t at times $t_1 + \tau, ..., t_n + \tau$. Then X_t is strictly stationary if for all $\tau, t_1, ..., t_n \in \mathbb{R}$ and $\forall n \in \mathbb{N}$

$$F_X(x_{t_1+\tau}, ..., x_{t_n+\tau}) = F_X(x_{t_1}, ..., x_{t_n}). \tag{A.1}$$

Hence a stationary process does not show any deterministic pattern, but only stochastic variations.

A weaker form, weak stationarity, is defined by having only the mean, variance and covariance constant in time. In this case we have for all t, s: $\mathbb{E}(X_t) = \mu$ a finite (unconditional) constant, $\text{Var}(X_t) = \sigma^2$ a finite positive (unconditional) constant and $\text{Cov}(X_t, X_s) = \gamma(\mid t - s \mid)$ for the autocovariance function $\gamma(k)$. We note that $\gamma(0) = \text{Cov}(X_t, X_t) = \mathbb{E}[(X_t - \bar{x})(X_t - \bar{x})] = \mathbb{E}[X_t^2] - \bar{x}^2 = \sigma^2$, where $\bar{x} = \frac{1}{n}\sum_{t=1}^{n} X_t$ is the mean.

White Noise (WN) is the simplest stationary process. An independent and identically distributed (i.i.d.) random variable forms the most simple WN, and if the distribution is normal with zero mean it is known as Gaussian WN. A weak WN process X_t is defined by a vanishing covariance for all $t \neq s$: $\text{Cov}(X_t, X_s) = 0$.

A.2 Autocorrelation

The autocorrelation between two random variables X_{t+k}, X_t is defined as

$$\text{Cor}(X_{t+k}, X_t) = \frac{\text{Cov}(X_{t+k}, X_t)}{\sqrt{\text{Var}(X_{t+k})\text{Var}(X_t)}} \,. \tag{A.2}$$

For weakly stationary processes we can write the autocorrelation function ACF

$$\text{Cor}(X_{t+k}, X_t) \coloneqq \rho(k) \,. \tag{A.3}$$

A known and stable technique to estimate the autocorrelations $\hat{\rho}$ is the plug-in approach*:

$$\hat{\rho} = \frac{\hat{\gamma}(k)}{\hat{\gamma}(0)} \tag{A.4}$$

with

$$\hat{\gamma}(k) = \frac{1}{n}\sum_{t=1}^{n-k}(X_{t+k} - \bar{x})(X_t - \bar{x}) \,. \tag{A.5}$$

Note that the estimated autocorrelations† with the plug-in approach are shrunken towards zero for high lags. This is important to keep in mind as we will work with up to 1000 lags.

As is noted in Cont (2007), one has to be vigilant when estimating sample autocorrelation coefficients as it may be possible that these converge to random values (see also Resnick (1998)). For the case of autocorrelations of squared returns, the outcome can even be worse (Mikosch et al. (2000)).

Let X_t be a stationary series. Then, choose n such that $\sum_{k=1}^{n} X_t = 0$. It can be shown that for the stationary series X_t the sum of all the estimated autocorrelation coefficients for all lags k up to $k = n - 1$ is‡

$$\sum_{k=1}^{n-1}\hat{\rho}(k) = -\frac{1}{2} \,. \tag{A.6}$$

This implies the presence of artifacts. Bartlett (1946) showed that the coefficients $\hat{\rho}(k)$ are asymptotically normally distributed. Hence, taking the null hypothesis that a series is i.i.d. (and thus $\hat{\rho}(k) = 0 \;\forall k$) the 95% confidence level to accept the null is the interval $\pm 1.96/\sqrt{n}$, where n is the length of the series This will serve as basis for our confidence bands in ACF and PACF plots.

*The plug-in approach has the advantage of not increasing the variance of $\rho(k)$ as k increases, which is a known problem for the lagged scatter plot method. The brute force algorithm has order n^2. As the time series analyzed have 494999 returns, a Fast Fourier Transform convolution is preferred (order $n\log(n)$, $\approx$ 37750 times faster).

†$\hat{\rho(k)}$ are also called sample autocorrelation coefficients

‡Expand the LHS, and use that $\sum_{k=1}^{n} X_t = 0$ for a stationary series.

A.3 AR processes

A process X_t is an autoregressive (AR) process, when X_t is a weighted average of past instances and a white noise error term (also called innovation). The innovation represents the exogenous shocks. The simplest AR process is the AR(1)

$$X_t - \mu = \varphi(X_{t-1} - \mu) + \varepsilon_t \quad \forall t \ , \tag{A.7}$$

where μ is the mean of the process, φ a constant determining the amount of feedback. The error ε_t is a weak white noise process $\mathrm{WN}(0, \sigma_\varepsilon^2)$ with mean 0 and variance σ_ε^2. The term $\mu + \varphi(X_{t-1} - \mu)$ is the nonconstant conditional expectation of X_t given X_{t-1}, and σ_ε^2 is the constant conditional variance of X_t.

If X_t is weakly stationary, then $\mid \varphi \mid < 1$. This can be seen by taking the variance on both sides of Equation A.7 and noting that they have to be equal : $\sigma_X^2 = \varphi^2\sigma_X^2 + \sigma_\varepsilon^2$. The ACF of an AR(1) model will decay geometrically to zero if $\mid \varphi \mid < 1$ and alternate if $\varphi < 0$. We have an explosive behavior for $\mid \varphi \mid > 1$. For $\varphi = 1$, the process is not stationary and represents the random walk.

We can reformulate Equation A.7 by repeated substitution as follows

$$X_t - \mu = \varepsilon_t + \varphi\varepsilon_{t-1} + \varphi^2\varepsilon_{t-2} + ... = \sum_{h=0}^{\infty} \varphi^h \varepsilon_{t-h} \quad \forall t \ . \tag{A.8}$$

The more general AR(p) model regresses not only the immediate past value, but the p past ones. A stochastic process X_t is an AR(p) process if

$$X_t - \mu = \varphi_1(X_{t-1} - \mu) + \varphi_2(X_{t-2} - \mu) + ... + \varphi_p(X_{t-p} - \mu) + \varepsilon_t \quad \forall t \ , \tag{A.9}$$

with ε_t a weak $\mathrm{WN}(0, \sigma_\varepsilon^2)$.

For an AR(p) process to be weakly stationary, its mean and variance are constant while its covariance only depends on the lag. The unconditional mean is straightforwardly $\mathbb{E}[X_t] = \mu$. We note that the conditional mean may be different from μ:

$$\mu_t = \mathbb{E}[X_t \mid X_{t-1}, ..., X_{t-p}] = \mu + \sum_{i=1}^{p} \varphi_i \mathbb{E}[x_{t-i} - \mu] \ . \tag{A.10}$$

For the variance and covariance, let us take the AR(1) case for simplicity

$$\mathrm{Var}(X_t) = \mathrm{Var}(\mu + \sum_{h=0}^{\infty} \varphi^h \varepsilon_{t-h}) = \sigma_\varepsilon^2 \sum_{h=0}^{\infty} \varphi^{2h} = \frac{\sigma_\varepsilon^2}{1 - \varphi^2} \tag{A.11}$$

$$\mathrm{Cov}(X_t, X_{t-h}) = \mathrm{Cov}\Big(\sum_{i=0}^{\infty} \varepsilon_{t-i}\varphi^i, \sum_{j=0}^{\infty} \varepsilon_{t-h-j}\varphi^j \Big) = \varphi^{|h|} \frac{\sigma_\varepsilon^2}{1 - \varphi^2} \tag{A.12}$$

$$\mathrm{Cor}(X_t, X_{t-h}) \overset{(A.2)}{=} \varphi^{|h|} \ . \tag{A.13}$$

An AR process takes into account past information of up to lag p for the present value. This implies that there is correlation at all lags. In case the ACF shows correlations only at small lags, a moving average model (MA) is more suitable.

A.4 MA processes

We talk about a moving average (MA) process X_t, when X_t is a weighted average (moving average) of past values of the white noise process ε_t. The MA(1) model is

$$X_t - \mu = \varepsilon_t + \theta\varepsilon_{t-1} \quad \forall t \ , \tag{A.14}$$

where ε_t is a weak white noise process $\text{WN}(0, \sigma_\varepsilon^2)$.

The more general MA(q) model is

$$X_t - \mu = \varepsilon_t + \theta_1\varepsilon_{t-1} + ... + \theta_q\varepsilon_{t-q} \quad \forall t \ . \tag{A.15}$$

The autocovariance and autocorrelation functions both vanish for lags higher than q: $\gamma(h) = \rho(h) = 0 \ \forall \mid h \mid > q$.

Similarly to the stationarity of an AR(p) process, we check the MA(q) case as follows: $\mathbb{E}[X_t] = \mu =$ constant, $\text{Var}(X_t) = \sigma_\varepsilon^2(1 + \sum_{i=1}^{q} \theta_i^2) =$ constant. The computation for the covariance can be found in Dettling (2016), pp.86-87.

A.5 ARMA process

An ARMA(p,q) process combines the autoregressive and moving average model described in the previous sections. It allows to be more parsimonious than using only an AR or MA model. This model is defined as

$$X_t - \mu = \varphi_1(X_{t-1} - \mu) + ... + \varphi_p(X_{t-p} - \mu) + \varepsilon_t + \theta_1\varepsilon_{t-1} + ... + \theta_q\varepsilon_{t-q} \quad \forall t \ . \tag{A.16}$$

With the backwards operator B defined as $B^h X_t = X_{t-h}$, the above equation can be rewritten as

$$(1 - \varphi_1 B - ... - \varphi_p B^p)(X_t - \mu) = (1 + \theta_1 B + ... + \theta_q B^q)\varepsilon_t \ . \tag{A.17}$$

Ruppert (2011) derives the following two equations for an ARMA(1, 1) process

$$\rho(1) = \frac{(1 + \varphi_1\theta_1)(\varphi_1 + \theta_1)}{1 + \theta_1^2 + 2\varphi_1\theta_1} \ , \tag{A.18}$$

and for $h \geq 2$

$$\rho(h) = \varphi_1 \rho(h-1) \,, \tag{A.19}$$

which will help us derive a key insight for the GARCH(1, 1) process.

A.6 Partial autocorrelation

The partial autocorrelation function PACF measures the linear dependency between two random variables X_{t+k}, X_t while removing the linear dependency between X_{t+k-1} and X_{t+1}.

$$\pi(k) := \text{Cor}(X_{t+k}, X_t \mid X_{t+k-1}, ..., X_{t+1}) \tag{A.20}$$

A theoretical relation exists between the autocorrelations $\rho(k)$ and the partial autocorrelations $\pi(k)$:

$$\pi(1) = \rho(1) \tag{A.21}$$

$$\pi(2) = (\rho(2) - \rho(1)^2)/(1 - \rho(1)^2) \,. \tag{A.22}$$

An important property of $\pi(k)$ is that the p^{th} coefficient of an AR(p) model is $\varphi_p = \pi(p)$. This means that for an $AR(3)$ model we have $\varphi_3 = \pi(3)$ but not necessarily $\varphi_2 = \pi(2)$ or $\varphi_1 = \pi(1)$. Also, $\pi(k) = 0$ for all $k > p$.

A.7 ACF and PACF

Analyzing the results of the previous section, we can deduct which model is appropriate by looking at the ACF and PACF plots:

- AR(p) when the ACF shows an infinite or exponential decay, and the PACF has a clear cut-off at lag p,
- MA(q) when the ACF shows a clear cut-off at lag q, and the PACF has an infinite or exponential decay,
- ARMA(p,q) when the ACF and PACF both show an infinite/exponential decay or a mix of decay and cut-offs.

A slowly decaying ACF may be due to nonstationarity.

A.8 GARCH

As we have seen, all ARMA models have a nonconstant conditional mean and a constant conditional variance. These models are thus not adequate to model financial time series displaying nonconstant

conditional variance (known as conditional heteroskedasticity). This is the case when volatility clustering is present.

An ARCH model (Autoregressive Conditional Heteroskedasticity) is then useful. Let a_t be an ARCH(1) model

$$a_t = \varepsilon_t \sqrt{\omega + \alpha a_{t-1}^2} \, , \tag{A.23}$$

with ε_t WN(0,1). We must have $\omega > 0$ and $\alpha \geq 0$ for the square root to be positive for all t. For stationarity the condition $\alpha < 1$ must hold. If we rewrite Equation A.23 as

$$a_t^2 = \varepsilon_t^2(\omega + \alpha a_{t-1}^2) \, , \tag{A.24}$$

we see a similarity to an AR(1) process in a_t^2. We have the following unconditional and conditional mean and variance:

$$\mathbb{E}[a_t] = 0 \tag{A.25}$$

$$\mathbb{E}[a_t \mid a_{t-1}, a_{t-2}, ...] = 0 \tag{A.26}$$

$$\mathbb{E}[a_t^2] = \omega + \alpha \mathbb{E}[a_{t-1}^2] \Rightarrow \mathbb{E}[a_t^2] = \frac{\omega}{1-\alpha} \tag{A.27}$$

$$\mathbb{E}[a_t^2 \mid a_{t-1}, a_{t-2}, ...] = \omega + \alpha a_{t-1}^2 = \sigma_t^2 \, . \tag{A.28}$$

As we can see we have the opposite behavior from an AR(1) process, with constant conditional mean and nonconstant conditional variance. ARCH is an example of an uncorrelated (independence of the conditional mean on the past) but dependent (dependence of conditional variance on the past) process. Equation A.28 shows how ARCH processes work: when a_{t-1} is unusually big, a_t will be too, which is a typical behavior in volatility clustering. Even though the a_t process does not have an interesting ACF (uncorrelated), the a_t^2 process does show a geometric decay similar to an AR(1) process, as $\rho_{a^2}(h) = \alpha^{|h|}$ (A.13).

When both the conditional mean and variance are nonconstant, we can use in the simplest case an AR(1)+ARCH(1) model

$$X_t - \mu = \varphi(X_{t-1} - \mu) + a_t \, , \tag{A.29}$$

where a_t is an ARCH(1) process such that $a_t = \varepsilon_t \sqrt{\omega + \alpha a_{t-1}^2}$, with ε_t i.i.d $\mathcal{N}(0,1)$. It is interesting to note that, as a_t is uncorrelated and thus has the ACF of white noise, X_t will display the same ACF as an AR(1): $\rho_{X_t}(h) = \varphi^{|h|} \ \forall h$. Also, a_t^2 has the ARCH(1) ACF: $\rho_{a^2}(h) = \alpha^{|h|} \ \forall h$.

To generalize the processes discussed in this section, we can define the general ARCH(p) process a_t with ε_t Gaussian WN(0,1)

$$a_t = \varepsilon_t \sigma_t = \varepsilon_t \sqrt{\omega + \sum_{i=1}^{p} \alpha_i a_{t-i}^2} \, . \tag{A.30}$$

As for the ARCH(1) case, the ACF of a_t^2 has the same structure as the ACF of a AR(p) process.

To improve parsimony in ARCH models, we can include autoregressive terms of the volatility which allows for persistent volatility to be more accurately modeled. This generalization of ARCH is named GARCH(p,q) and is defined as

$$a_t = \varepsilon_t \sigma_t = \varepsilon_t \sqrt{\omega + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2} \, . \tag{A.31}$$

In (Ruppert, 2011, p. 418-420), the parallel between GARCH and ARMA processes is derived. In summary, if a_t is GARCH, then a_t^2 is ARMA with weak WN. Also, a necessary condition for the stationarity of the process a_t is given by

$$\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1 \, . \tag{A.32}$$

Equation A.18 and Equation A.19 are used to derive the following two key equations of a GARCH(1,1) process,

$$\rho_{a^2}(1) = \frac{\alpha(1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2} \, , \tag{A.33}$$

and for $h \geq 2$

$$\rho_{a^2}(h) = (\alpha\beta)^{h-1} \rho_{a^2}(1) \, . \tag{A.34}$$

Equation A.33 means that $\rho_{a^2}(1)$ is not uniquely defined. Equation A.34 implies that a higher value of $\alpha + \beta$ translates to a slower decay of ρ_{a^2} for lags higher than 1.

Appendix B

Additional Figures

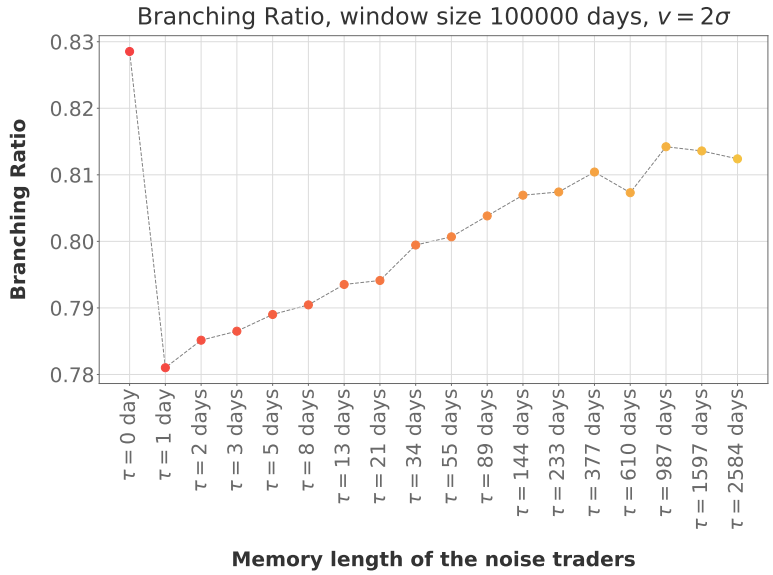


Figure B.1: Asymptotic values for the branching ratio, approximated with Equation 4.24, for a window of size 100000 and calibrated on the absolute returns above $v = 2\sigma$. For all memory parameters, the branching ratio is high but sub critical. The noise traders with no memory behave as embodied by a infinite memory (see Equation 2.4). The shorter strictly positive memories create the least endogeneity. The longer memories tend to create the most endogenous markets, regardless of the select thresholds The values for a window of size 100000 shown here are slightly lower than for 10000 (Figure 4.11) for all memory parameters but present the same trend. The branching ratios show a consistent and converging behavior at very large windows. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length.

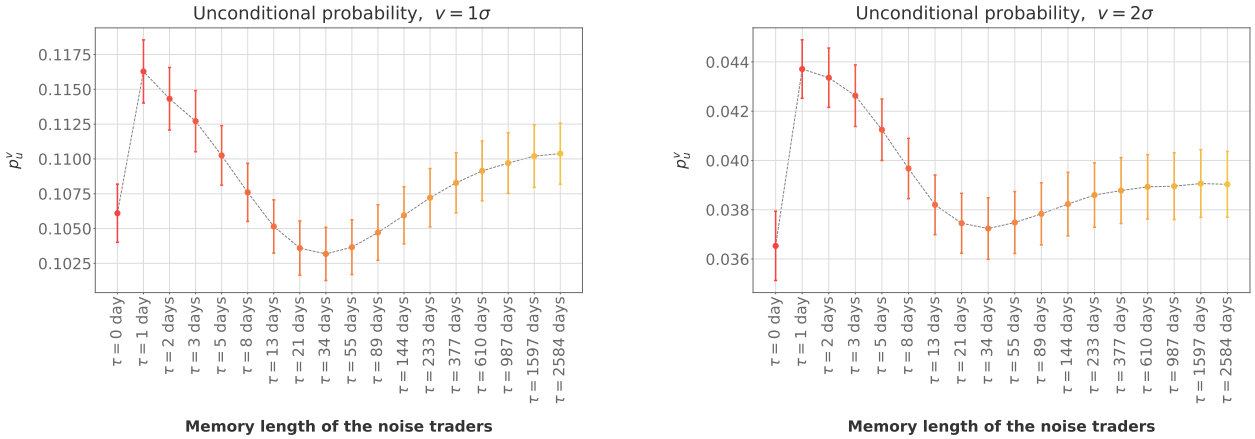


Figure B.2: Unconditional probability to generate a return of absolute value above $v = 1\sigma$ (left) and $v = 2\sigma$(right) for each value of the parameter τ in Table 2.1. The computations are run over 494999 returns and 1000 seeds for each τ and v. The distribution for each τ is plotted in Figure 4.2. The numerical values and their errors (the standard deviation of the corresponding distribution) are found in Table 4.1. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length.

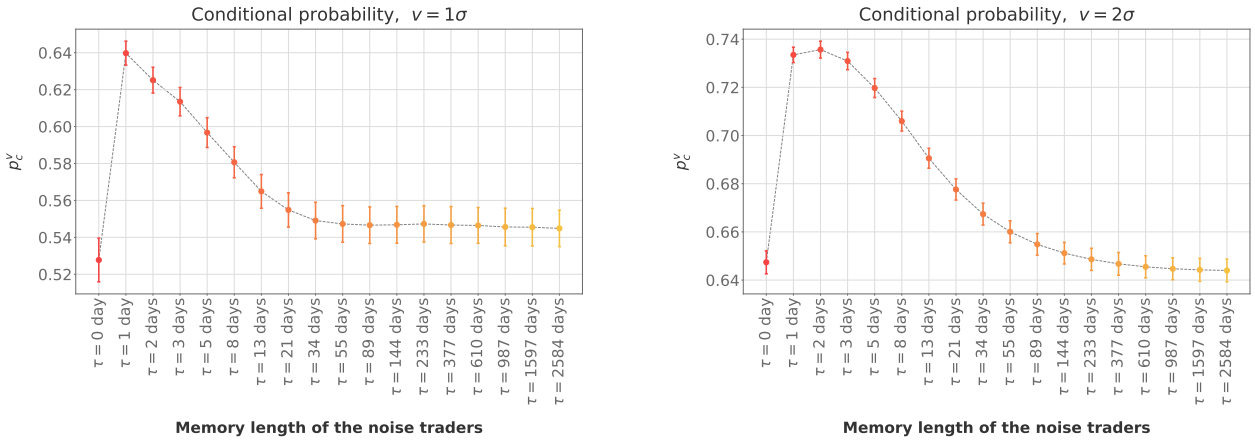


Figure B.3: Conditional probability of to generate a return of absolute value at time t above v when the absolute value of the return at time $t-1$ was also above v, for $v = 1\sigma$ (left) and $v = 2\sigma$(right) for each value of the parameter τ in Table 2.1. The computations are run over 494999 returns and 1000 seeds for each τ and v. The distribution for each τ is plotted in Figure 4.3. The numerical values and their errors (the standard deviation of the corresponding distribution) are found in Table 4.1. The simulations are run varying only the memory parameter τ of the noise traders for the values listed in Table 2.1, while keeping all the other market variables fixed (except for the random seed). In each simulation, all the noise trader agents have the same memory length.

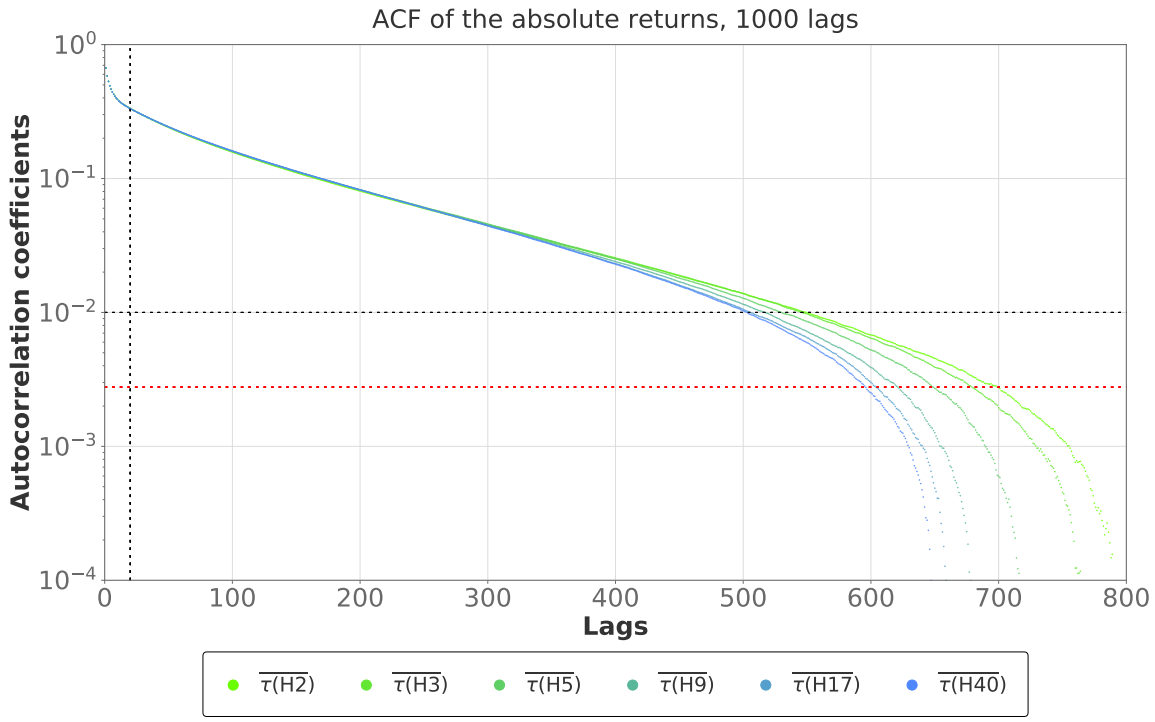


Figure B.4: ACF coefficients of the absolute returns for all memory parameters up to lag 800 for the six markets H1$(\overline{\tau(\mathrm{H}i)})$. The red horizontal dashed line represents the upper band of the 95% confidence level to accept the null hypothesis that the coefficients are i.i.d. (see section A.2). The black vertical dashed line is at lag 20, the start lag $s_{\overline{\tau(\mathrm{H}i)}}$ used to pinpoint a range of lags yielding a positive combined p-value for the decay exponent β. The black horizontal dashed line corresponds to ACF coefficients with value 0.01. This cut-off is the fix end lag $f_{\overline{\tau(\mathrm{H}i)}}$ of the ranges used to pinpoint a positive combined p-value. A slower decay of the ACF of the absolute returns implies stronger volatility clustering. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing.

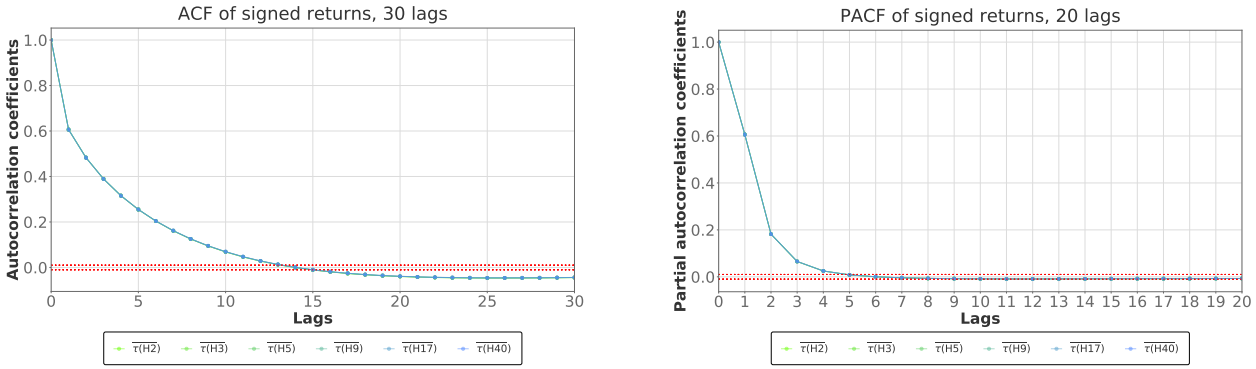


Figure B.5: ACF (left) and PACF (right) coefficients of signed returns, for each market H1($\overline{\tau(\mathrm{H}i)}$). The simulations are run for the values listed in Table 1.1, with the memory parameter and the random seed changing. The coefficients displayed correspond to the average coefficient over 1000 simulation of 494999 returns generated for each H1($\overline{\tau(\mathrm{H}i)}$). The errors, not clearly visible at this scale, are the standard error of the distribution of the 1000 coefficients for each lag. The red horizontal dashed lines represent the 95% confidence level to accept null hypothesis that the coefficients are i.i.d. The first ACF lags are significantly non vanishing, indicating a strong non stationarity in the returns generated. The first lag for which the ACF and PACF coefficient is zero for each H1($\overline{\tau(\mathrm{H}i)}$) and the lag with the lowest coefficients are given in Table C.2.

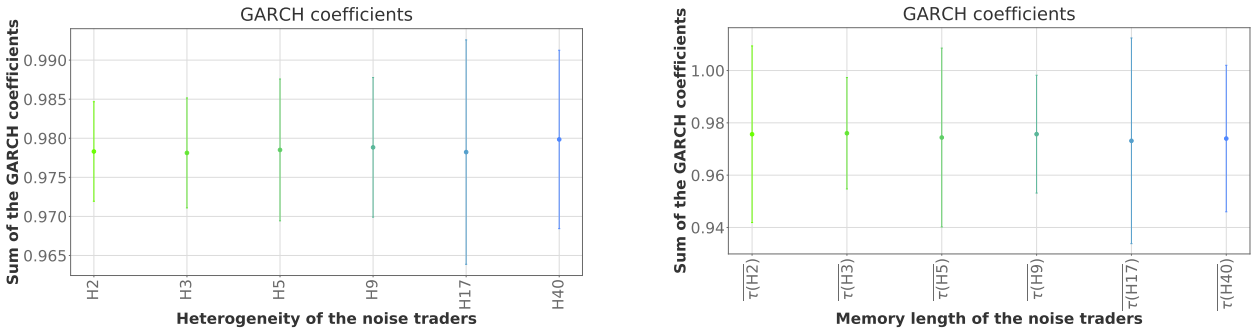


Figure B.6: Average sum of the two GARCH(1,1) coefficients for each heterogeneous market Hi (left) and market H1$(\overline{\tau(\mathrm{H}i)})$ (right). The simulations are run for the values listed in Table 1.1, with the number of noise trader groups (given in Table 5.1), the memory of the noise traders and the random seed changing. It is known from Equation A.34 that the sum of the coefficients of a GARCH(p, q) fit on the signed returns quantifies the decay of the autocorrelation of absolute returns. A GARCH(1,1) model is fitted on every simulation of every market Hi and H1$(\overline{\tau(\mathrm{H}i)})$. The sum of the two GARCH coefficients is averaged over the 1000 simulations. The error is taken as the standard deviation of the distribution of the sums. The results displayed do not allow any clear interpretation. Note that the mean of the sum is below but close to the critical value 1.0 (Equation A.32) for all memory lengths (the confidence bands of some markets H1$(\overline{\tau(\mathrm{H}i)})$ overlap with the critical value), indicating a variance marginal distribution close to infinite. Mikosch et al. (2000) argue at length that in such a case the sample autocorrelations are deceptive estimators for the signed and absolute returns. The ACF of the GARCH(1,1) is an unreliable estimator (tardy convergence), becomes senseless, and one should deflect drawing conclusions from the ACF of the absolute returns (see also Resnick (1998) and Cont (2007)). The non-stationarity of the returns (Figure 5.4) may also play a part (see Equation A.25 and more generally section A.8 where ε_t is always defined as WN(0,1)). The results are disregarded.

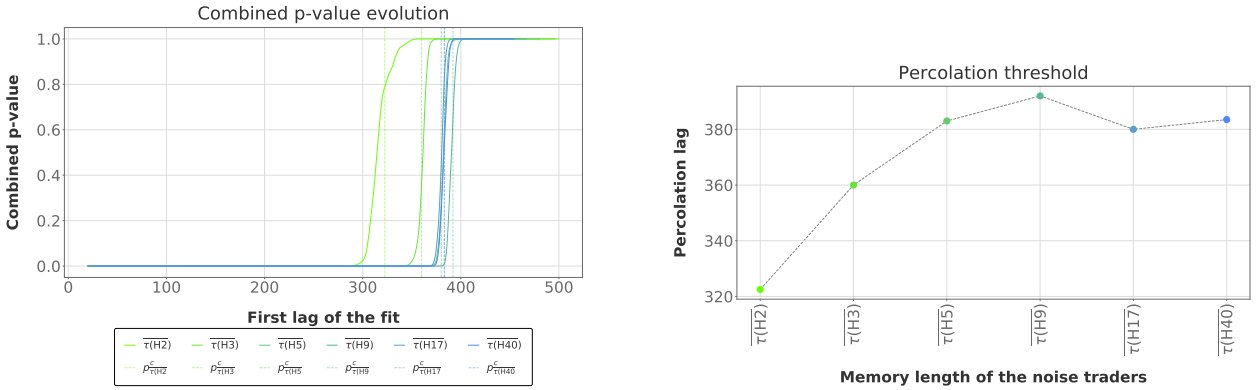


Figure B.7: The left figure shows the evolution of the combined p-value of the linear fit on the range $[\log(y_{\overline{\tau(\mathrm{H}i)}}(x_{s_{\overline{\tau(\mathrm{H}i)}}})), \log(y_{\overline{\tau(\mathrm{H}i)}}(x_{f_{\overline{\tau(\mathrm{H}i)}}}))]$ of the points $(x_{\overline{\tau(\mathrm{H}i)}i}, \log(y_{\overline{\tau(\mathrm{H}i)}}))$ displayed in Figure B.4, for the starting lags $s_{\overline{\tau(\mathrm{H}i)}} \in [20, f_{\overline{\tau(\mathrm{H}i)}} - m]$ given on the x-axis of the present plot, the final lag $f_{\overline{\tau(\mathrm{H}i)}}$ corresponding to the lag preceding the first lag for which the ACF coefficient is below 0.01 (black dashed line in Figure 5.7) for a given $\overline{\tau(\mathrm{H}i)}$, and minimal amount of lags fitted $m = 50$. The evolution of the combined p-value for each $\overline{\tau(\mathrm{H}i)}$ displays an unexpected percolating behavior with an irreversible and sharp transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds $p^c_{\overline{\tau(\mathrm{H}i)}}$ (vertical dashed lines). The value $p^c_{\overline{\tau(\mathrm{H}i)}}$ is the mean lag between the highest lag $s_{\overline{\tau(\mathrm{H}i)}}$ for which the combined p-value is below 0.01 and the lowest lag $s_{\overline{\tau(\mathrm{H}i)}}$ with combined p-value of at least 0.99. The plot on the right displays the values of the percolation thresholds. The simulations are run for the values listed in Table 1.1, with the memory of the noise traders and the random seed changing.

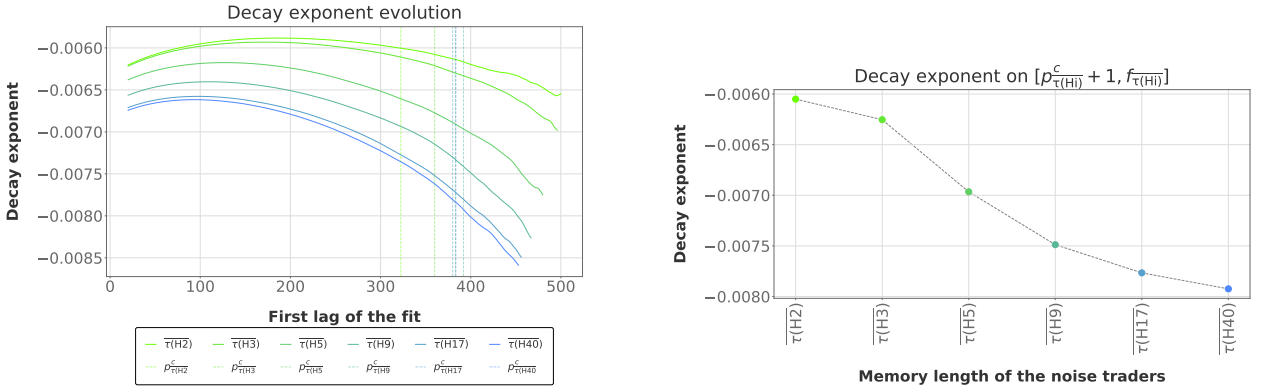


Figure B.8: The left figure shows the values of the decay exponent β of the linear fit on the range$[\log\left(y_{\overline{\tau(\mathrm{Hi})}}(x_{s_{\overline{\tau(\mathrm{Hi})}}})\right), \log\left(y_{\overline{\tau(\mathrm{Hi})}}(x_{f_{\overline{\tau(\mathrm{Hi})}}})\right)]$ of the points $(x_{\overline{\tau(\mathrm{Hi})}i}, \log\left(y_{\overline{\tau(\mathrm{Hi})}}\right))$ displayed in Figure B.4, for the starting lags $s_{\overline{\tau(\mathrm{Hi})}} \in [20, f_{\overline{\tau(\mathrm{Hi})}} - m]$ given on the x-axis of the present plot, the final lag $f_{\overline{\tau(\mathrm{Hi})}}$ corresponding to the lag preceding the first lag for which the ACF coefficient is below 0.01 (black dashed line in Figure 5.7) for a given $\overline{\tau(\mathrm{Hi})}$, and minimal amount of lags fitted $m = 50$. The evolution of the combined p-value for each $\overline{\tau(\mathrm{Hi})}$ displays an unexpected percolating behavior with an irreversible and sharp transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds $p^c_{\overline{\tau(\mathrm{Hi})}}$ (vertical dashed lines). The value $p^c_{\overline{\tau(\mathrm{Hi})}}$ is the mean lag between the highest lag $s_{\overline{\tau(\mathrm{Hi})}}$ for which the combined p-value is below 0.01 and the lowest lag $s_{\overline{\tau(\mathrm{Hi})}}$ with combined p-value of at least 0.99. The plot on the right displays the decay exponent computed on the range $[p^c_{\overline{\tau(\mathrm{Hi})}} + 1, f_{\overline{\tau(\mathrm{Hi})}}]$. The simulations are run for the values listed in Table 1.1, with the memory of the noise traders and the random seed changing.

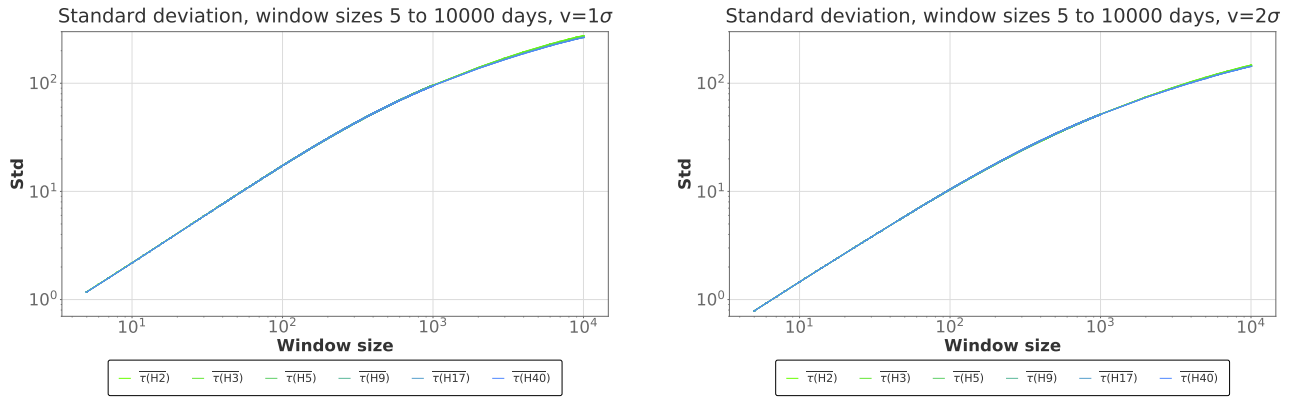


Figure B.9: Evolution of the standard deviation for the six markets H1$(\overline{\tau(\mathrm{Hi})})$ for the threshold $v = 1\sigma$ (left) and $v = 2\sigma$ (right). The density of lines changing at $n = 1000$ is due to computing the standard deviation (and the error bars) for window sizes that are multiples of 1000 above $n = 1000$. From $n = 5$ to $n = 1000$, every window size is computed. The simulations are run for the values listed in Table 1.1, with the memory of the noise traders and the random seed changing.

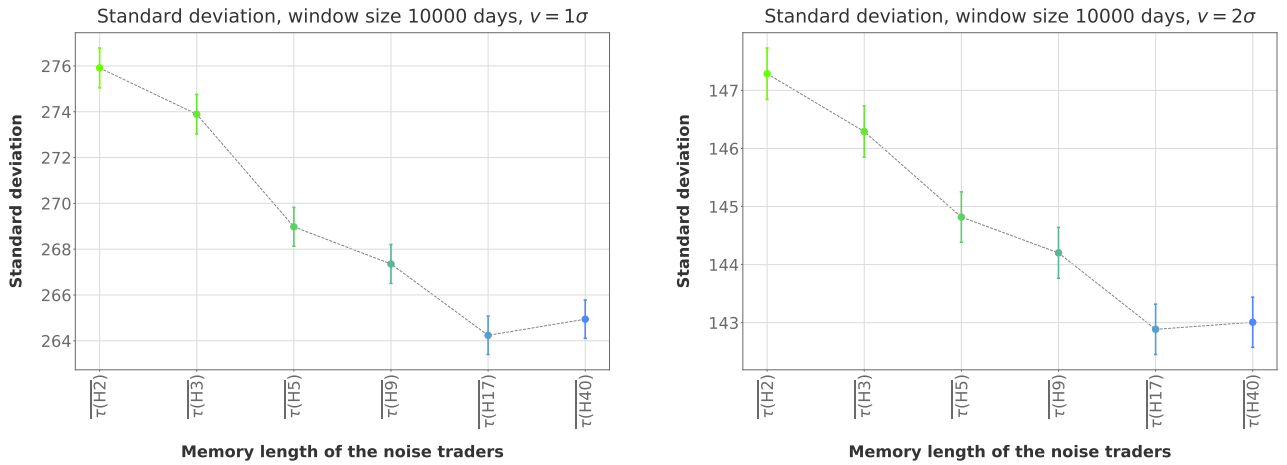


Figure B.10: Volatility clustering measured through the sample standard deviation of the frequency distribution, for the six markets H1$(\overline{\tau(\mathrm{H}i)})$. Larger memories show strong clustering (see Table 5.1 for the values of $\overline{\tau(\mathrm{H}i)}$) generate more clustering, in agreement with the results obtained in Figure 4.8. The simulations are run for the values listed in Table 1.1, with the memory of the noise traders and the random seed changing.

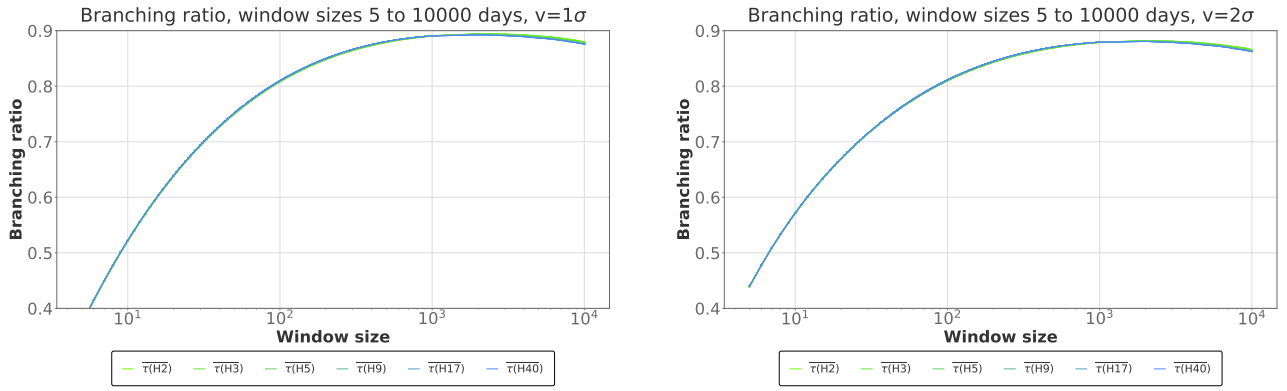


Figure B.11: Evolution of the branching ratio over very long windows for the six markets H1$(\overline{\tau(\mathrm{H}i)})$ calibrated on the absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). The trend stabilizes from window size 1000. A snapshot of the values for window size $m = 10000$ is presented in Figure B.12. The simulations are run for the values listed in Table 1.1, with the memory of the noise traders and the random seed changing.

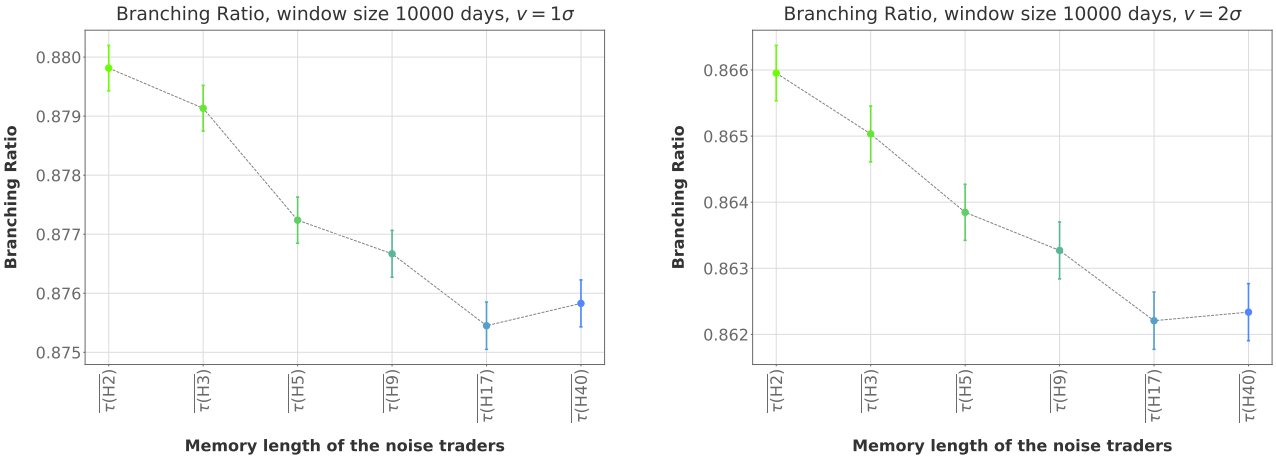


Figure B.12: Asymptotic values for the branching ratio of the six markets H1$(\overline{\tau(\mathrm{H}i)})$ computed through a non parametric approximation of the self-exciting Hawkes point processes (Equation 4.24) in one dimension on a window of size $m = 10000$, calibrated on the absolute returns above $v = 1\sigma$ (left) and $v = 2\sigma$ (right). In agreement with the results obtained in Figure 4.11, larger memories show higher branching ratio (see Table 5.1 for the values of $\overline{\tau(\mathrm{H}i)}$. The simulations are run for the values listed in Table 1.1, with the memory of the noise traders and the random seed changing.

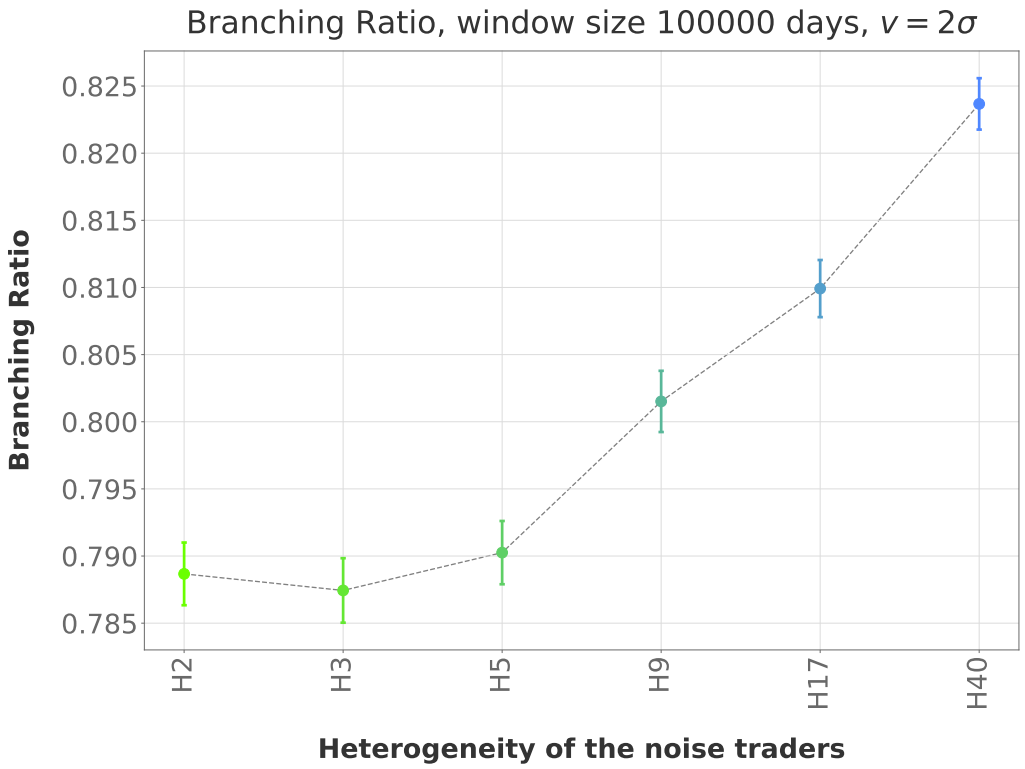


Figure B.13: Branching ratio for the six markets Hi listed in Table 5.1 approximated with Equation 4.24, calibrated on the absolute returns above $v = 2\sigma$. The approximation shows a consistent and converging behavior at very large windows. The values for a window of size 100000 shown here are slightly lower than for 10000 (Figure 5.15) for all memory parameters but present the same trend. The simulations are run for the values listed in Table 1.1, with the memory of the noise traders and the random seed changing.

Appendix C

Additional Tables

	ACF of signed returns			PACF of signed returns		
Heterogeneity	cut-off lag	lowest coeff	lag of lowest coeff	cut-off lag	lowest coeff	lag of min coeff
H2	11	-0.070	20	2	-0.039	3
H3	11	-0.070	20	2	-0.039	3
H5	11	-0.070	20	2	-0.039	3
H9	11	-0.071	20	2	-0.039	3
H17	11	-0.074	20	2	-0.041	3
H40	11	-0.074	20	2	-0.040	3

Table C.1: Behavior of the signed autocorrelation and partial autocorrelations coefficients near the crossing of the i.i.d. null hypothesis (cut-off lag) for markets Hi with a heterogeneity i of memories of the noise traders. All markets Hi display the same cut-off lag and reach their lowest ACF and PACF coefficients at the same lag. Comparing the results with Table C.2 shows that the PACF coefficients have a faster decay in markets Hi than in markets H1$(\overline{\tau(\mathrm{H}i)})$. Even though the PACF coefficients reach lower values in the markets Hi than in markets H1$(\overline{\tau(\mathrm{H}i)})$, these minima occur much earlier (at lag 3 for Hi, between lag 11 and 12 for H1$(\overline{\tau(\mathrm{H}i)})$. This might suggest lower p, q coefficients to model the dynamic of the nonconstant conditional mean and the constant conditional variance in Hi markets. The simulations are run for the values listed in Table 1.1, with the random seed, the number of noise trader groups and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

	ACF of signed returns			PACF of signed returns		
Heterogeneity	cut-off lag	lowest coeff	lag of lowest coeff	cut-off lag	lowest coeff	lag of min coeff
$\overline{\tau(\mathrm{H2})}$	14	-0.047	26	5	-0.009	11
$\overline{\tau(\mathrm{H3})}$	14	-0.047	25	5	-0.009	12
$\overline{\tau(\mathrm{H5})}$	14	-0.046	25	5	-0.009	12
$\overline{\tau(\mathrm{H9})}$	14	-0.046	26	5	-0.009	11
$\overline{\tau(\mathrm{H17})}$	14	-0.045	26	5	-0.009	12
$\overline{\tau(\mathrm{H40})}$	14	-0.046	26	5	-0.009	12

Table C.2: Behavior of the signed autocorrelation and partial autocorrelations coefficients near the crossing of the i.i.d. null hypothesis (cut-off lag) for markets H1($\overline{\tau(\mathrm{H}i)}$) with all the noise traders sharing the same memory $\overline{\tau(\mathrm{H}i)}$ corresponding to the average memory of the market Hi. See Table C.1 for a analysis of the values. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

	ACF of signed returns			PACF of signed returns		
τ	cut-off lag	lowest coeff	lag of lowest coeff	cut-off lag	lowest coeff	lag of min coeff
0	14	-0.045	25	5	-0.009	12
1	11	-0.070	20	2	-0.037	3
2	12	-0.092	20	3	-0.037	4
3	13	-0.101	22	4	-0.026	5
5	14	-0.101	25	4	-0.018	9
8	16	-0.088	29	5	-0.014	11
13	17	-0.068	33	6	-0.010	12
21	17	-0.049	35	6	-0.007	13
34	17	-0.040	32	6	-0.007	12
55	15	-0.038	29	6	-0.007	13
89	15	-0.041	27	6	-0.008	12
144	14	-0.043	26	5	-0.009	12
233	14	-0.045	26	5	-0.009	12
377	14	-0.046	25	5	-0.009	12
610	14	-0.046	26	5	-0.009	12
987	14	-0.046	26	5	-0.009	12
1597	14	-0.047	25	5	-0.009	11
2584	14	-0.047	25	5	-0.009	12

Table C.3: Behavior of the signed autocorrelation and partial autocorrelations coefficients near the crossing of the i.i.d. null hypothesis (cut-off lag) for different value of the memory parameters of the noise traders, for the ABM described in chapter 1. As expected, longer memories are linked to an autoregressive process of higher lags. This can be traced back to the price momentum part of the social interaction of the noise traders (Equation 1.18). A higher τ leads to older price returns taken into consideration. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

τ	**Percolation lag** $p^c_{\overline{\tau(\mathrm{H}i)}}$	**End lag** $f_{\overline{\tau(\mathrm{H}i)}}$	**Number of lags fitted**	**Decay exponent** β
$\overline{\tau(\mathrm{H}2)}$	322	550	228	-0.00600
$\overline{\tau(\mathrm{H}3)}$	360	546	186	-0.00621
$\overline{\tau(\mathrm{H}5)}$	383	530	147	-0.00691
$\overline{\tau(\mathrm{H}9)}$	392	517	125	-0.00741
$\overline{\tau(\mathrm{H}17)}$	380	506	126	-0.00769
$\overline{\tau(\mathrm{H}40)}$	383	503	120	-0.00783

Table C.4: The range $[\log(y_{\overline{\tau(\mathrm{H}i)}}(x_{s_{\overline{\tau(\mathrm{H}i)}}})), \log(y_{\overline{\tau(\mathrm{H}i)}}(x_{f_{\overline{\tau(\mathrm{H}i)}}}))]$ of the points $(x_{\overline{\tau(\mathrm{H}i)}}, \log(y_{\overline{\tau(\mathrm{H}i)}}))$ displayed in Figure B.4 are fitted for varying starting lag $s_{\overline{\tau(\mathrm{H}i)}}$. The evolution of the combined p-value for each $\overline{\tau(\mathrm{H}i)}$ (see Figure B.7) displays an unexpected percolating behavior with an irreversible and sharp transition from a combined p-value of 0 to a combined p-value of 1 at the percolation thresholds $p^c_{\overline{\tau(\mathrm{H}i)}}$ (vertical dashed lines in Figure B.7). The percolation thresholds $p^c_{\overline{\tau(\mathrm{H}i)}}$ are the mean lag between the highest starting lag for which the range yields a combined p-value below 0.01 and the lowest starting lag for which the range yields a combined p-value above 0.99 (see Figure B.7). The end lag $f_{\overline{\tau(\mathrm{H}i)}}$ represents the lag preceding the first lag for which the ACF coefficients of the absolute returns are below the cut-off 0.01. The number of lag fitted is the difference between the end lag and the percolation lag. The decay exponent β is computed on the ranges $[p^c_{\overline{\tau(\mathrm{H}i)}} + 1, f_{\overline{\tau(\mathrm{H}i)}}]$. The simulations are run for the values listed in Table 1.1, with the random seed and the memory parameters changing. In each simulation, all the noise trader agents have the same memory length.

Bibliography

Diego Ardila, Zalan Forro, and Didier Sornette. The acceleration effect and gamma factor in asset pricing. *Swiss Finance Institute Research Paper*, (15-30), 2015.

Emmanuel Bacry, Khalil Dayri, and Jean-François Muzy. Non-parametric kernel estimation for symmetric hawkes processes. application to high frequency financial data. *The European Physical Journal B*, 85(5):157, 2012.

Emmanuel Bacry, Iacopo Mastromatteo, and Jean-François Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(01):1550005, 2015.

Maurice S Bartlett. On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society*, 8(1):27–41, 1946.

Sushil Bikhchandani and Sunil Sharma. Herd behavior in financial markets. *IMF Staff papers*, 47(3): 279–310, 2000.

Tim Bollerslev, Ray Y Chou, and Kenneth F Kroner. Arch modeling in finance: A review of the theory and empirical evidence. *Journal of econometrics*, 52(1-2):5–59, 1992.

H. Peter Boswijk, Cars H. Hommes, and Sebastiano Manzan. Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics and Control*, 31(6):1938 – 1970, 2007. ISSN 0165-1889. doi: https://doi.org/10.1016/j.jedc.2007.01.001. Tenth Workshop on Economic Heterogeneous Interacting Agents.

John Y Campbell and Albert S Kyle. Smart money, noise trading and stock price behaviour. *The Review of Economic Studies*, 60(1):1–34, 1993.

Carl Chiarella, Roberto Dieci, and Xue-Zhong He. Chapter 5 - heterogeneity, market mechanisms, and asset price dynamics. In Thorsten Hens and Klaus Reiner Schenk-Hoppé, editors, *Handbook of Financial Markets: Dynamics and Evolution*, Handbooks in Finance, pages 277 – 344. North-Holland, San Diego, 2009. doi: https://doi.org/10.1016/B978-012374258-2.50009-9.

Eungchum Cho. Variance of sample variance. 2018.

Eungchun Cho, Moon Jung Cho, and John Eltinge. The variance of sample variance from a finite population. *International Journal of Pure and Applied Mathematics*, 21(3):389, 2005.

Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. 2001.

Rama Cont. Volatility clustering in financial markets: empirical facts and agent-based models. In *Long memory in economics*, pages 289–309. Springer, 2007.

Rama Cont, Marc Potters, and Jean-Philippe Bouchaud. Scaling in stock market data: stable laws and beyond. In *Scale invariance and beyond*, pages 75–85. Springer, 1997.

Jean-Remy Conti. Conti long-term behavior of an artificial market, composed of fundamentalists and noise traders. Master's thesis, ETH Zürich, 2018.

J. Bradford de Long, Andrei Shleifer, Lawrence H. Summers, and Robert J. Waldmann. Positive feedback investment strategies and destabilizing rational speculation. *The Journal of Finance*, 45(2): 379–395, 1990. ISSN 00221082, 15406261.

Morris H DeGroot and Mark J Schervish. *Probability and statistics*. Pearson Education, 2012.

Marcel Dettling. Applied time series analysis. *Applied time series analysis*, 2016.

Francis X Diebold and Marc Nerlove. The dynamics of exchange rate volatility: a multivariate latent factor arch model. *Journal of Applied econometrics*, 4(1):1–21, 1989.

Zhuanxin Ding, Clive WJ Granger, and Robert F Engle. A long memory property of stock market returns and a new model. *Journal of empirical finance*, 1(1):83–106, 1993.

Eugene F Fama. Efficient capital markets: Ii. *The journal of finance*, 46(5):1575–1617, 1991.

Vladimir Filimonov and Didier Sornette. Quantifying reflexivity in financial markets: Toward a prediction of flash crashes. *Physical Review E*, 85(5):056108, 2012.

Vladimir Filimonov and Didier Sornette. Apparent criticality and calibration issues in the hawkes self-excited point process model: application to high-frequency financial data. *Quantitative Finance*, 15(8):1293–1314, 2015.

Dominique M Guillaume and Olivier V Pictet. Unveiling nonlinearities through time scale transformations. 1995.

Dominique M Guillaume, Michel M Dacorogna, Rakhal R Davé, Ulrich A Müller, Richard B Olsen, and Olivier V Pictet. From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and stochastics*, 1(2):95–129, 1997.

John Gurland and Ram C Tripathi. A simple approximation for unbiased estimation of the standard deviation. *The American Statistician*, 25(4):30–32, 1971.

Stephen J Hardiman and Jean-Philippe Bouchaud. Branching-ratio approximation for the self-exciting hawkes process. *Physical Review E*, 90(6):062807, 2014.

Stephen J Hardiman, Nicolas Bercot, and Jean-Philippe Bouchaud. Critical reflexivity in financial markets: a hawkes process analysis. *The European Physical Journal B*, 86(10):442, 2013.

Georges Harras and Didier Sornette. How to grow a bubble: A model of myopic adapting agents. *Journal of Economic Behavior & Organization*, 80(1):137–152, 2011.

Georges Harras, Claudio J Tessone, and Didier Sornette. Noise-induced volatility of collective dynamics. *Physical Review E*, 85(1):011150, 2012.

Theodore E Harris. *The theory of branching processes*. Courier Corporation, 2002.

AG Hawkes. Cluster models for earthquakes-regional comparisons. *Bull. Int. Stat. Inst.*, 45(3):454–461, 1973.

Alan G Hawkes. Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, 33(3):438–443, 1971a.

Alan G Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58 (1):83–90, 1971b.

Agnès Helmstetter and Didier Sornette. Subcritical and supercritical regimes in epidemic models of earthquake aftershocks. *Journal of Geophysical Research: Solid Earth*, 107(B10):ESE–10, 2002.

Patrick Hewlett. Clustering of order arrivals, price impact and trade path optimisation. 2006.

Cars H Hommes. Heterogeneous agent models in economics and finance. *Handbook of computational economics*, 2:1109–1186, 2006.

Brian Hurst, Yao Hua Ooi, and Lasse Heje Pedersen. A century of evidence on trend-following investing. *The Journal of Portfolio Management*, 44(1):15–29, 2017.

Zhi-Qiang Jiang, Wei-Xing Zhou, Didier Sornette, Ryan Woodard, Ken Bastiaensen, and Peter Cauwels. Bubble diagnosis and prediction of the 2005–2007 and 2008–2009 chinese stock market bubbles. *Journal of economic behavior & organization*, 74(3):149–162, 2010.

Taisei Kaizoji, Matthias Leiss, Alexander Saichev, and Didier Sornette. Super-exponential endogenous bubbles in an equilibrium model of fundamentalist and chartist traders. *Journal of Economic Behavior & Organization*, 112:289–310, 2015.

Ralf Kohrt. The market impact of exploiting financial bubbles. Master's thesis, ETH Zürich, 2016.

James T Kost and Michael P McDermott. Combining dependent p-values. *Statistics & Probability Letters*, 60(2):183–190, 2002.

Albert S Kyle. Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, pages 1315–1335, 1985.

Blake LeBaron. Agent-based computational finance: Suggested readings and early research. *Journal of Economic Dynamics and Control*, 24(5-7):679–702, 2000.

Blake LeBaron. Time scales, agents, and empirical finance. 2006.

Thomas Lux and Michele Marchesi. Volatility clustering in financial markets: a microsimulation of interacting agents. *International journal of theoretical and applied finance*, 3(04):675–702, 2000.

Burton G Malkiel and Eugene F Fama. Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, 25(2):383–417, 1970.

Thomas Mikosch et al. Limit theory for the sample autocorrelations and extremes of a garch (1, 1) process. *The Annals of Statistics*, 28(5):1427–1451, 2000.

Yoshiko Ogata. The asymptotic behaviour of maximum likelihood estimators for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 30(1):243–261, 1978.

Madis Ollikainen. Multiple market regimes in an equilibrium model of fundamentalist and noise traders. Master's thesis, ETH Zürich, 2016.

Adrian Pagan. The econometrics of financial markets. *Journal of empirical finance*, 3(1):15–102, 1996.

Calyampudi Radhakrishna Rao, Calyampudi Radhakrishna Rao, Mathematischer Statistiker, Calyampudi Radhakrishna Rao, and Calyampudi Radhakrishna Rao. *Linear statistical inference and its applications*, volume 2. Wiley New York, 1973.

Sidney I Resnick. Why non-linearities can ruin the heavy-tailed modeler's day. *A Practical Guide to Heavy Tails: Statistical Techniques and Applications. Birkhäuser, Boston*, pages 219–239, 1998.

David Ruppert. *Statistics and data analysis for financial engineering*, volume 13. Springer, 2011.

Egle Samanidou, Elmar Zschischang, Dietrich Stauffer, and Thomas Lux. Agent-based models of financial markets. *Reports on Progress in Physics*, 70(3):409, 2007.

Didier Sornette. Discrete-scale invariance and complex dimensions. *Physics reports*, 297(5):239–270, 1998.

Didier Sornette. *Critical phenomena in natural sciences: chaos, fractals, selforganization and disorder: concepts and tools*. Springer Science & Business Media, 2006.

Didier Sornette. *Why stock markets crash: critical events in complex financial systems*, volume 49. Princeton University Press, 2017.

Didier Sornette and S Utkin. Limits of declustering methods for disentangling exogenous from endogenous events in time series with foreshocks, main shocks, and aftershocks. *Physical Review E*, 79(6):061110, 2009.

Didier Sornette, Anders Johansen, and Jean-Philippe Bouchaud. Stock market crashes, precursors and replicas. *Journal de Physique I*, 6(1):167–175, 1996.

Meir Statman. The diversification puzzle. *Financial Analysts Journal*, 60(4):44–53, 2004. ISSN 0015198X.

Jie-Jun Tseng and Sai-Ping Li. Quantifying volatility clustering in financial time series. *International Review of Financial Analysis*, 23:11–19, 2012.

Michele Tumminello, Fabrizio Lillo, Jyrki Piilo, and Rosario N Mantegna. Identification of clusters of investors from their real trading activity in a financial market. *New Journal of Physics*, 14(1): 013041, 2012.

Rebecca Westphal and Didier Sornette. Market impact and performance of arbitrageurs of financial bubbles in an agent-based model. *Swiss Finance Institute Research Paper*, (19-29), 2019.

Spencer Wheatley, Alexander Wehrli, and Didier Sornette. The endo–exo problem in high frequency financial price fluctuations and rejecting criticality. *Quantitative Finance*, pages 1–14, 2019.

Jiancang Zhuang, Yosihiko Ogata, and David Vere-Jones. Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association*, 97(458):369–380, 2002.

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Authored by (in block letters):
For papers written by groups the names of all authors are required.

Name(s):	**First name(s):**

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the 'Citation etiquette' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date	**Signature(s)**

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.