Simulation of Stylized Facts in Agent-Based Computational Economic Market Models

Maximilian Beikirch
RWTH Aachen University

Philipp Otte
Forschungszentrum Jülich

Torsten Trimborn
RWTH Aachen University

Simulation of Stylized Facts in Agent-Based Computational Economic Market Models

Maximilian Beikirch*†, Simon Cramer*†, Martin Frank‡, Philipp Otte§¶†, Emma Pabich*, Torsten Trimborn‖**

December 7, 2018

Abstract

We study the qualitative and quantitative appearance of stylized facts in several agent-based computational economic market (ABCEM) models. We perform our simulations with the *SABCEMM* (Simulator for Agent-Based Computational Economic Market Models) tool recently introduced by the authors (Trimborn et al. 2018a). The SABCEMM simulator is implemented in C++ and is well suited for large scale computations. Thanks to its object-oriented software design, the SABCEMM tool enables the creation of new models by plugging together novel and existing agent and market designs as easily as plugging together pieces of a puzzle. We present new ABCEM models created by recombining existing models and study them with respect to stylized facts as well. The code is available on GitHub (Trimborn et al. 2018b), such that all results can be reproduced by the reader.

Keywords: agent-based models, Monte-Carlo simulations, economic market models, stylized facts, SABCEMM, finite size effects, simulator

*RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany

†ORCiD IDs: Maximilian Beikirch: 0000-0001-6055-4089, Simon Cramer: 0000-0002-6342-8157, Philipp Otte: 0000-0002-1586-2274, Emma Pabich: 0000-0002-0514-7402, Torsten Trimborn: 0000-0001-5134-7643

‡Karlsruhe Institute of Technology, Steinbuch Center for Computing, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany

§Forschungszentrum Jülich GmbH, Institute for Advanced Simulation, Jülich Supercomputing Centre, 52425 Jülich, Germany

¶MathCCES, RWTH Aachen University, Schinkelstraße 2, 52056 Aachen, Germany

‖IGPM, RWTH Aachen University, Templergraben 55, 52056 Aachen, Germany

**Corresponding author: trimborn@igpm.rwth-aachen.de

1 Introduction

The empirical observation of stylized facts in financial data dates back more than 100 years. The observation of inequality in income may be seen as the first stylized fact documented by Vilfredo Pareto in 1897 (Pareto 1897). Stylized facts are commonly accepted as persistent empirical patterns in financial data. Furthermore, they are universal in that sense that they can be observed on different markets and even on different time scales all over the world.
The first true empirical observations of stylized facts have been probably done by Fama and Mandelbrot in the 1960s (Mandelbrot 1997; Brada et al. 1966; Eugene 1963). They have shown that the stock return distribution is not well fitted by a Gaussian distribution and obtained the well-known fat-tail characteristic of stock return data. This stylized fact can be quantified by the inverse cumulative distribution function of logarithmic stock returns $F_c(r),\ r \in \mathbb{R}$, where we denote the corresponding random variable $R \in \mathbb{R}$:

$$F_c(r) := \int_r^\infty \Phi(\tilde{r})\ d\tilde{r} \sim \frac{1}{r^\mu}, \quad \mu > 0.$$

Here, Φ denotes the distribution function of the logarithmic stock return distribution and μ the Pareto exponent. The question how to quantify the deviations from Gaussianity is very crucial. One frequently used measure of the fatness of the tail and the peak at the mean of a distribution is the excess kurtosis, given as the normalized fourth moment of stock returns R minus a correction term, defined by

$$\begin{aligned}\kappa &:= \frac{E[(R-\bar{R})^2]}{\sigma^4} - 3,\\ \bar{R} &:= E[R],\\ \sigma^2 &:= E[(R-\bar{R})^2].\end{aligned}$$

The correction term is needed to obtain Gaussian behavior for $\kappa = 0$ called mesokurtic. The stock return distribution exhibits leptokurtic behavior which corresponds to $\kappa > 0$. Thus, the stock return distribution has a higher peak around the mean value and a heavier tail than the Gaussian distribution as figure 1 reveals. The tail exponent of the inverse cumulative distribution function of logarithmic stock returns can be estimated by the Hill estimator (see appendix definition 1). Examples of the excess kurtosis and Hill estimator of real stock price data is given in table 1.

A further possibility in order to visualize the fat-tail property of stock returns is to make use of quantile-quantile plots (qq-plots). The qq-plot in figure 1 plots the data against a Gaussian distributions and the deviations from the straight line clearly indicates the fat-tail behavior.

A further example of a stylized fact in stock prices is volatility clustering, obtained by Mandelbrot as well (Mandelbrot 1997). This stylized fact can be quantified with the help of the auto-correlation function. Since the stock return distribution is assumed to be a stationary stochastic process one can ask for the correlation between stock returns at different points of time. The auto-correlation for the stationary stochastic process $R(t),\ t > 0$ is given by:

$$C(l) := Corr(R(t+l), R(t)) = \frac{Cov(R(t+l), R(t))}{E[(R(t)-\bar{R})^2]} = \frac{E[(R(t+l)-\bar{R})\ (R(t)-\bar{R})]}{E[(R(t)-\bar{R})^2]},\ l > 0$$

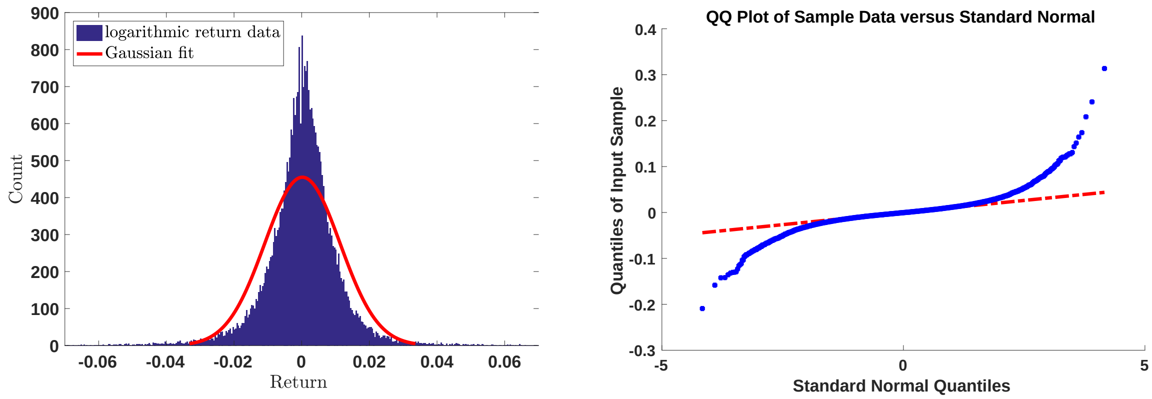


Figure 1: Histogram of daily Dow Jones data with a Gaussian fit (left hand side) and a quantile-quantile plot of the logarithmic return data. Dow Jones data (May 27, 1896-November 14, 2018). Source: `https://stooq.com` (accessed on November 15, 2018).

	DAX	Dow Jones	S&P 500
excess kurtosis	6.41	19.94	26.79
Hill estimator	2.94	2.61	2.86

Table 1: Overview of different daily logarithmic stock return data of different indices. DAX (January 15,1980- November 14, 2018); Dow Jones (May 27, 1896-November 14, 2018); S&P 500(January 15,1980- November 14, 2018); source: `https://stooq.com` (accessed on November 15th 2018).

The correlation $Corr$ is given by the normalized covariance Cov of two random variables. The auto-correlation function $C(l) \in [-1, 1]$ depends on the time shift called lag $l > 0$ of the stochastic process. Empirical data reveals that raw returns are not auto-correlated but in fact absolute or quadratic raw returns possess significant auto-correlation. Interestingly, in the case of absolute or squared returns the auto-correlation function possesses an algebraic decay similar to the stock return distribution.

$$Corr(|R(t+l)|, |R(t)|) \sim \frac{1}{l^\beta},\ \beta > 0.$$

The figure 2 depicts the empirical auto-correlation function of raw and absolute daily returns of DAX data. We clearly obtain no auto-correlation for raw returns and a slowly decreasing auto-correlation with respect to the time lag for absolute returns.

Besides the previously presented stylized facts there are at least thirty stylized facts documented (Chen et al. 2012; Lux 2008). For a detailed discussion on stylized facts we refer to (Cont 2001; Ehrentreich 2007; Campbell et al. 1997; Pagan 1996; Lux 2008).

Although stylized facts are “almost universally accepted among economists and physicists” (Maldarella and Pareschi 2012), the origins of stylized facts remain widely undiscovered (Pa-

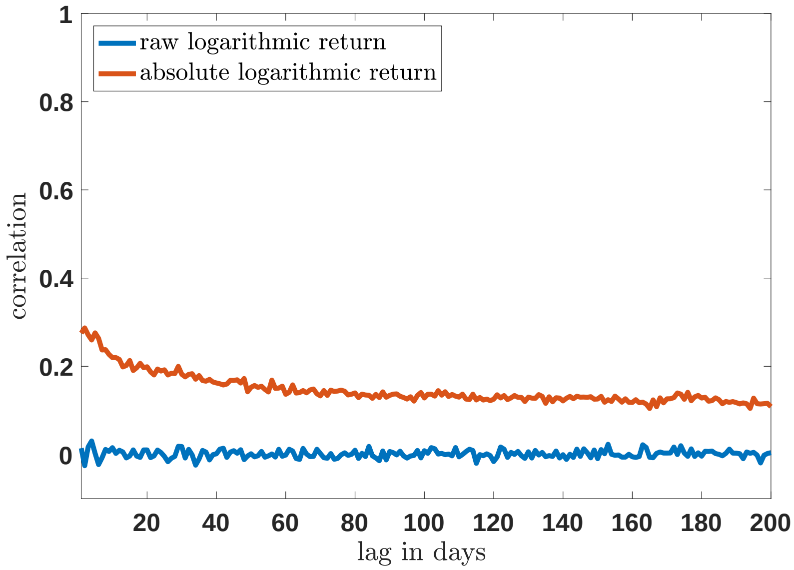


Figure 2: Auto-correlation of raw and absolute returns of daily Dow Jones data (May 27, 1896-November 14, 2018). Source: `https://stooq.com` (accessed on November 15th 2018).

gan 1996; Cowan and Jonard 2002; Maldarella and Pareschi 2012). For example many stylized facts cannot be explained by the famous Efficient Market Hypothesis by Eugene Fama (Fama 1965) which has been the dominant paradigm in macroeconomics for many years. Furthermore, several studies indicate that stylized facts play a crucial role in the creation of financial crashes (Farmer and Foley 2009; LeBaron 2006).

One possible approach to gaining insights into the creation of stylized facts are computational agent-based models. This approach which borrows tools from statistical mechanics, such as Monte-Carlo simulations, has become very popular over the last decade (Farmer and Foley 2009; Hommes 2006; Tesfatsion 2002). These modern financial market models of interacting heterogeneous agents share many similarities with interacting particle systems from physics (Sornette 2014; Zschischang and Lux 2001; Lux et al. 2008). These models usually consider bounded rational agents in the sense of Simon (Simon 1955) and are influenced by behavioral finance (LeBaron 2006; Farmer and Foley 2009; Hommes 2006; Chen et al. 2012). Furthermore, these asset pricing models are often inspired by physical theories or models such as kinetic theory or the famous Ising model. Many agent-based financial market models are able to replicate the most prominent stylized facts of financial markets such as fat-tails of asset returns or volatility clustering. Thus, these computational agent-based models are able to shed light on the origins of stylized facts. As many studies indicate behavioral aspects of financial agents may be one reason for the creation of stylized facts (Cross et al. 2005; Lux 2008; Chen et al. 2012). There has been even developed new theories such as the interacting agent hypothesis (Lux and Marchesi 1999; Ehrentreich 2007) or the heterogeneous market hypothesis proposed by Hommes (Hommes 2001; Ehrentreich 2007) as alternatives to the efficient market hypothesis. For a comprehensive introduction to agent-based models and stylized facts we refer to (Ehrentreich 2007; Chen et al. 2012; Janssen and Ostrom 2006; Cont

2007; LeBaron 2000, 2006).

The drawback of this computational approach is that all results are only based on numerical experiments. Thus, the simulations often depend on a large amount of pseudo random numbers and it is of major importance to ensure to have a good pseudo random number generator. In addition, earlier studies (Egenter et al. 1999; Zschischang and Lux 2001; Challet and Marslii 2002; Kohl 1997; Hellthaler 1996) have shown that obtained stylized facts in several models are only numerical artifacts. More precisely, these studies revealed that for example the very influential Lux-Marchesi model and the Levy-Levy-Solomon model exhibit finite size effects (Egenter et al. 1999; Zschischang and Lux 2001). Finite size effects generally describe that different numbers of agents may lead to qualitative different model outputs. For that reason it is of paramount importance to simulate ABCEM models with a large number of agents. Furthermore, there is no objective comparison possible between different models, since the models are implemented in different languages and simulated on different machines. In addition, we experienced difficulties while reproducing the results published in literature. This may have several reasons. First of all ABCEM models are usually non-linear dynamical systems, very sensitive to their parameters. Secondly, ABCEM models heavily depend on pseudo random numbers and thus it is impossible to reproduce published results exactly. Finally, many publications provide incomplete information regarding the implementation details, e.g. initial values of model quantities.

For these reasons, we have recently introduced the open source SABCEMM simulator (Trimborn et al. 2018a) available on GitHub (Trimborn et al. 2018b). The SABCEMM simulator is built on a novel unified ABCEM model introduced in (Trimborn et al. 2018a) as well. The key idea is the definition of building blocks, for example market mechanisms or agent designs, to be able to recombine building blocks of different models as easily as plugging together pieces of a puzzle. This enables the user to create new ABCEM models or implement existing ABCEM models from literature with a minimum amount of additional coding. This is achieved by implementing the building blocks within the SABCEMM C++ source code and combining the building blocks via an easy to use XML-based configuration file. Due to the implementation of SABCEMM using C++, SABCEMM enables the user to run models with several million agents on a standard notebook or workstation. This lends SABCEMM particularly well for analysis of statistics of and sensitivity analyses for ABCEM models free of finite size effects. We point out that the SABCEMM tool supports numerous high quality pseudo random number generators. This enables the user to carry out fair comparisons between different models.

There have been four ABCEM models from literature implemented in the SABCEMM simulator, namely the Levy-Levy-Solomon (LLS) model (Levy et al. 1994), the Cross model (Cross et al. 2005), the Harras model (Harras and Sornette 2011) and the Franke-Westerhoff model (Franke and Westerhoff 2012). The goal of the present paper is to illustrate the extensive simulation possibilities of the SABCEMM simulator. Thus, we first aim to reproduce known results in literature. Secondly, we run additional test on existing models regarding the long time behavior or finite size effects. Finally, we utilize the object-oriented design using building blocks to create new ABCEM models and study them with respect to stylized facts.

The outline of the paper is as follows: In section 2 we give a short introduction to the

SABCEMM simulator. Then we present the simulation results of each model with respect to the reproducibility of the most prominent stylized facts, namely fat-tails, absence of autocorrelation and volatility clustering. Especially, we create new models and test them with respect to stylized facts. Finally, we compare the artificial stock price data of several models with respect to stylized facts. We finish this paper with a short conclusions of this work.

2 The SABCEMM Simulator

The recently introduced open source simulator SABCEMM (Trimborn et al. 2018a), available on GitHub (Trimborn et al. 2018b), is especially designed for the large-scale simulation of ABCEM models. The simulator is implemented in C++ and comes with an object oriented design. The distinct features of SABCEMM in comparison to other simulation tools such as the JAMEL or JASA simulators (Abar et al. 2017) are:

1. **Generality**: It is built on the basis of a generalized ABCEM model, suitable for implementation of a very wide spectrum of ABCEM models.

2. **Recombination**: It allows recombination of the building blocks of different ABCEM models via configuration files for evaluation of novel ABCEM models.

3. **Comparability**: It provides a common foundation, including random number generators, for fair comparisons of different ABCEM models.

4. **Extensibility**: Implementation of additional ABCEM models is facilitated by object-oriented design.

5. **Efficiency**: Suitability for simulations with a large number ($> 10^6$) of agents, allowing for testing for finite size effects.

In the following, we present the main conceptional ideas behind the simulator and explain the advantages of this software tool.

SABCEMM is well suited for any economic market model which consists of at least one *agent* and one *market mechanism*. As an agent we understand an investor who has a supply of or demand for a certain good or asset, which is traded at the market. Thus, a market mechanism has to determine the price by the demand and supply of all market participants. More precisely, we differentiate between the so called *price adjustment process* and the *excess demand calculator*. The latter one aggregates the supply and demand of all market participants (agents) to one quantity, the excess demand. The former one is rather the method of how the market price is fixed based on this excess demand. A schematic picture, which illustrates the presented ideas is shown in figure 3. The concept of an *environment* has been first introduced by Otte et al. (Trimborn et al. 2018a) and needs to be explained in detail. An environment is some additional coupling between the agents. The most famous example is possibly herding, which is frequently used in ABCEM models. We emphasize that such an environment is not mandatory. For a rigorous mathematical definition of the meta-model which is the skeletal structure of the SABCEMM simulator we refer to (Trimborn et al. 2018a). In addition, a detailed discussion of technical details and computational aspects of SABCEMM can be found in (Trimborn et al. 2018a) as well. Finally, we present some examples of the previously mentioned building blocks such as **agent**, **excess demand**, **price adjustment process** and **herding**. We want to begin with an example of a financial agent.

Example 1. *Frequently used financial agents in ABCEM models are fundamentalist and chartists (Lux 1998; Hommes 2006). A fundamental agent believes that to every stock or good there is a fair market price (or fundamental value) $P^F > 0$ and that the market price will converge to this fundamental value. Hence, a fundamentalist wants to sell stocks if the*

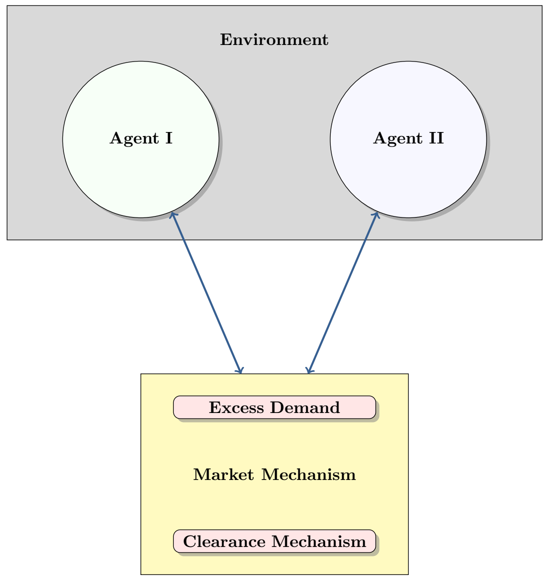


Figure 3: Schematic picture of the abstract ABCEMM model (Trimborn et al. 2018a).

stock price is above the fundamental value and he buys stocks if the market price is below the fundamental value. Hence the agent's demand can be calculated as follows

$$ed^F(t_k) := a\ (P^F(t_k) - P(t_k)),$$

for a positive constant $a > 0$ *and a logarithmic stock price* $P(t_k) \in \mathbb{R}$ *with the discretized time* $t_k := t_0 + k\ \Delta t,\ \Delta t > 0\ k \in \mathbb{N}$*. Here,* $t_0 > 0$ *denotes the initial time. A chartist assumes that the future stock return is best approximated by extrapolating past returns. One simple example of such an investor is*

$$ed^C(t_k) := b\ (P(t_k) - P(t_{k-1})),$$

for a positive constant $b > 0$*. Examples of such agents can be found in the models (Chiarella et al. 2006; Beja and Goldman 1980; Hommes 2001, 2006; Lux 1995; Franke and Westerhoff 2009).*

Another building bock is the aggregated demand ED of all agents. The aggregated excess demand ED of an ABCEM model of N agents is defined as the mean value of agents' excess demands $ed_i \in \mathbb{R}$.

$$ED(t_k) := \frac{1}{N} \sum_{i=1}^{N} ed_i(t_k).$$

Example 2. *A simple example is the Franke-Westerhoff model (Franke and Westerhoff 2012) where the population of agents is fully described by two agents ($N = 2$), a chartist and a fundamentalist. Hence, the excess demand is defined as*

$$ED(t_k) = \frac{1}{2}\left(ed^C(t_k) + ed^F(t_k)\right).$$

Furthermore, we present an example of a price adjustment mechanism. More precisely we give an example of a irrational market as discussed in (Trimborn et al. 2018a), which can be interpreted as a time discretized version of an ordinary differential equation with time step $\Delta t = 1$.

Example 3. *The evolution of the logarithmic stock price in the Harras model is given by:*

$$P(t_{k+1}) = P(t_k) + \frac{1}{\lambda}\ ED^H(t_k).$$

Here, the constant $\lambda >$ denotes the market depth and ED^H the excess demand of the Harras model. For a definition of the excess demand of the Harras model the reader is referred to (Harras and Sornette 2011).

Finally, we present the herding mechanism in the Cross model which can be regarded as an example of an environment.

Example 4. *In the Cross model (Cross et al. 2005) each agent is described by a herding pressure $c_i \geq 0$ and the individual position $\sigma_i \in \{0, 1\}$ (long or short in the market). The evolution of the herding pressure reads*

$$\begin{cases} c_i(t_{k+1}) = c_i(t_k) + \Delta t |ED^C(t_k)|, & \text{if } \sigma_i(t_k)\ ED^C(t_k) \leq 0 \\ c_i(t_{k+1}) = c_i(t_k), & \text{otherwise.} \end{cases}$$

Thus, the herding pressure is increased if the investment decision of the agent σ_i has the opposite sign as the aggregated excess demand ED^C, which is in this model defined as the average of the agents' positions. This corresponds to the situation that the agent's position being in the minority. The agent switches its position if the herding pressure c_i has reached a threshold $\alpha_i > 0$. After a switch the herding pressure is reseted to zero. This herding mechanism is an additional coupling among the agents, complementary to the coupling by the market mechanism. For a full presentation of the model we refer to appendix A.2.

We like to close this paragraph with the remark that any building block of the implemented models can be recombined in a reasonable manner thanks to the object orientation of SABCEMM. In section 4, we provide an example of such new models.

3 Validation and Novel Tests for Known ABCEM Models

In this section we present simulations of the Cross, LLS, Harras and Franke-Westerhoff model. We discuss each model separately and demonstrate the advantages of our simulation framework. Thus, we study the behavior of the LLS and Cross model in the case of many agents. This way, we investigate if the models may exhibit finite-size effects. Furthermore, we study several model variants of the Franke-Westerhoff model with respect to common stylized facts. We emphasize that we provide the reader with all necessary information to reproduce the results. Thus, we define the implemented models in detail and present all parameter values and initial states for each model in the appendix A.2. We ran our simulations on an Intel Xeon 64 bit architecture. The input files for all simulations can be found on the GitHub repository of the SABCEMM source code (Trimborn et al. 2018b). Furthermore, our simulation data is available as a data publication (Trimborn et al.; Beikirch et al.).

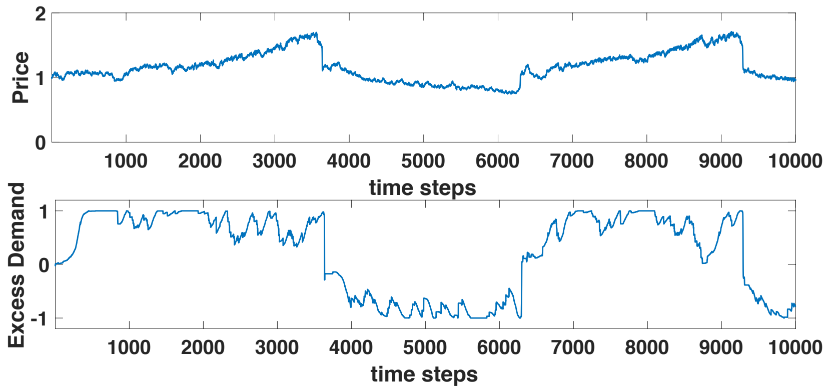
We give a short introduction to each model and discuss the appearance of stylized facts for each model separately. For a detailed definition of the implemented models and parameter settings we refer to appendix A.2.

3.0.1 Cross Model

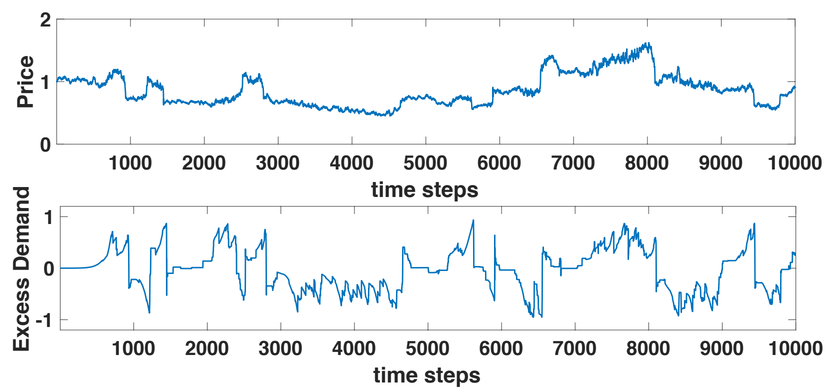
This section presents results for the Cross model which is inspired by the Ising model (Ising 1925) from physics. In the Cross model, each agent is characterized by his position, long $\sigma_i = +1$ or short $\sigma_i = -1$ in the market, respectively. Their investment propensity is determined by two tensions: one related to rational agent behavior and the other to irrational agent behavior. They both mimic the role of temperature in the Ising model. The irrational agent behavior takes into account the herding propensity of financial investors. The price process is driven through the change of the excess demand and is additionally perturbed by white noise. The authors Cross et al. show that their model can replicate the most prominent stylized facts of financial markets, namely fat-tails, uncorrelated price returns and volatility clustering. For further modeling details, we refer to (Cross et al. 2005, 2007).

In our simulations, we obtain the same qualitative results as presented in (Cross et al. 2005, 2007). As figure 4a reveals, the price dynamic is influenced heavily by the evolution of the excess demand over time. Furthermore, the absence of auto-correlation in raw price returns can be verified by figure 6a. In addition, figure 5a reveals the fat-tail in asset returns. By adding the *heteroskedasticity parameter* θ, one couples the noise with the excess demand. This leads to volatility clustering as we can see in figure 6a. A detailed introduction of the model can be found in appendix A.2.

Finite Size Effects In (Cross et al. 2007) the authors claim that their model has no finite size effects. All their simulations are performed with only 100 agents. In order to verify their statement, we analyze the model with different numbers of up to five million agents. We ran our simulations with $10,000$ time steps in order to have a sufficiently large sample size. Our simulations support the findings of Cross et al. (see figs. 4b, 5b and 6b). Thus, we obtain no qualitative difference between the model outputs conducted with 100 or five million agents.



(a) Parameters as in table 4.



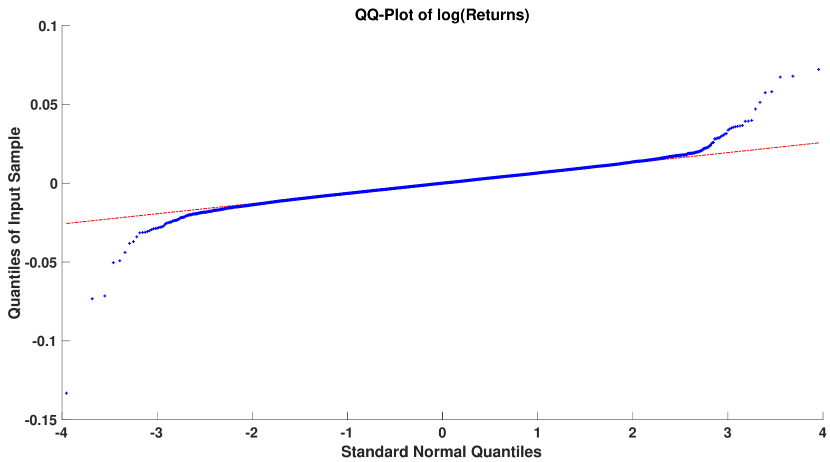
(b) Parameters as in table 4, except $N = 5,000,000$ and $\theta = 2$.

Figure 4: Development of price and excess demand for the Cross base model.

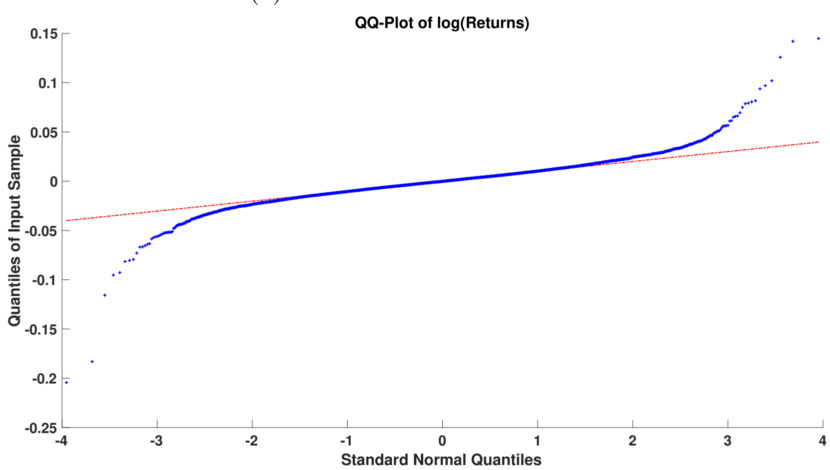
3.0.2 LLS Model

In this section, we present results for the LLS model which is one of the earliest and most influential econophysical ABCEM models. In addition, the LLS model is an example of a *rational market*, i.e. the excess demand is zero at each time step (compare eq. (2)). For further details we refer to (Trimborn et al. 2018a). The LLS model is subject to critical discussions in literature (Zschischang and Lux 2001). We discuss these crucial findings in detail in this section.

The LLS model considers the wealth evolution of the financial agents. Every agent has



(a) Parameters as in table 4.



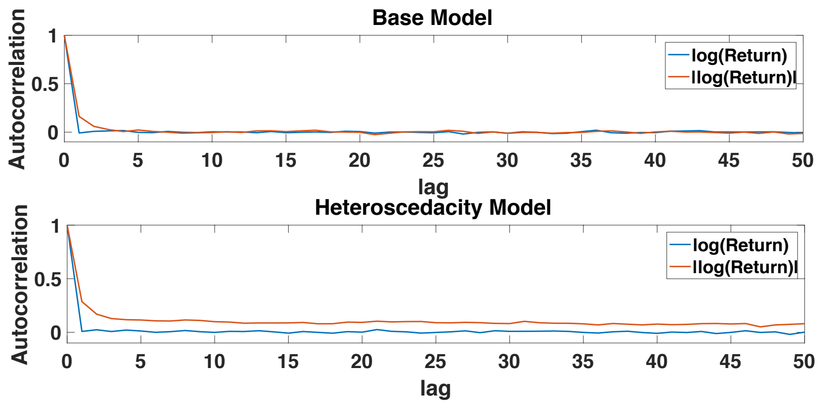
(b) Parameters as in table 4, except $N = 5,000,000$ and $\theta = 2$.

Figure 5: Fat-tails observed in QQ-plots of the logarithmic returns for the Cross base model.

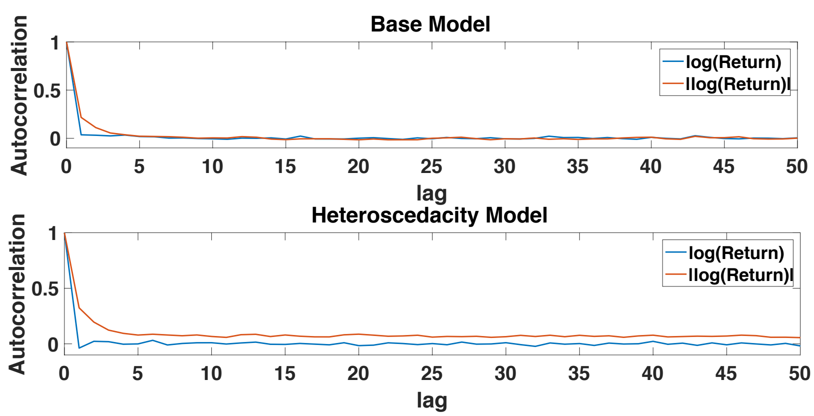
to decide in each time step which fraction of wealth he wants to invest in stocks with the remaining wealth being invested in a safe bond. The investment decision is determined by a utility maximization. For modeling details we refer to (Levy et al. 1994, 1995, 1996, 2000).

We define the model and parameter sets in detail in appendix A.2, which is the identical choice as in the earlier studies (Levy et al. 1995, 1996). We consider only one type of financial agent. The agent has a fixed memory span of $m = 15$. Figure 7 shows the simulation in the case of noise and no noise added to the investment decision. We observe that the noise leads to oscillatory behavior, which coincides with earlier findings in (Levy et al. 1995). Figure 8 shows results for three types of agents with different memory spans $m_1 = 10,\ m_2 = 141,\ m_3 = 256$. The results are qualitatively identical to the results in (Levy et al. 1996).

Finite Size Effects It was discovered earlier that the LLS model exhibits finite size effects (Zschischang and Lux 2001). In our simulations, we identify two different kinds of effects caused by a different number of agents. Our simulations are conducted first with 99 agents



(a) Upper graph: Cross base model with parameters see table 4; lower graph: Cross heteroscedacity model with parameters see table 4 except $\theta = 2$.



(b) Upper graph: Cross base model with parameters see table 4 and $N = 5,000,000$; lower graph: Cross heteroscedacity model with parameters see table 4 except $\theta = 2$ and $N = 5,000,000$.

Figure 6: Auto-correlation of log-returns and absolute log-returns in the Cross base model and the heteroscedacity models.

and then with 999 agents. First, the qq-plot of the logarithmic stock return of both simulations depicted in figure 9 clearly shows that the number of agents has a tremendous effect on the tail behavior. Secondly, Levy at al. (Levy et al. 1996) claimed that the investor group with the maximum memory becomes the dominant one, meaning they own the maximum amount of wealth. In our simulations, the wealth evolution of different agent groups changes with varying number of agent, as the figures 10a and 10b reveal. Thus, we can conclude that the

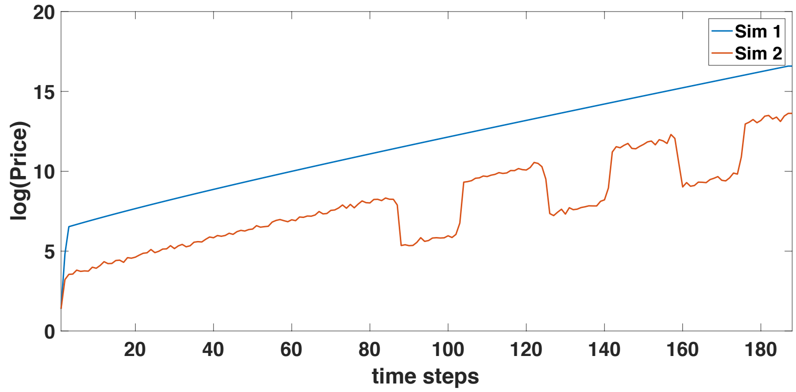


Figure 7: Price evolution of the LLS model with noise $\sigma_\gamma = 0.2$ (red) and without noise $\sigma_\gamma = 0$ (blue). Parameters as in table 5

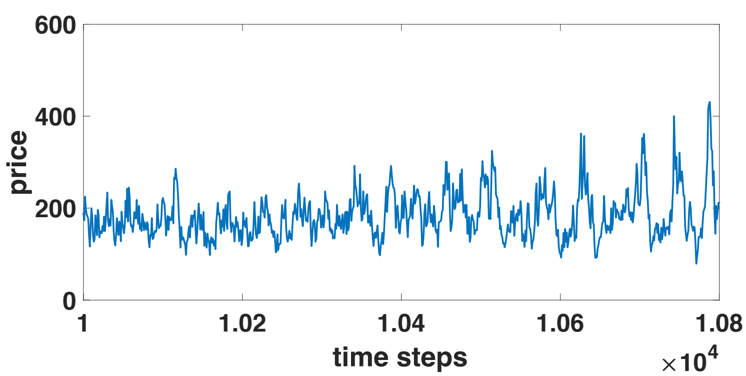


Figure 8: Price evolution of the LLS model with three different investor types. Parameters as in table 6.

qualitative output of the model changes with respect to the number of agent, which is an undesirable model characteristic.

Discussion of Model Behavior The simulation in figure 7 reveals that the deterministic model is characterized by a constant investment proportion. The optimal investment propor-

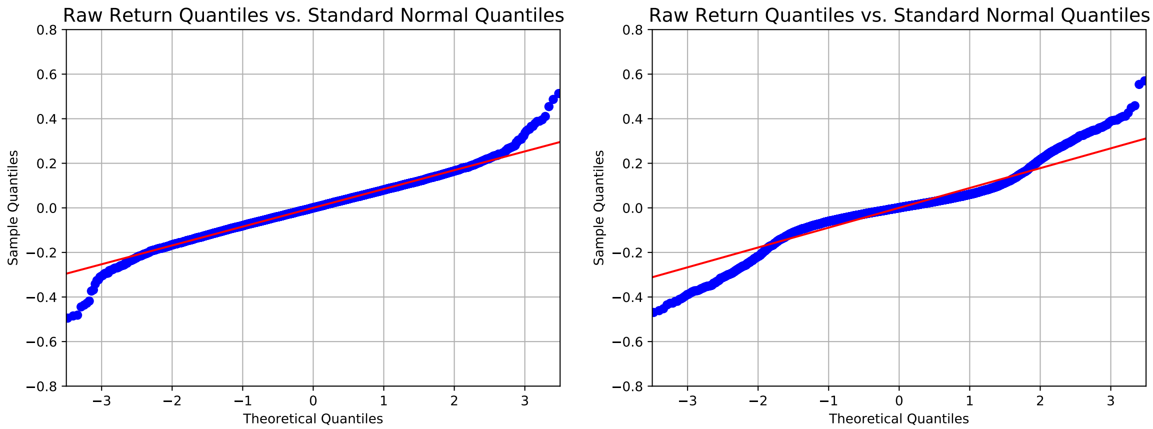


Figure 9: QQ-Plot of log-returns. Simulations conducted with 99 agents (left) and 999 agents (right), respectively. The parameters are set as in table 6.

tion is always located at the boundaries $\gamma \in \{0.01, 0.99\}$, determined through the initialization of the return history. This is an absolutely reasonable result, thus the wealth evolution in the original LLS model (Levy et al. 1994) is linear

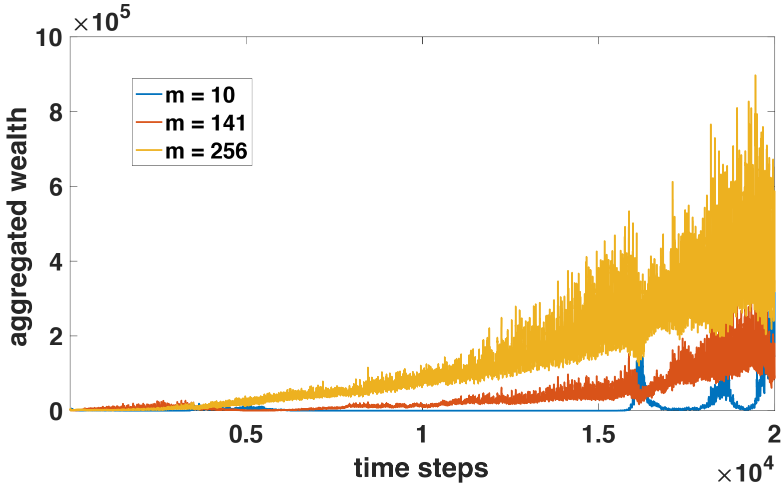
$$w(t_{k+1}) = w(t_k) + (1 - \gamma(t_k))\ r + \gamma(t_k)\ w(t_k)\ \frac{S(t_{k+1}) - S(t_k) + D(t_k)}{S(t_k)}, \tag{1}$$

and the chosen logarithmic utility function is monotonically increasing. In fact, additive noise on the optimal solutions leads to oscillatory behavior. Thus, the investors change between the two possible extreme investments of being fully invested in stocks or bonds. We point out that the noise level is crucial in order to obtain this oscillatory behavior. Figure 11 illustrates the model output for different noise levels.

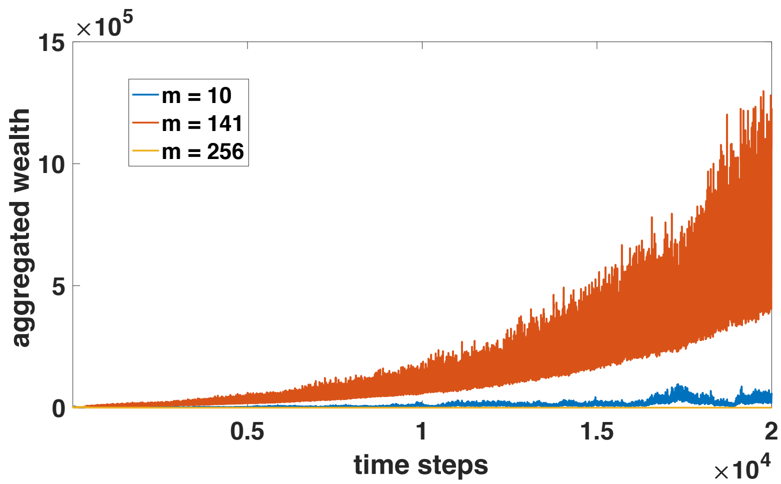
Nevertheless figure 8 seems to indicate chaotic price behavior. The previous simulations (see figs. 9, 10a and 10b) clearly reveal finite size effects. In our simulations we obtain that also in the noisy case approximately 90% of the investment decisions (pre-noise) are located at the boundaries. Mathematically this is an unsatisfying result.

3.0.3 Harras Model

The Harras model is inspired by phase transitions observed in the Ising model (Ising 1925), similar to the Cross model. The goal of the authors is to investigate the development of crashes and bubbles. The investment decision of each agent is based on their expectation of future price behavior. To be more precise, their opinion is determined by three types of information: private information, global information like news and friendship information. The friendship information is implemented through microscopic interactions on a lattice topology of agents. This introduces spatial correlations among the different investors. The topology is a good example for an environment as introduced in section 2. Besides their opinion the agents in the Harras model are equipped with their personal wealth and their number of stocks. These quantities are aggregated in the excess demand, that drives the market price in a deterministic fashion. For further modeling details we refer to the original model (Harras and Sornette 2011).



(a) Results for $N = 99$ agents.



(b) Results for $N = 999$ agents.

Figure 10: Aggregated wealth of three equally sized agent groups for different total numbers of agents and parameters as found in table 6.

In their simulations the authors have obtained that depending on the weight of the friendship information the price behavior changes. Above a certain threshold the model develops bubbles and price crashes, called *dragon-kings* by Harras and Sornette. Furthermore, the authors show that the model can generate fat-tails in asset returns and volatility clustering as well.

We cannot reproduce all results of Harras and Sornette. We obtain that the tails of the

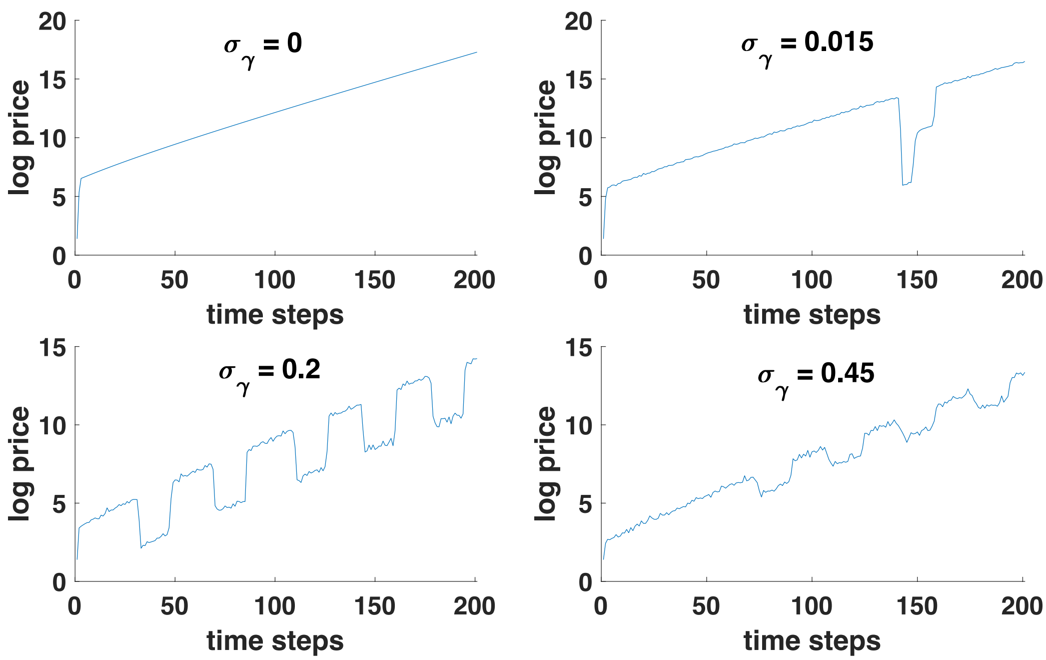


Figure 11: Simulations of the basic LLS model with varying noise levels. Parameters as set in table 5.

logarithmic stock return distribution are slimmer than for the Gaussian distribution (see qq-plot of figure 13), but we obtain volatility clustering as indicated by the auto-correlation plot in figure 13. We cannot observe bubbles and crashes as presented in (Harras and Sornette 2011). The differences may be caused by incomplete information regarding the parameter settings and modeling details. We refer to the appendix A.2 for details.

3.0.4 Franke-Westerhoff Model

In this work we consider the Franke-Westerhoff model as discussed in (Franke and Westerhoff 2012). Earlier model variants can be found in (Franke and Westerhoff 2009, 2011). The model studies the evolution of two groups of traders, one group with a chartist strategy and the other with a fundamental strategy. This evolution can be interpreted in the sense of the meta-model introduced in (Trimborn et al. 2018a) as the evolution of two representative agents. In each time step a fraction of traders may adapt their investment strategy by a switching process based on socio-economic factors. These factors are for example a comparison of the agents' estimated wealth or a herding mechanism. The logarithmic stock price is then driven by the aggregated excess demand of both groups. As a distinct feature of the Franke-Westerhoff model they employ the concept of *structural stochastic volatility* as introduced in (Franke and Westerhoff 2009). This means that the demand of chartist and fundamentalist has a

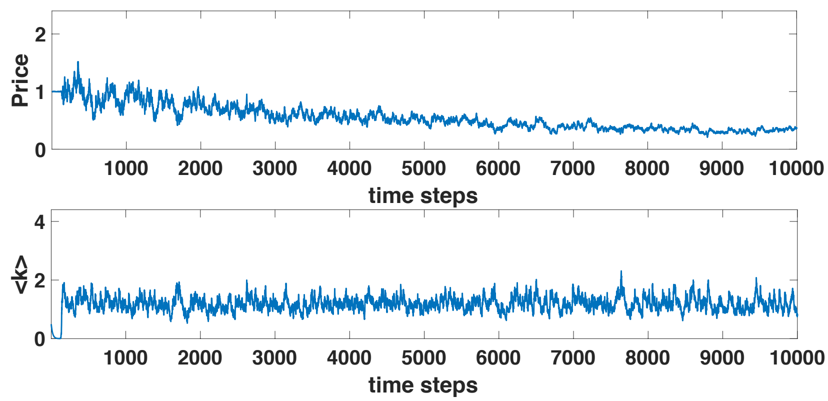


Figure 12: Harras model with no friendship information ($C_1 = 0$). Parameters as in table 7.

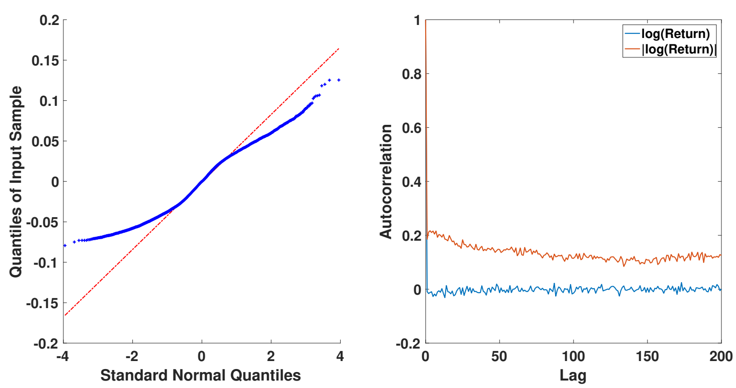


Figure 13: Harras model with no friendship information ($C_1 = 0$). Parameters as in table 7.

stochastic component which takes into account heterogeneity within agent groups and uncertain events. Furthermore, the variance of these random terms differ between both investor groups.

The second distinct feature of the Franke-Westerhoff model is the possibility to choose between two well known switching mechanism. The discrete choice approch (DCA) introduced by Brock and Hommes (Brock and Hommes 1997) or the transition probability approach (TPA) introduced by Weidlich and Haag (Weidlich and Haag 2012) and Lux (Lux 1995).

The authors have shown in several publications (Franke and Westerhoff 2009, 2011, 2012) that their model fits the statistical features of real financial markets such as volatility clustering or fat-tails of stock returns extremely well. Furthermore, they have even employed the

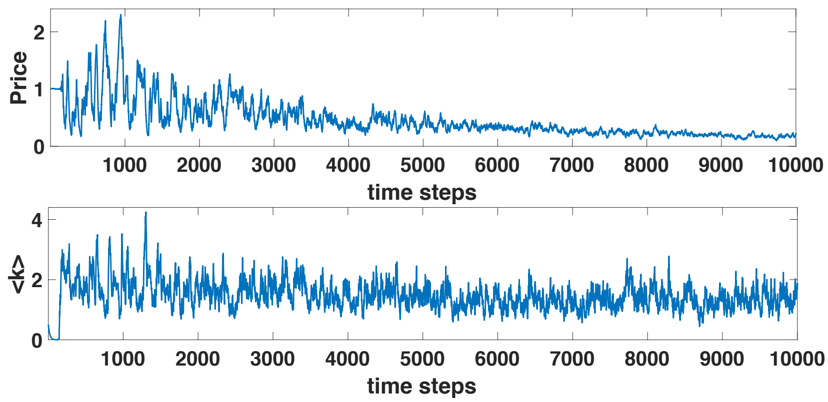
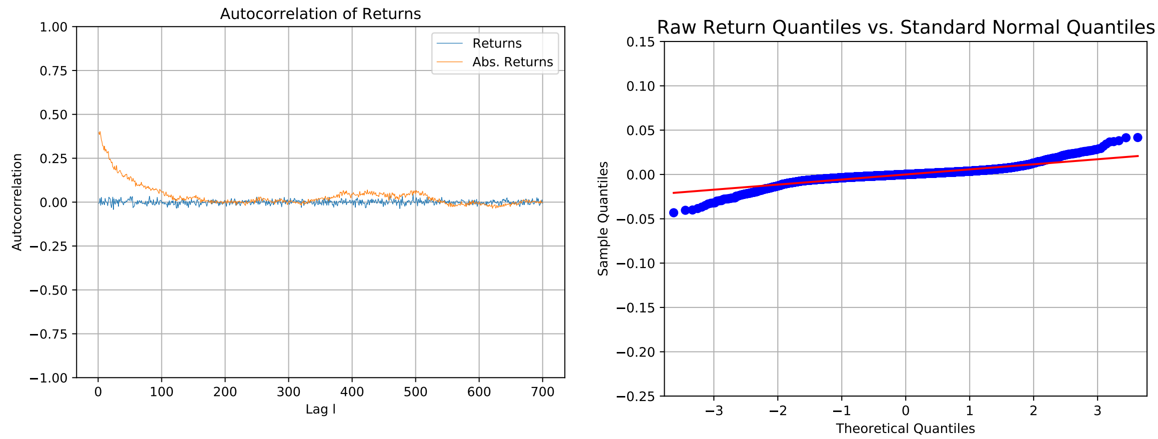


Figure 14: Harras model with friendship information ($C_1 = 4$). Parameters as in table 7 except $C_1 = 4$.

method of simulated moments (Franke 2009) in order to fit their model parameters to original financial data.

First, we present the simulations of the Franke-Westerhoff model with discrete choice approach and the behavioral factors herding (H), predisposition (P) and misalignment (M). This model choice can be conveniently abbreviated by DCA-HPM. Figure 15 shows a plot of the auto-correlation function of raw and absolute logarithmic returns. The auto-correlation of raw returns is approximately zero which indicates absence of auto-correlation. In comparison to the raw returns, we obtain a slow algebraic decay in the case of absolute returns. This is known as volatility clustering. In addition, the quantile-quantile plot indicates fat-tails for the logarithmic stock returns.

As a second aspect we investigate the impact of both switching mechanisms. We have simulated the HPM case for the DCA and TPA approach. Figure 16 reveals that the mean fraction of chartists is slightly larger in the DCA approach than in the TPA method (see table 2 as well). Furthermore, figure 16 shows that he fractions of chartist in the DCA case are much more volatile. For our simulations we have used same random number seed in both runs. Qualitatively, the results of both switching mechanisms do not differ. As presented by Franke and Westerhoff in (Franke and Westerhoff 2011) we run a model contest as well. We have averaged our simulation over 200 runs. Table 2 depicts the different values for the excess kurtosis, the hill estimator and the average fractions of chartists. We do not obtain a model which significantly dominates all other choices with respect to reproducing stylized facts. Finally, we have conducted a test in order to study the long time behavior of the model. In fact figure 17 shows that there is no qualitative change in the evolution of prices in the long run.



(a) Auto-correlation of returns and absolute returns.

(b) QQ plot of raw returns.

Figure 15: Macroscopic statistics for a DCA-HPM simulation. Parameters as in table 11.

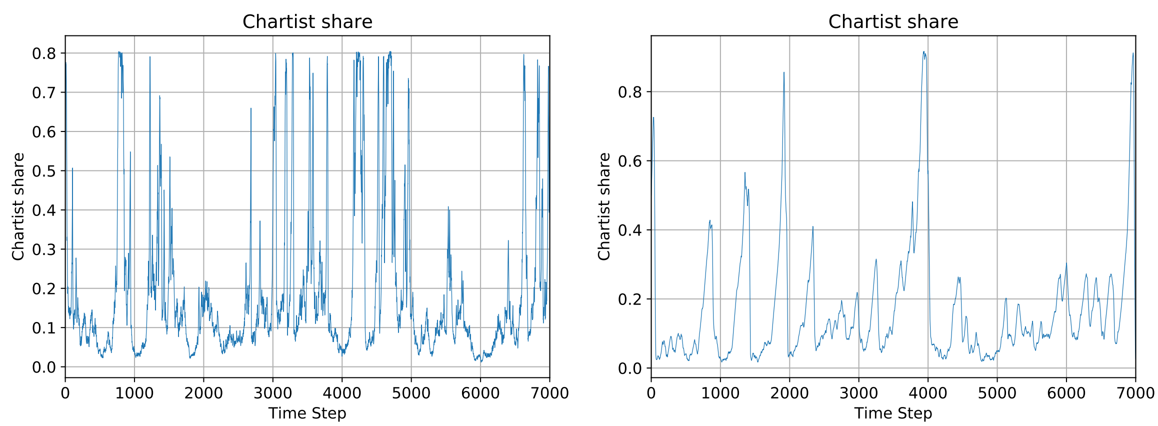


Figure 16: Chartist share for DCA-HPM (left hand side) and TPA-HPM (right hand side) switching mechanisms. Parameters as in table 11 and 14, respectively.

4 Novel ABCEM Models

In this section we demonstrate the flexibility of the SABCEMM simulator by creating new models and adding new features to the Cross model. More precisely, we consider the Cross agents in combination with new market mechanisms. Especially, we can show that the precise form of the market mechanism can have a direct impact on the appearance of stylized facts. In addition, we modify the Cross agents by adding a wealth evolution to each agent and study the statistical properties of the wealth evolution.

Wealth Evolution We consider the original Cross model as defined in appendix A.2. We add a wealth evolution of the type

$$w_i(t_{k+1}) = w_i(t_k) + \Delta t \ \left[(1-\gamma)\ r + \gamma \frac{S(t_k) - S(t_{k-1})}{\Delta t\ S(t_k)} \right] w_i(t_k),$$

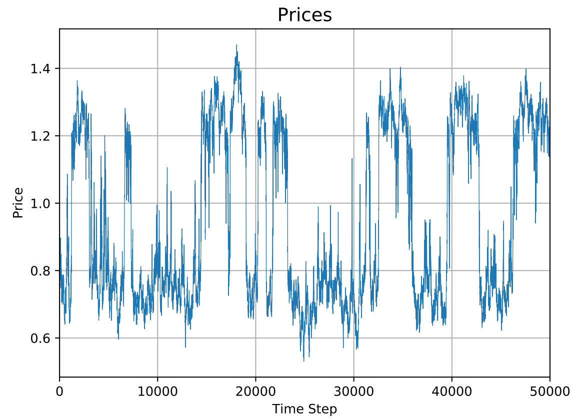


Figure 17: Logarithmic prices for a DCA-HPM simulation (50,000 time steps). Parameters as in table 11, with 50,000 time steps.

	excess kurtosis	Hill estimator	avarage chartist share
TPA-W	5.9512	3.344	0.2813
TPA-WP	7.025	3.1957	0.2507
TPA-HPM	8.614	2.5833	0.1503
DCA-W	8.2023	3.173	0.2577
DCA-WP	7.7600	3.1314	0.2285
DCA-HPM	10.033	2.481	0.1674
DCA-WHP	8.01	3.1192	0.2227
TPA average	7.1967	3.041	0.2274
DCA average	7.99	2.895	0.2179

Table 2: Model contest, parameters as in tables 8 to 14. Each simulation run lasted for 7,000 time steps. The random seed is chosen differently for each repetition, but identical for each of the seven model variants.

to each Cross agent $i = 1, ..., N$. The positive constant $r > 0$ denotes the interest rate and $\gamma \in (0, 1)$ a fixed fraction of stock investments for all agents. The goals was to study the influence of the non-Gaussian return distribution on the wealth distribution. With increasing γ, we obtain an increasing excess kurtosis of the wealth distribution. In fact the excess kurtosis approaches the excess kurtosis of the stock price, which is approximately 6. In figure 18 we averaged the results over 100 runs. The qq-plot in figure 20 clearly shows the fat-tail behavior for a fixed $\gamma = 1$.

Furthermore, we could generalize the microscopic excess demand of the Cross agents by adding a wealth dependency. This would lead to an additional coupling between wealth and stock price behavior. We leave this question open for further research.

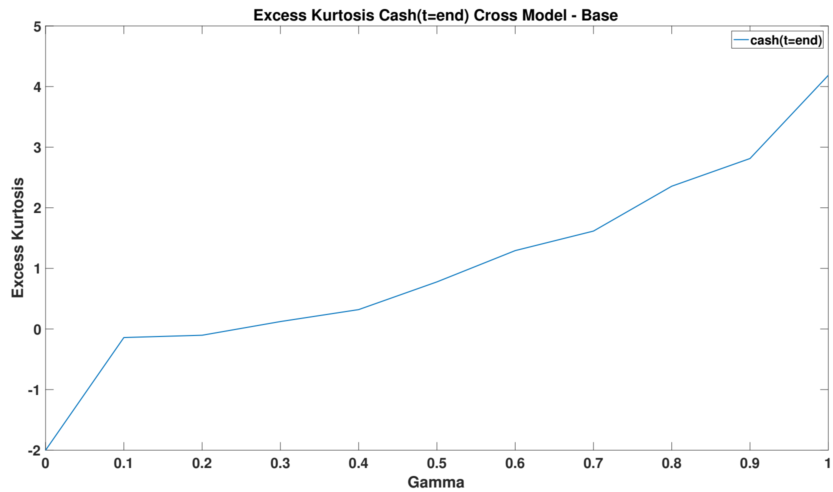


Figure 18: Excess kurtosis for the wealth at the final timestep of the Cross model. Parameters as in table 4 with $w_i(t = 0) = 1$ and $r = 0.01$.

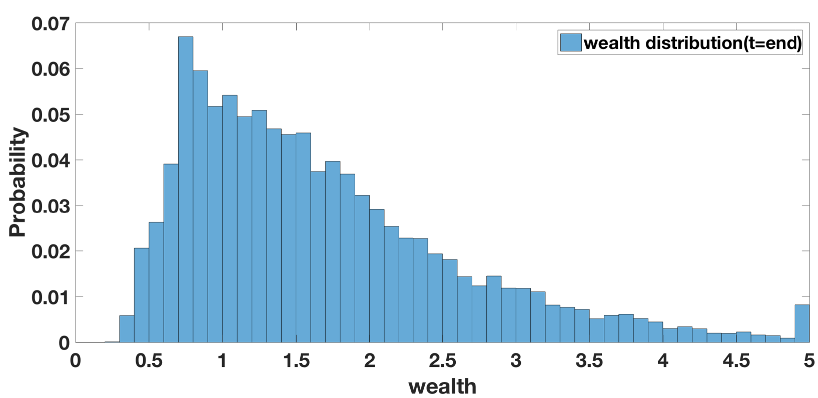


Figure 19: Histogram of the wealth distribution at the final timestep. Parameters as in table 4 with $w_i(t = 0) = 1$ and $r = 0.01$.

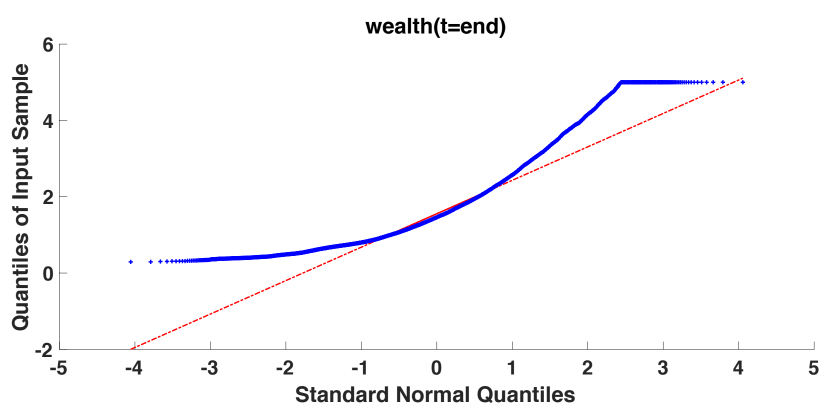


Figure 20: QQ-Plot of the wealth distribution. Parameters as in table 4 with $w_i(t = 0) = 1$ and $r = 0.01$.

SDE Discretization In this section, we again consider the Cross agents, but we change the clearance mechanism. We set the pricing rule to be

$$S(t_{k+1}) = S(t_k) + \Delta t\ F_{Cross}(t_k, S, ED) + \sqrt{\Delta t}\ S(t_k)\ (1 + \theta\ |ED(t_k)|)\ \eta, \quad \eta \sim \mathcal{N}(0,1).$$

We choose the Euler-Maruyama discretization of the SDE:

$$dS = F_{Cross}(S, ED)\ dt + S\ (1 + \theta\ |ED|)\ dW,$$

where W is a Wiener process and the SDE should be interpreted in the Itô sense. For our first simulation, we set the drift operator F to read:

$$F^1_{Cross}(t, S, ED) := S(t)\ \frac{d}{dt} ED(t).$$

As the figures 21 and 22 reveal, the behavior of this model is identical to the original Cross model. Thus, we have obtained a pricing rule which is a proper time discretization of a time continuous model.

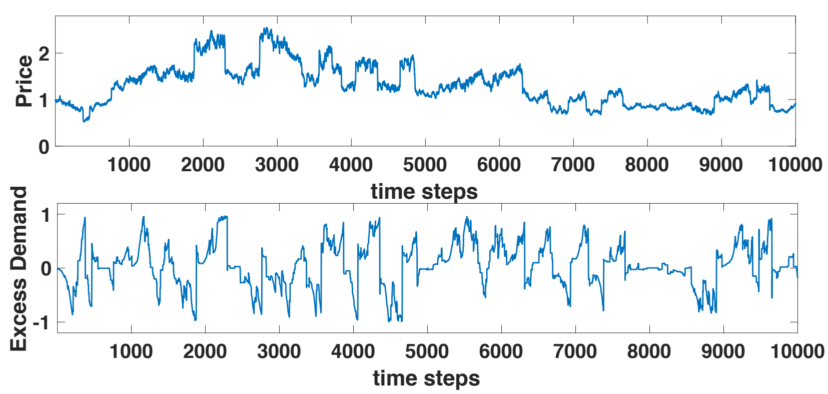


Figure 21: Cross agents with SDE pricing rule. Parameters as in table 4 except $\theta = 2$.

In a second test we set the drift coefficient to

$$F^2_{Cross}(t, S, ED) := S(t)\ ED(t).$$

We study the different impact of the two models on the stock price behavior. Our studies reveal that the the choice F^2_{Cross} leads to Gaussian price behavior. We measured this with the excess kurtosis, which we averaged over 100 runs. This is an interesting result, since it reveals the great influence of the drift coefficient on the stock price behavior.

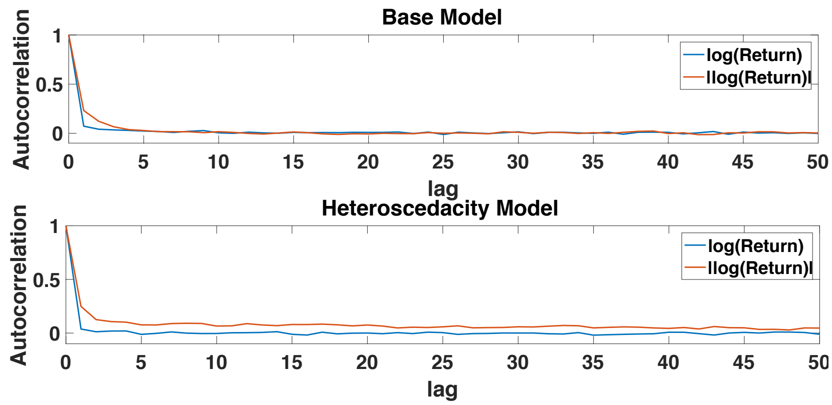


Figure 22: Auto-correlation of log-returns and absolute log-returns in the base model (upper graph, parameters see table 4) and full model (parameters see table 4 except $\theta = 2$). We obtain the same behavior as previously in section 3.0.1.

	excess kurtosis
F^1_{Cross}, $\theta = 0$	27.0434
F^1_{Cross}, $\theta = 2$	22.5530
F^2_{Cross}, $\theta = 0$	-0.0044
F^2_{Cross}, $\theta = 2$	1.2580

Table 3: Averaged excess kurtosis over 100 runs with different drift functions.

5 Conclusion

We have presented various simulation results generated with the new simulator SABCEMM. Therefore, we have given a short introduction of the SABCEMM software. Afterwards we have carried out several simulations of the LLS model, Cross model, Harras model and Franke-Westerhoff model. Here, we ran simulations with up to several million agents in order to exclude finite size effects. We could demonstrate the efficiency of the SABCEMM tool with respect to the recombination of building blocks. Thus, we have created new models simply by interchanging the market mechanism.

Our findings and especially the obtained stylized facts coincide with the findings in literature. The only exception is the stock price behavior in the Harras model since we do not obtain *dragon-kings* and fat-tails in the stock return distribution. The extended numerical studies such as simulations with many agents or long time simulations have supported the validity of the particular model. The study of the new ABCEM models which use the agent design of the Cross model have revealed several new insights. Technically, we have shown the feasibility of creating reasonable new models by simply changing one line of code. Economically, enables us to study the origins of stylized facts with respect to different building blocks. Our studies have shown that the choice of the market mechanism heavily influences the fat-tail property of the stock return data. Hence, we conclude that the precise form of the market mechanism is of imminent importance in order to generate realistic stock price data. Furthermore, we have seen that the fat-tail behavior in stock returns translates one to one to the tail of the wealth. Interestingly, the fraction of investments in the stock return determines the size of the tail. Hence, the more money is invested in stocks the more prominent is the tail in the wealth distribution.

These observations lead us to the conclusions that the combination of further building blocks may help to find the origins of stylized facts and to understand the mechanism which create them.

Acknowledgement

Torsten Trimborn gratefully acknowledges support by the Hans-Böckler-Stiftung and the RWTH Aachen University Start-Up grant. The work was partially funded by the Excellence Initiative of the German federal and state governments.

A Appendix

A.1 Mathematical Appendix

Definition 1. *The Hill estimator of a sample $X \in \mathbb{R}^n_{>0}$ sorted as $x_1 \geq x_2 \geq ... x_n$ is defined as*

$$H := \left(\frac{1}{k} \sum_{i=1}^{k} \ln(x_i) - \ln(x_k) \right)^{-1},$$

with $k := \lfloor 0.05 * n \rfloor$. For details we refer to (Hill 1975).

A.2 Models

Cross Model We present the Cross model as defined in (Cross et al. 2005).

We assume a fixed number of $N \in \mathbb{N}$ agents. Each agent decides in each time step, whether he wants to be long or short in the market. Thus, the investment propensity σ_i, $1 \leq i \leq N$ of each agent switches between $\sigma_i = \pm 1$. The excess demand of all investors at time $[0, \infty)$ is then defined as:

$$ED(t_k) := \frac{1}{N} \sum_{i=1}^{N} \sigma_i(t_k).$$

Furthermore, the model introduces two pressures, the *herding pressure* and the *inaction pressure*, which control the switching mechanism.
The *inaction pressure* is defined by the interval

$$I_i = \left[\frac{m_i}{1 + \alpha_i}, m_i \; (1 + \alpha_i) \right],$$

where m_i denotes the stock price of the last switch of agent i and $\alpha_i > 0$ is the so called *inaction threshold*. The *herding pressure* is given by:

$$\begin{cases} c_i(t_{k+1}) = c_i(t_k) + \Delta t |ED(t_k)|, & \text{if } \sigma_i(t_k) \; ED(t_k) < 0 \\ c_i(t_{k+1}) = c_i(t_k), & \text{otherwise.} \end{cases}$$

In addition, one defines the *herding threshold* β_i. The thresholds are chosen once randomly from an i.i.d. random variable, which is uniformly distributed.

$$\alpha_i \sim \mathrm{Unif_c}(A_1, A_2), \; A_2 > A_1 > 0,$$
$$\beta_i \sim \mathrm{Unif_c}(B_1, B_2), \; B_2 > B_1 > 0.$$

The constants B_1 and B_2 have to scale with time, since they correspond to the time units an investor can resist the herding pressure.

$$B_1 := b_1 \cdot \Delta t,$$
$$B_2 := b_2 \cdot \Delta t.$$

Switching mechanism The switching is then induced if

$$c_i > \beta_i \text{ or } S(t) \notin I_i.$$

After a switch the herding pressure is reset to zero and the inaction interval gets updated as well. The stock price is then driven by the excess demand:

$$\begin{aligned} &S(t_{k+1}) = S(t_k)\ \exp\{(1+\theta\ |ED(t_k)|)\ \sqrt{\Delta t}\ \eta(t_k) + \kappa\ \Delta t\ \frac{\Delta ED(t_k)}{\Delta t}\}, \\ &\sqrt{\Delta t}\ \eta \sim \mathcal{N}(0, \Delta t) \\ &\Delta ED(t_k) := \frac{1}{N}\sum_{i=1}^{N} \sigma_i(t_k) - \frac{1}{N}\sum_{i=1}^{N} \sigma_i(t_{k-1}), \end{aligned}$$

where κ denotes the market depth and $\Delta t > 0$ the time step.

Cross model extensions: One alternative pricing function is given by:

$$S(t_{k+1}) = S(t_k)\ + \Delta t\ \kappa\ \frac{\Delta ED(t_k)}{\Delta t}\ S(t_k) + \sqrt{\Delta t}\ F_{Cross}(t_k, S, ED)\ S(t_k)\ \eta,$$

Furthermore, we have added the wealth evolution, for a fixed interest rate $r > 0$ and fixed investment fraction $\gamma \in (0,1)$:

$$w_i(t_{k+1}) = w_i(t_k) + \Delta t\ \left[(1-\gamma)\ r + \gamma \frac{S(t_k) - S(t_{k-1})}{\Delta t\ S(t_k)}\right] w_i(t_k).$$

LLS Model We have implemented the model as defined in (Levy et al. 1994, 1995). We have added one possible time scale to the model. In order to obtain the original model one needs to set $\Delta t = 1$.

The model considers $N \in \mathbb{N}$ financial agents who can invest $\gamma_i \in [0.01, 0.99],\ i = 1, ..., N$ of their wealth $w_i \in \mathbb{R}_{>0}$ in a stocks and have to invest $1 - \gamma_i$ of their wealth in a safe bond with interest rate $r \in (0,1)$. The investment propensities γ_i are determined by a utility maximization and the wealth dynamic of each agent at time $t \in [0, \infty)$ is given by

$$\begin{aligned} w_i(t_k) &= w_i(t_{k-1}) \\ &+ \Delta t \left((1-\gamma_i(t_{k-1}))\ r\ w_i(t_{k-1}) + \gamma_i(t_{k-1})\ w_i(t_{k-1})\ \underbrace{\frac{\frac{S(t_k)-S(t_{k-1})}{\Delta t} + D(t_k)}{S(t_k)}}_{=:x(S,t_k,D)} \right). \end{aligned}$$

The dynamics is driven by a multiplicative dividend process. Given by:

$$D(t_k) := (1 + \Delta t\ \tilde{z})\ D(t_{k-1}),$$

where $\tilde{z}$ is a uniformly distributed random variable with support $[z_1, z_2]$. The price is fixed by the so called *market clearance condition*, where $n \in \mathbb{N}$ is the fixed number of stocks and $n_i(t)$ the number of stocks of each agent.

$$n = \sum_{i=1}^{N} n_i(t_k) = \sum_{k=1}^{N} \frac{\gamma_k(t_k)\ w_k(t_k)}{S(t_k)}. \tag{2}$$

The utility maximization is given by

$$\max_{\gamma_i \in [0.01, 0.99]} E[\log(w(t+\Delta t, \gamma_i, S^h))].$$

with

$$E[\log(w(t_{k-1}, \gamma_i, S^h))] = \frac{1}{m_i} \sum_{j=1}^{m_i} U_i \Bigg((1-\gamma_i(t_k)) w_i(t_k, S^h)\,(1+r\Delta t) \\ + \gamma_i(t_k) w_i(t_k, S^h) \Big(1 + x\big(S, t_k - j\Delta t, D\big)\ \Delta t\Big) \Bigg).$$

The constant m_i denotes the number of time steps each agent looks back. Thus, the number of time steps m_i and the length of the time step Δt defines the time period each agent extrapolates the past values. The superscript h indicates, that the stock price is uncertain and needs to be fixed by the market clearance condition. Finally, the computed optimal investment proportion gets blurred by a noise term.

$$\gamma_i(t_k) = \gamma_i^*(t_k) + \epsilon_i,$$

where ϵ_i is distributed like a truncated normally distributed random variable with standard deviation σ_γ.

Utility maximization Thanks to the simple utility function and linear dynamics we can compute the optimal investment proportion in the cases where the maximum is reached at the boundaries. The first order necessary condition is given by:

$$f(\gamma_i) := \frac{d}{dt} E[\log(w(t_{k+1}, \gamma_i, S^h))] = \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{\Delta t\ (x\big(S, t_k - j\Delta t, D\big) - r)}{\Delta t\ (x\big(S, t_k - j\Delta t, D\big) - r)\ \gamma_i + 1 + \Delta t\ r}.$$

Thus, for $f(0.01) < 0$ we can conclude that $\gamma_i = 0.01$ holds. In the same manner, we get $\gamma_i = 0.99$, if $f(0.01) > 0$ and $f(0.99) > 0$ holds. Hence, solutions in the interior of $[0.01, 0.99]$ can be only expected in the case: $f(0.01) > 0$ and $f(0.99) < 0$. This coincides with the observations in (Samanidou et al. 2007).

Harras model We present the Harras model as defined in (Harras and Sornette 2011).

We consider N financial agents where each agent is equipped with a personal opinion $\psi_i(t_k)$, and t_k denotes a discrete time step. The personal opinion is created through the personal information of each agent $\epsilon_i(t_k)$, public information $n(t_k)$ and the expected action of the surrounded neighbor j by the agent i, $E_i[\sigma_j(t_k)], \sigma_j \in \{-1, 0, 1\}$. The opinion of the i-th agent at time t_k reads:

$$\psi_i(t_k) = c_{1,i} \sum_{j=1}^{4} k_{ij}(t_{k-1})\ E[\sigma_i(t_k)] + c_{2,i} u(t_{k-1})\ n(t_k) + c_{3,i}\ \epsilon_i(t_k). \tag{3}$$

During the evaluation of our simulations we noticed a significant difference in the magnitude of the price's volatility. Our investigation leads us to the conclusion that the opinion of the i-th agent at time t_k should instead be:

$$\psi_i(t_k) = c_{1,i}\frac{1}{4}\sum_{j=1}^{4} k_{ij}(t_{k-1})\ E[\sigma_i(t_k)] + c_{2,i}u(t_{k-1})\ n(t_k) + c_{3,i}\ \epsilon_i(t_k).$$

The weights $(c_{1,i}, c_{2,i}, c_{3,i})$ are chosen initially for each agent from three uniformly distributed random variables on the domains $[0, C_l],\ l \in \{1, 2, 3\}$. The private and public information $\epsilon_i(t_k), n(t_k)$ are modeled as standard normally distributed i.i.d. random variables. The agents are grouped on a virtual square lattice with periodic boundary conditions, such that each agent has four neighbors. We update of the opinion of each agent is performed in each time step in random order. The additional factor k_{ij} weights the predicted action of the j-th agent based on the past performance. In the same manner the factor u weights the public information stream. The update rule of these weighting factors is given by:

$$\begin{aligned} u(t_k) &= \alpha\ u(t_{k-1}) + (1-\alpha)\ n(t_{k-1})\frac{ED(t_k)}{\sigma_{ED}(t_k)}, \\ k_{i,j}(t_k) &= \alpha\ k_{i,j}(t_{k-1}) + (1-\alpha)\ E_i[\sigma_j(t_{k-1})]\frac{ED(t_k)}{\sigma_{ED}(t_k)}, \end{aligned}$$

with the constant $0 < \alpha < 1$ and the volatility

$$\begin{aligned} &\sigma^2_{ED}(t_k) = \alpha\ \sigma^2_{ED}(t_{k-1}) + (1-\alpha)\ (ED(t_{k-1}) - \langle ED(t_k)\rangle)^2, \\ &\langle ED(t_k)\rangle = \alpha\ \langle ED(t_{k-1})\rangle + (1-\alpha)\ ED(t_{k-1}), \end{aligned}$$

where the brackets $\langle\cdot\rangle$ denote the expected excess demand ED. The agent's action on the market is then determined by a threshold of each agent. The threshold $\bar{\psi}_i$ is drawn from a uniform distribution in the interval $[0, \Omega]$. The trading decision of each agent is characterized by $\sigma_i = \pm 1$, where $\sigma_i = 1$ represents a buy order and $\sigma = -1$ a sell order. We have

$$\sigma_i(t_k) = \begin{cases} 1, & \psi_i(t_k) > \bar{\psi}_i, \\ -1, & \psi_i(t_k) < -\bar{\psi}_i, \\ 0, & \text{else.} \end{cases}$$

Furthermore, each agent tracks his number of stocks q_i and the cash w_i and thus the trading volume $v_i(t_k)$ of each agent (in units of stocks) is given by

$$v_i(t_k) = \begin{cases} g\ \frac{w_i(t_k)}{S(t_k)}, & \sigma_i = 1, \\ g\ q_i(t_k), & \sigma = -1. \end{cases}$$

Here, $S(t_k)$ denotes the stock price and $g \in (0, 1)$ a fixed fraction of wealth each agent wants to invest in stocks. The stock price is then driven by the excess demand $ED(t_k)$

$$\log(S(t_k)) = \log(S(t_{k-1})) + ED(t_{k-1}),$$

where ED is defined as:

$$ED(t_k) := \frac{1}{\lambda\ N}\sum_{i=1}^{N} \sigma_i(t_k)\ v_i(t_k).$$

The constant $\lambda > 0$ represents the market depth. Finally, we want to state the update mechanism of w_i and q_i.

$$w_i(t_k) = w_i(t_{k-1}) - \sigma_i(t_{k-1})\ v_i(t_{k-1})\ S(t_k),$$
$$q_i(t_k) = q_i(t_{k-1}) + \sigma_i(t_{k-1})\ v_i(t_{k-1}).$$

Unknown implementation details: While implementing the model from (Harras and Sornette 2011), we were missing some crucial implementation details. First of all, no information was given concerning the initialization of most variables. We choose them according to table 7. Further the term $E[\sigma_i(t_k)]$ used during the calculation of the opinion (3) is not defined. We assume this to be the decision σ_i of agent i. Since the update of the agents is done in random order σ_i can either be $\sigma_i(t_{k-1})$ if the agent i has yet to be updated or $\sigma_i(t_k)$ if the agent i has already been updated. This assumption was later confirmed by Harras via mail correspondence. Furthermore, the update order can not be deduced from the publication (Harras and Sornette 2011). We have ordered the process as defined through the previously introduced discrete time steps.

Franke-Westerhoff model The Franke-Westerhoff model (Franke and Westerhoff 2011) considers tow types of agents, chartists and fundamentalists. The demand of each agent reads

$$d^f(t_k) = \phi(P_f - P(t_k)) + \epsilon_k^f, \quad \phi \in \mathbb{R}^+, \quad \epsilon_k^f \sim \mathcal{N}(0, \sigma_f^2), \tag{4}$$
$$d^c(t_k) = \chi(P(t_k) - P(t_{k-1})) + \epsilon_k^c, \quad \chi \in \mathbb{R}^+, \quad \epsilon_k^c \sim \mathcal{N}(0, \sigma_c^2), \tag{5}$$

where $P(t_k)$ denotes the logarithmic market price and P_f denotes the fundamental price. The noise terms ϵ_k^f and ϵ_k^c are normally distributed, with zero mean and different standard deviations σ_c^2 and σ_f^2. The second important features are the fractions of the chartist or fundamental population. In that sense the two agents can be seen as representative agents of a population. The fraction of chartists $n^C(t_k) \in [0, 1]$ and the fraction of fundamentalitst $n^F(t_k) \in [0, 1]$ have to fulfill $n^C(t_k) + n^F(t_k) = 1$. Hence, the excess demand can be define as:

$$ED^F(t_k) := \frac{1}{2}(ed^f(t_k) + ed^c(t_k)) \tag{6}$$

$$ed^f(t_k) := 2\ n^f(t_k)d^f(t_k)$$
$$ed^c(t_k) := 2\ n^c(t_k d^c(t_k). \tag{7}$$

The pricing equation is then given by the simple rule

$$P(t_k) = P(t_{k-1}) + \mu ED^F(t_{k-1}). \tag{8}$$

Finally, we need to specify the switching mechanism. We have implemented two possible switching mechanisms, the transition probability approach (TPA) (Weidlich and Haag 2012; Lux 1995) and the discrete choice approach (DCA) (Brock and Hommes 1997) approach. In both cases we consider the so called switching index $a(t_k) \in \mathbb{R}$ which describes the attractiveness of the fundamental strategy over the chartist strategy. Thus, a positive $a(t_k)$ reflects an advantage of the fundamental strategy in comparison to the chartist and if $a(t_k)$ is negative we have the opposite situation.

DCA In the DCA case we obtain

$$n^f(t_k) = \frac{1}{1+exp(-\beta a(t_{k-1}))},$$
$$n^c(t_k) = \frac{1}{1+exp(\beta a(t_{k-1}))}, \tag{9}$$

where the parameter $\beta > 0$ measures the intensity of choice.

TPA In the TPA case, we first define switching probabilities

$$\pi^{cf}(a(t_{k-1})) = min[1, \nu \exp(a(t_{k-1}))],$$
$$\pi^{fc}(a(t_{k-1}) = min[1, \nu \exp(-a(t_{k-1}))],$$

where π^{xy} is the probability that an agent with strategy x switches to strategy y. The flexibility parameter ν is a scaling factor for $a(t_k)$. Then the time evolution of chartist and fundamentalist shares is given by:

$$n^f(t_k) = n^f(t_{k-1}) + n^c(t_{k-1})\pi^{cf}(a(t_{k-1})) - n^f(t_{k-1})\pi^{fc}(a(t_{k-1}))$$
$$n^c(t_k) = n^c(t_{k-1}) + n^f(t_{k-1})\pi^{fc}(a(t_{k-1})) - n^c(t_{k-1})\pi^{cf}(a(t_{k-1}))$$

Finally, we have to specify how the switching index $a(t_k)$ is calculated. The switching index $a(t_k)$, encodes how favourable a fundamentalist strategy is over a chartist strategy. The switching index is determined linearly out of the the four principles *wealth comparison, predisposition, herding and misalignment.*

$$\alpha(t_k) = \alpha_w\ (w^f(t_k) - w^c(t_k)) + \alpha_0 + \alpha_h\ (n^f(t_k) - n^c(t_k)) + \alpha_m\ (P(t_k) - P_f)^2,$$

where $\alpha_w, \alpha_0, \alpha_h, \alpha_m > 0$ are weights respectively scaling factors. The sign $\alpha_p \in \mathbb{R}$ determines the predisposition with respect to a fundamental or chartist strategy. The hypothetical wealth $w^f(t_k), w^c(t_k)$ is determined as follows:

$$\begin{aligned}
&w^f(t_k) := m\ w^f(t_{k-1}) + (1-m)g^f(t_k),\\
&g^f(t_k) := [exp(P(t_k)) - exp(P(t_{k-1}))]d^f(t_{k-2}),\\
&w^c(t_k) := m\ w^c(t_{k-1}) + (1-m)g^c(t_k),\\
&g^c(t_k) := [exp(P(t_k)) - exp(P(t_{k-1}))]d^c(t_{k-2}).
\end{aligned}$$

Here, the memory variable $m \in [0, 1]$ weights the past performance with the most recent return. For details regarding the modeling we refer to (Franke and Westerhoff 2011).

A.3 Parameter sets

Cross Model

LLS Model The initialization of the stock return is performed by creating an artificial history of stock returns. The artificial history is modeled as a Gaussian random variable with mean μ_h and standard deviation σ_h. Furthermore, we have to point out that the increments of the dividend is deterministic, if $z_1 = z_2$ holds. We used the C++ standard random number generator for all simulations of the LLS model if not otherwise stated.

Parameter	Value
N	1000
A_1	0.1
A_2	0.3
b_1	25
b_2	100
$w_i(t=0)$	1 $\forall 1 \le i \le N$
time steps	$10,000$
Δt	$4 \cdot 10^{-5}$
κ	0.2
θ	0
$S(t=0)$	1

(a) Parameters of Cross model.

Variable	Initial Value
$ED(t=0)$	$\frac{1}{N}\sum\limits_{i=1}^{N}\gamma_i(0)$
$c_i(t=0)$	$B_1 + \texttt{rand}\ (B_2 - B_1),\ \forall 1 \le i \le N$
$m_i(t=0)$	$S(t=0),\ \forall 1 \le i \le N$
$\sigma_i(t=0)$	$\mathrm{Unif_d}(\{-1,1\})$

(b) Initial values of Cross model.

Table 4: Cross basic setting.

Parameter	Value
N	100
m_i	15
σ_γ	0 or 0.2
r	0.04
$z_1 = z_2$	0.05
Δt	1
time steps	200

(a) Parameters of LLS model.

Variable	Initial Value
μ_h	0.0415
σ_h	0.003
$\gamma(t=0)$	0.4
$w_i(t=0)$	1000
$n_i(t=0)$	100
$S(t=0)$	4
$D(t=0)$	0.2

(b) Initial values of LLS model.

Table 5: Basic setting of the LLS model.

Harras Model

Franke-Westerhoff Model

Parameter	Value
N	99
m_i	$10,\ 1 \leqslant i \leqslant 33$ $141,\ 34 \leqslant i \leqslant 66$ $256,\ 67 \leqslant i \leqslant 99$
σ_γ	0.2
r	0.0001
$z_1 = z_2$	0.00015
Δt	1
time steps	$20,000$

(a) Parameters of LLS model.

Variable	Initial Value
μ_h	0.0415
σ_h	0.003
$\gamma_i(t=0)$	0.4
$w_i(t=0)$	1000
$n_i(t=0)$	100
$S(t=0)$	4
$D(t=0)$	0.004

(b) Initial values of LLS model.

Table 6: Setting for the LLS model (3 agent groups).

Parameter	Value
C_1	0
C_2	1
C_3	1
Ω	2
g	0.02
α	0.95
$w_i(t_0)$	$1 \quad \forall 1 \leq i \leq N$
$q_i(t_0)$	$1 \quad \forall 1 \leq i \leq N$
N	2500
λ	0.25
$S(t=0)$	1
$q_i(t=0)$	$1 \quad \forall 1 \leq i \leq N$
time steps	$10,000$

(a) Parameters in the Harras model.

Variable	Initial Value
$k_{ij}(t=0)$	$\mathrm{Unif_c}(0,1)$
$E_i[\sigma_j(t=0)]$	$\mathrm{Unif_d}(\{-1,0,1\})$
$\sigma^2_{ED}(t=0)$	0.1
$\langle ED(t=0)\rangle$	0
$ED(t=0)$	0
$ED(t=-1)$	0
$u_i(t=-1)$	0
$u_i(t=0)$	0
$v_i(t=0)$	0
$\sigma_i(t=0)$	0

(b) Initial values Harras.

Table 7: Harras basic setting.

Parameter	Value
ϕ	1.0
χ	1.2
η	0.991
α_w	1580
α_0	0
α_h	0
α_m	0
σ_f	0.681
σ_c	1.724
β	1
ν	–
P_f	1
μ	0.01
time steps	$7,000$

(a) Parameters for DCA-W.

Variable	Initial Value
$P(t=0)$	1

(b) Initial values DCA-W.

Table 8: DCA-W.

Parameter	Value
ϕ	1.0
χ	0.9
η	0.987
α_w	2668
α_0	2.1
α_h	0
α_m	0
σ_f	0.752
σ_c	1.726
β	1
ν	–
P_f	1
μ	0.01
time steps	$7,000$

(a) Parameters for DCA-WP.

Variable	Initial Value
$P(t=0)$	1

(b) Initial values DCA-WP.

Table 9: DCA-WP.

Parameter	Value
ϕ	1.0
χ	0.9
η	0.987
α_w	2668
α_0	2.1
α_h	1.28
α_m	0
σ_f	0.741
σ_c	2.087
β	1
ν	–
P_f	1
μ	0.01
time steps	$7,000$

(a) Parameters for DCA-WHP.

Variable	Initial Value
$P(t=0)$	1

(b) Initial values DCA-WHP.

Table 10: DCA-WHP.

Parameter	Value
ϕ	0.12
χ	1.5
η	0
α_w	0
α_0	-0.327
α_h	1.79
α_m	18.43
σ_f	0.758
σ_c	2.087
β	1
ν	–
P_f	1
μ	0.01
time steps	$7,000$
random seed	2661

(a) Parameters for DCA-HPM.

Variable	Initial Value
$P(t=0)$	1

(b) Initial values DCA-HPM.

Table 11: DCA-HPM.

Parameter	Value
ϕ	1.15
χ	0.81
η	0.987
α_w	1041
α_0	0
α_h	0
α_m	0
σ_f	0.715
σ_c	1.528
β	–
ν	0.05
P_f	1
μ	0.01
time steps	$7,000$

(a) Parameters for TPA-W.

Variable	Initial Value
$P(t=0)$	1

(b) Initial values TPA-W.

Table 12: TPA-W.

Parameter	Value
ϕ	1.0
χ	0.83
η	0.987
α_w	2668
α_0	0.376
α_h	0
α_m	0
σ_f	0.736
σ_c	1.636
β	–
ν	0.05
P_f	1
μ	0.01
time steps	$7,000$

(a) Parameters for TPA-WP.

Variable	Initial Value
$P(t=0)$	1

(b) Initial values TPA-WP.

Table 13: TPA-WP.

Parameter	Value
ϕ	0.18
χ	2.3
η	0.987
α_w	0
α_0	-0.161
α_h	1.3
α_m	12.5
σ_f	0.79
σ_c	1.9
β	$-$
ν	0.05
P_f	1
μ	0.01
time steps	$7,000$
random seed	2661

(a) Parameters for TPA-HPM.

Variable	Initial Value
$P(t=0)$	1

(b) Initial values TPA-HPM.

Table 14: TPA-HPM.

References

Sameera Abar, Georgios K Theodoropoulos, Pierre Lemarinier, and Gregory MP O'Hare. Agent based modelling and simulation tools: a review of the state-of-the-art software. *Computer Science Review*, 24:13–33, 2017.

Maximilian Beikirch, Simon Cramer, Martin Frank, Philipp Joachim Otte, Emma Pabich, and Torsten Trimborn. Dataset for simulation of stylized facts in agent-based computational economic market models. URL `https://publications.rwth-aachen.de/record/749924`.

Avraham Beja and M Barry Goldman. On the dynamic behavior of prices in disequilibrium. *The Journal of Finance*, 35(2):235–248, 1980.

Josef Brada, Harry Ernst, and John Van Tassel. Letter to the editor—the distribution of stock price differences: Gaussian after all? *Operations Research*, 14(2):334–340, 1966.

William A Brock and Cars H Hommes. A rational route to randomness. *Econometrica: Journal of the Econometric Society*, pages 1059–1095, 1997.

John Y Campbell, Andrew Wen-Chuan Lo, Archie Craig MacKinlay, et al. *The econometrics of financial markets*, volume 2. princeton University press Princeton, NJ, 1997.

D. Challet and M. Marslii. Criticality and finite size effects in a simple realistic model of stock market. *Working Paper Department of Theoretical Physics, Oxford University, INFM, Trieste-SISSA Unit, Trieste*, 2002.

Shu-Heng Chen, Chia-Ling Chang, and Ye-Rong Du. Agent-based economic models and econometrics. *The Knowledge Engineering Review*, 27(02):187–219, 2012.

Carl Chiarella, Roberto Dieci, and Laura Gardini. Asset price and wealth dynamics in a financial market with heterogeneous agents. *Journal of Economic Dynamics and Control*, 30(9):1755–1786, 2006.

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.

Robin Cowan and Nicolas Jonard. *Heterogenous agents, interactions and economic performance*, volume 521. Springer Science & Business Media, 2002.

Rod Cross, Michael Grinfeld, Harbir Lamba, and Tim Seaman. A threshold model of investor psychology. *Physica A: Statistical Mechanics and its Applications*, 354:463–478, 2005.

Rod Cross, Michael Grinfeld, Harbir Lamba, and Tim Seaman. Stylized facts from a threshold-based heterogeneous agent model. *The European Physical Journal B*, 57(2):213–218, 2007.

E Egenter, T Lux, and D Stauffer. Finite-size effects in monte carlo simulations of two stock market models. *Physica A: Statistical Mechanics and its Applications*, 268(1):250–256, 1999.

Norman Ehrentreich. *Agent-based modeling: The Santa Fe Institute artificial stock market model revisited*, volume 602. Springer Science & Business Media, 2007.

Fama Eugene. The distrigbution of the daily differences of the logarithms of stock prices. In *Ph.D. Thesis*. Graduate School of Business University of Chicago, 1963.

Eugene F Fama. The behavior of stock-market prices. *The journal of Business*, 38(1):34–105, 1965.

J Doyne Farmer and Duncan Foley. The economy needs agent-based modelling. *Nature*, 460 (7256):685–686, 2009.

Reiner Franke. Applying the method of simulated moments to estimate a small agent-based asset pricing model. *Journal of Empirical Finance*, 16(5):804–815, 2009.

Reiner Franke and Frank Westerhoff. Validation of a structural stochastic volatility model of asset pricing. *Christian-Albrechts-Universität zu Kiel. Department of Economics*, 2009.

Reiner Franke and Frank Westerhoff. Estimation of a structural stochastic volatility model of asset pricing. *Computational Economics*, 38(1):53–83, 2011.

Reiner Franke and Frank Westerhoff. Structural stochastic volatility in asset pricing dynamics: Estimation and model contest. *Journal of Economic Dynamics and Control*, 36(8):1193–1211, 2012.

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.

T Hellthaler. The influence of investor number on a microscopic market model. *International Journal of Modern Physics C*, 6(06):845–852, 1996.

Bruce M Hill. A simple general approach to inference about the tail of a distribution. *The annals of statistics*, pages 1163–1174, 1975.

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

Carsien Harm Hommes. Financial markets as nonlinear adaptive evolutionary systems. 2001.

Ernst Ising. Beitrag zur theorie des ferromagnetismus. *Zeitschrift für Physik A Hadrons and Nuclei*, 31(1):253–258, 1925.

Marco A Janssen and Elinor Ostrom. Empirically based, agent-based models. *Ecology and society*, 11(2), 2006.

R Kohl. The influence of the number of different stocks on the levy–levy–solomon model. *International Journal of Modern Physics C*, 8(06):1309–1316, 1997.

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

Blake LeBaron. Agent-based financial markets: Matching stylized facts with style. *Post Walrasian Macroeconomics: Beyond the DSGE Model*, pages 221–235, 2006.

Haim Levy, Moshe Levy, and Sorin Solomon. *Microscopic simulation of financial markets: from investor behavior to market phenomena*. Academic Press, 2000.

Moshe Levy, Haim Levy, and Sorin Solomon. A microscopic model of the stock market: cycles, booms, and crashes. *Economics Letters*, 45(1):103–111, 1994.

Moshe Levy, Haim Levy, and Sorin Solomon. Microscopic simulation of the stock market: the effect of microscopic diversity. *Journal de Physique I*, 5(8):1087–1107, 1995.

Moshe Levy, Nathan Persky, and Sorin Solomon. The complex dynamics of a simple stock market model. *International Journal of High Speed Computing*, 8(01):93–113, 1996.

Thomas Lux. Herd behaviour, bubbles and crashes. *The economic journal*, pages 881–896, 1995.

Thomas Lux. The socio-economic dynamics of speculative markets: interacting agents, chaos, and the fat tails of return distributions. *Journal of Economic Behavior & Organization*, 33 (2):143–165, 1998.

Thomas Lux. Stochastic behavioral asset pricing models and the stylized facts. Technical report, Economics working paper/Christian-Albrechts-Universität Kiel, Department of Economics, 2008.

Thomas Lux and Michele Marchesi. Scaling and criticality in a stochastic multi-agent model of a financial market. *Nature*, 397(6719):498–500, 1999.

Thomas Lux et al. Applications of statistical physics in finance and economics. Technical report, Kiel working paper, 2008.

Dario Maldarella and Lorenzo Pareschi. Kinetic models for socio-economic dynamics of speculative markets. *Physica A: Statistical Mechanics and its Applications*, 391(3):715–730, 2012.

Benoit B Mandelbrot. The variation of certain speculative prices. In *Fractals and scaling in finance*, pages 371–418. Springer, 1997.

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

Vilfredo Pareto. Cours d'économie politique, professé à l'université de lausanne. tome second, 1897.

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

Herbert A Simon. A behavioral model of rational choice. *The quarterly journal of economics*, pages 99–118, 1955.

Didier Sornette. Physics and financial economics (1776–2014): puzzles, ising and agent-based models. *Reports on Progress in Physics*, 77(6):062001, 2014.

Leigh Tesfatsion. Agent-based computational economics: Growing economies from the bottom up. *Artificial life*, 8(1):55–82, 2002.

T. Trimborn, P. Otte, S. Cramer, M. Beikirch, E. Pabich, and M. Frank. SABCEMM-A Simulator for Agent-Based Computational Economic Market Models. *ArXiv e-prints*, January 2018a.

T. Trimborn, P. Otte, S. Cramer, M. Beikirch, E. Pabich, and M. Frank. Simulator for agent based computational economic market models (sabcemm). `https://github.com/SABCEMM/SABCEMM`, 2018b.

Torsten Trimborn, Philipp Otte, Simon Cramer, Max Beikirch, Emma Pabich, and Martin Frank. Data set of sabcemm - a simulation framework for agent-based computational economic market models. `http://dx.doi.org/10.18154/RWTH-2017-09142`.

Wolfgang Weidlich and Günter Haag. *Concepts and models of a quantitative sociology: the dynamics of interacting populations*, volume 14. Springer Science & Business Media, 2012.

Elmar Zschischang and Thomas Lux. Some new results on the levy, levy and solomon microscopic stock market model. *Physica A: Statistical Mechanics and its Applications*, 291(1): 563–573, 2001.