Agent-Based Modelling
Effects of Trader Circulation in an Artificial Stock Market

Master Thesis
Submitted in Partial Fulfillment of the Requirements for the Degree of
MASTER OF SCIENCE
in Banking and Finance

Supervised by: Michael Razen, PhD
Department of Banking and Finance
at the University of Innsbruck School of Management

Submitted by

Sebastian SCHIEFER, BSc
Innsbruck, October 2018
Declaration of Academic Honesty

I hereby declare that this master’s thesis has been written only by the undersigned and without any assistance from third parties. I confirm that no sources have been used in the preparation of this thesis other than those indicated in the thesis itself.

This master’s thesis has heretofore not been submitted or published elsewhere, neither in its present form, nor in a similar version.

Date: __________________________ Signature: __________________________
Abstract / Zusammenfassung

Abstract

This thesis deals with the agent-based model by Lux and Marchesi (2000), which simulates an artificial stock market where two different trading groups trade a single asset. The original model is adapted in order to tackle the question what effects traders leaving and entering the market have on the price and return series as well as on the market composition of the market. It is shown that by relaxing some of the unrealistic assumptions of the original model, volatility is induced into the market by the circulation of the traders.

Zusammenfassung

1 Motivation

The agent-based modelling approach is a relatively new form of modelling financial markets. Although the idea of this research form has been around for quite some time and is not only used in economics, see for example the famous segregation model by Schelling in the late 1970s, it was due to the advances in computational power and the rise and insights of behavioural finance, that gained the agent-based modelling approach a significant momentum boost. In the financial literature these models are used to break up the traditional convention of a representative agent, i.e. a fully rational agent that is in place of all market participants (Müller and Pyka, 2017). However, a whole branch of economics, the so called behavioural economics, is dedicated to the irrationalities of people and show that we are not always making the theoretically best choice but rather adhere to biases and heuristics (see Tversky and Kahneman, 1974).

Several crisis in the past have shown that ordinary models do not provide a satisfying rationale to answer the questions that emerged with the “irrational exuberance”, a phrase coined by Ex-Federal Reserve Board chairman Alan Greenspan during the dot-com bubble. Taking for example the Chinese Warrants Bubble in 2005-2008, Xiong and Yu (2011) have shown that newly arriving investors contributed largely to excessive price volatility. The impact of traders leaving and entering such a market could not be observed with a traditional model and a representative agent. Even
worse, whenever a fully rational agent wants to trade with another fully rational trader, no agreement would be reached. This is due to the fact that the counterparty always assumes that the supplier has superior information and that he would be worse off. These no-trade situations can not be explained with representative equilibrium models (Hommes, 2008). Agent-based models take the bounded rationality and heterogeneity of agents into account and therefore provide a more realistic setup and try to overcome the shortcomings of the traditional models. That is, the market is modelled from a bottom-up perspective where the interactions of various heterogeneous agents on the micro-level emerge to a more complex behaviour on the aggregate (Epstein and Axtell, 1996). Using an agent-based approach yields numerous benefits over a human subject experiment. Most importantly, a simulation can be repeated several times and at lower costs with a greater control of the parameters (Duffy, 2006).

The purpose of this master thesis is to examine the effects of agents on the asset price, the return and the market composition when leaving and entering an artificial stock market.

Therefore the work at hand is subdivided into 4 chapters. Chapter 2 on the following page starts with a paper by Lux and Marchesi and builds the foundation of this master thesis. A detailed summary of the characteristics of the model and the results are provided. Chapter 3 on page 19 deals with the recreation and validation of the results of the original model. Furthermore, in order to fulfil the purpose mentioned above, the adjustments to the model are presented. Subsequently in Chapter 4 on page 34 the results of the adjustments are compared with data from the original model and the differences are highlighted and interpreted. Chapter 5 on page 51 concludes the work at hand and gives possible directions for future research.
2 The Lux-Marchesi-Model

This part of the thesis deals with the seminal paper of Lux and Marchesi (2000), which was published in the International Journal of Theoretical and Applied Finance. It builds on previously published papers by Lux (1995, 1998). Their paper builds the foundation for this master thesis and their model is the main focus of attention and will be adjusted in the subsequent chapter.

In this chapter an in-depth explanation of the original model is provided in Section 2.1, after that in Section 2.2 on page 10 the simulation process is explained. This chapter concludes with a discussion in Section 2.3 on page 14 about the results and their importance.

2.1 The Model

Lux and Marchesi intended to provide a possible explanation for the findings of clustered volatility and ARCH-effects in financial time series data, since there is still no satisfying rationale for them. They focused on three properties of financial time series, that can be found in various sectors of the market, ranging from prices and returns of a stock market, to commodity futures and prices of precious metals:

(i) Asset prices or logs of prices follow a unit root process.

(ii) Raw returns are unpredictable and exhibit no significant autocorrelation.
(iii) Returns do not follow a standard normal distribution, but a power law in the outer parts, resulting in fat tails.

(iv) Large deviations in returns tend to be followed by large deviations again, and small deviations tend to be followed by small deviations, yielding volatility clusters.

While traditional models, using a representative agent, are able to explain the stylized facts (i) and (ii), they have difficulties with explaining facts (iii) and (iv) (see Hommes, 2006). By using a multi-agent framework, which builds on previously published papers by Lux (1995, 1998), they provide a possible explanation for all of the aforementioned stylized facts of financial time series.

The basic idea of the model consists of a predefined amount of speculators \(N\) who can be distinguished into two different types of traders, fundamentalists \(n_f\) and chartists \(n_c\), each of them trading a single asset in a simulated market. The chartists can be further divided into optimistic \(n_{co}\) and pessimistic \(n_{cp}\) traders, depending on their view of the future development of the market, resulting in the relation \(n_c = n_{co} + n_{cp}\). Note that the labels of the variables in the work at hand deviate from the original denotations in Lux and Marchesi (2000), for the sake of an easier understanding and readability.

This setup of chartists and fundamentalists in agent-based models has already brought forth interesting insights in financial market dynamics. A paper by Joshi et al. (2000) has shown that the use of technical trading can arise due to a prisoner’s dilemma. Papers by Taylor and Allen (1992), Cheung and Chinn (2001) and Menkhoff and Schmidt (2005) support the existence of chartism empirically, despite the non-conformity with the Efficient Market Hypothesis.

During the simulation process, the amount of traders \(N\) stays constant over time.
However, the composition of $N = n_f + n_c$ is allowed to vary over time, due to the set of rules for every type of trader which will be further discussed in the next paragraph.

One of the dynamics of the model stems from the alternation of the trading strategies. Chartists are allowed to switch between an optimistic and a pessimistic behaviour. Lux and Marchesi model this behaviour by using utility functions. In the case of chartists, the function contains a term that represents the opinion of other traders, simulating herd behaviour. This assumption is not far off from reality and there are various reasons for the herding effect. One of the earliest papers dealing with this subject, and a possible explanation for the effect, is the paper of Banerjee (1992). The author has shown that taking into account the information of others is rational for the individual person, because they might have additional information that the person is missing out. The decision makers rather choose the same strategy as others than using their own strategy. This results in the aforementioned herding behaviour. However, Banerjee also shows that the resulting equilibrium is inefficient. The existence of herding behaviour is also backed up by empirical studies such as by Chiang and Zheng (2010), who show that there is strong evidence for herding behaviour in almost all stock markets around the world.

The opinion index is given by Equation (2.1) and is the difference between optimistic and pessimistic chartists divided by the total amount of technical traders.

$$x = \frac{n_{co} - n_{cp}}{n_c}, x \in [-1, 1]$$ (2.1)

The second term that influences the utility function is the actual price trend in the form of the price change in continuous time given by $\dot{p} = \frac{dp}{dt}$.

The probability of agents switching between a bullish and a bearish behaviour can
therefore be written as in Equation (2.2).

\[
\pi_{p \rightarrow o} = \nu_1 \left( \frac{n_c}{N} \cdot \exp(U_1) \right) \\
\pi_{o \rightarrow p} = \nu_1 \left( \frac{n_c}{N} \cdot \exp(-U_1) \right)
\]

where \( U_1 = \alpha_1 x + \alpha_2 \frac{\dot{p}}{\nu_1} \) (2.2)

The parameter \( \nu_1 \) represents the time frequency of revaluation of the traders opinion. The parameters \( \alpha_1 \) and \( \alpha_2 \) are weights, measuring the importance of the respective factors of the utility function. If these factors contradict each other, the incentive to follow the majority is weaker than it is, if both factors, the opinion index and the price trend, are in conformity. Therefore it is much likelier that an optimistic agent changes his behaviour, once there are currently more pessimistic traders in the market and the price shows a downward-trend. The opposite is the case for the change of a pessimistic to an optimistic chartist. Hence, a bullish (bearish) agent buys (sells) whenever he believes the future development of the market to be on the rise (fall).

The profits of technical traders are calculated as excess profits compared to an alternative investment given by \((d + \dot{p})/p - r\), with \(d\) being the dividend payment of the traded asset, \(\dot{p}\) as mentioned above as the price change, \(p\) as the actual price and \(r\) being the return from an alternative investment. However, this only holds for optimistic traders as they continue to extend their portfolio. On the other hand, pessimistic chartists try to avoid potential losses by looking for alternative investment opportunities. The respective profit is therefore \(r - (d + \dot{p})/p\).

Furthermore agents are allowed to switch between a chartist strategy and a fundamentalist strategy. Unlike their technical counterparts, fundamentalists buy (sell) whenever they believe the asset to be undervalued (overvalued). Additionally, they assume that in the long run, the prices will revert to the fundamental price \(p_f\). The
probabilities of fundamentalists switching between the subgroups of chartists are given in Equation (2.3). Note that the first part of the utility function includes the profit of the respective chartist subgroup in order to compare the excess profits of both strategies.

\[
\pi_{f \rightarrow o} = \nu_2 \left( \frac{n_f}{N} \cdot \exp(U_2) \right) \quad \text{where} \quad U_2 = \alpha_3 \left( \left( \frac{d + \hat{p}}{\nu_2} \right)/p - r - s \left| \frac{p_f - p}{p} \right| \right)
\]

\[
\pi_{o \rightarrow f} = \nu_2 \left( \frac{n_o}{N} \cdot \exp(-U_2) \right)
\]

\[
\pi_{f \rightarrow p} = \nu_2 \left( \frac{n_f}{N} \cdot \exp(U_3) \right) \quad \text{where} \quad U_3 = \alpha_3 \left( r - \left( \frac{d + \hat{p}}{\nu_2} \right)/p - s \left| \frac{p_f - p}{p} \right| \right)
\]

\[
\pi_{p \rightarrow f} = \nu_2 \left( \frac{n_p}{N} \cdot \exp(-U_3) \right)
\]

Again, \( \nu_2 \) is a parameter for the revaluation frequency of the agent and \( \alpha_3 \) is the weight given to the difference between the profits of the chartist and fundamentalist strategy. Since fundamentalists believe that the price reverts to its fundamental value, their profit can be written as \( s \left| \frac{p_f - p}{p} \right| \). This holds, irrespective of whether the asset is currently overpriced or underpriced. Unlike chartists, whose profits are realized immediately, the profits of fundamentalists have to be discounted with the factor \( s \), since the potential gains are realized only after the price has reverted to its fundamental value.

The price formation process is managed by an auctioneer, based on demand and supply. Once the auctioneer senses an imbalance between demand and supply, he adjust the current price by one cent up (down) within one unit of time. Similarly to above, this is modelled by transition probabilities given in Equation (2.4).

\[
\pi_{tp} = \max \left[ 0, \beta \left( ED + \epsilon \right) \right]
\]

\[
\pi_{ip} = -\min \left[ 0, \beta \left( ED + \epsilon \right) \right]
\]
In this case $\beta$ is the reaction speed of the market maker regarding the imbalances in the market. The variable $ED$, representing aggregated excess demand, consists of the respective surplus of chartists and fundamentalists and can therefore be formalized as $ED = ED_c + ED_f$. Furthermore, a noise term $\epsilon$ is added to simulate the imprecision of the auctioneer in assessing excess demand. This noise term is assumed to follow a Normal distribution with mean equals to $\mu$ and standard deviation equals to $\sigma$. The excess demand of chartists is given by $ED_c = (n_{co} - n_{cp})t_c$, where $t_c$ is the fixed amount of shares that all chartists buy or sell. On the other hand, excess demand for fundamentalists can be written as $ED_f = n_f\gamma(p_f - p)$, where $\gamma$ is the reaction strength to the deviation of the fundamental price to the current price.

The authors further investigated the theoretical results of the model, by deriving differential equations for the opinion index $x$, the chartist index $z = n_c/N$ and the price $p$. They pin down three stationary solutions: (i) the price is equals to the fundamental price and there is no dissension about where the market is moving; (ii) fundamentalists are driven out of the market and there is, again, a balance in the market opinion; (iii) technical traders are driven out of the market and the price is equals to the fundamental price. However, in the simulation the equilibria (ii) and (iii) are avoided by adding additional constraints in the agent distribution process, explained in Section 2.2 on page 10, since they are absorbing states and are not interesting for the development of the simulation.

Furthermore they defined conditions for the instability of equilibrium (i) in Equation (2.5) on the following page.
\[ \alpha_1 > 1 + \alpha_3 \frac{\nu_2 T_r}{\nu_1 T_f p_f} \]

or

\[ 2 \cdot z_{\text{max}} \nu_1 \left( \alpha_1 + \alpha_2 \frac{\beta}{\nu_1} z_{\text{max}} T_c - 1 \right) + 2 \cdot (1 - z_{\text{max}}) \alpha_3 \beta z_{\text{max}} \frac{T_c}{p_f} - \beta (1 - z_{\text{max}}) T_f > 0 \]

This can be interpreted as follows: Once the weight given to the opinion index reaches a certain threshold, i.e. the incentive to follow the crowd is too high, equilibrium (i) is not stable any more.

It is therefore possible to calculate a range of values for the chartist index, where values below \( z_{\text{max}} \) define the area of stable equilibria, and for values above \( z_{\text{max}} \) the equilibria are unstable. The exact threshold for \( z_{\text{max}} \) can be calculated by solving the second part of Equation (2.5) for \( z_{\text{max}} \). The solution to this quadratic equation is given in Equation (2.6).

\[ a = 2 T_c \beta \alpha_2 - \frac{(2 T_f \beta \alpha_3)}{p_f} \]

\[ z_{\text{max}} = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{where} \quad b = T_f \beta - 2 \nu_1 + 2 \alpha_1 \nu_1 + \frac{2 T_c \beta \alpha_3}{p_f} \]

\[ c = -T_f \beta \]
2.2 The Simulation Process

In this section the individual steps of the simulation are explained. Figure 2.1 on the next page depicts the sequence of the simulation process. At first, the user sets the variables, such as the number of observations, the fundamental price or the reaction speed of the price adjustment through the auctioneer. Table 2.1 presents one of the original parameter sets used in Lux and Marchesi (2000). Unfortunately, there are no relevant empirical observations for, e.g. the frequencies of opinion revaluation or the weights given to trend following, that could have been used for the specific parameters. However, the authors claim that the results of the simulation are robust to a variety of parameter sets and that the main findings are unchanging. The parameter set below is fine-tuned in a way that the simulated returns approximately match the observable returns on an asset market or a foreign exchange market at a daily frequency, i.e. simulated absolute returns do not exceed 20%-30% in most cases.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N$</td>
<td>Number of agents</td>
<td>500</td>
</tr>
<tr>
<td>$p_f$</td>
<td>Fundamental price</td>
<td>10.0</td>
</tr>
<tr>
<td>$d$</td>
<td>Dividend payment</td>
<td>0.004</td>
</tr>
<tr>
<td>$r$</td>
<td>Return of alternative investment</td>
<td>0.0004</td>
</tr>
<tr>
<td>$s$</td>
<td>Discount factor</td>
<td>0.75</td>
</tr>
<tr>
<td>$\nu_1$</td>
<td>Time frequency of optimist/pessimist revaluation</td>
<td>3.0</td>
</tr>
<tr>
<td>$\nu_2$</td>
<td>Time frequency of chartist/fundamentalist revaluation</td>
<td>2.0</td>
</tr>
<tr>
<td>$\alpha_1$</td>
<td>Weight given to opinion index x</td>
<td>0.6</td>
</tr>
<tr>
<td>$\alpha_2$</td>
<td>Weight given to price trend</td>
<td>0.2</td>
</tr>
<tr>
<td>$\alpha_3$</td>
<td>Weight given to price differential</td>
<td>0.5</td>
</tr>
<tr>
<td>$T_c = Nt_c$</td>
<td>Trading volume of chartists</td>
<td>10.0</td>
</tr>
<tr>
<td>$T_f = N\gamma$</td>
<td>Trading volume of fundamentalists</td>
<td>5.0</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Reaction speed of auctioneer</td>
<td>6.0</td>
</tr>
<tr>
<td>$\epsilon \sim N(\mu, \sigma)$</td>
<td>Imprecision of auctioneer</td>
<td>$N(0, 0.05)$</td>
</tr>
</tbody>
</table>

Table 2.1: Original parameter set of Lux and Marchesi (2000)
CHAPTER 2. THE LUX-MARCHESI-MODEL

Figure 2.1: Flowchart for the simulation of Lux and Marchesi (2000)
Once the simulation starts, agents are randomly distributed among the different groups of traders. The number of chartists is chosen sufficiently small in order to ensure that the simulation starts at a stable equilibrium, i.e. the fraction of technical traders is below $z_{\text{max}}$. Furthermore, to exclude equilibria of type (ii) and (iii), a lower limit to both types of agents is set. That means that switching between strategies is only allowed if there are more than $n_{\text{min}}$ traders in each strategy. The lower limit $n_{\text{min}}$ is arbitrarily set to four by the authors. The distribution of agents can therefore be formalized as in Equation (2.7).

$$
\begin{align*}
  n_c & = \max \left[ N \cdot z_{\text{max}} \cdot \xi, 2 \cdot n_{\text{min}} \right] \quad \text{where } \xi \sim U([0,1]) \\
  n_{co} & = \min \left[ \max \left[ n_c \cdot \xi, n_{\text{min}} \right], n_c - n_{\text{min}} \right] \\
  n_{cp} & = n_c - n_{co} \\
  n_f & = N - n_c
\end{align*}
$$

At first the total amount of chartists is randomly allocated. Since there have to be at least $n_{\text{min}}$ traders in the optimistic and pessimistic group, the lower limit is $2 \cdot n_{\text{min}}$. In order to allow the simulation to start at a stable state, the upper limit is calculated by multiplying the total amount of agents $N$ with the threshold $z_{\text{max}}$ and a uniformly distributed random variable $\xi$ on the interval $[0, 1]$.

Chartists are then split up in optimistic and pessimistic traders in a similar fashion. There are never less than $n_{\text{min}}$ agents in the bullish (bearish) group and there can never be more than $n_c - n_{\text{min}}$ in total. Note that the calculation of $n_{co}$ and $n_{cp}$ are interchangeable.

Once the chartists are partitioned, the amount of fundamentalists results from the difference of total agents and chartists.
Subsequently, the actual simulation starts with the outer loop. The outer loop can be thought of as “trading days” and simulates the predefined number of return observations. Within this loop the inner loop represents micro time steps of this “day”, e.g. minutes. In their paper Lux and Marchesi used a time increment of 0.01, for what they called “normal times”, but decreased the time increment to 0.002 during high volatile sequences. The reason for this is, that they did not want to restrict the price changes that happen during these time steps. Also, computation power has to be taken into consideration. In this micro interval, agents form an opinion about the market development, change their strategy, buy or sell the asset and therefore influence the price development, while the market maker adjusts the price based on supply and demand. It is therefore necessary, that all transition probabilities, i.e. the probabilities that agents switch their strategy and the price adjustment probability, are divided by this micro interval. The actual change of behaviour occurs by comparing a uniformly distributed random variable $\xi$ on the interval $[0, 1]$ to the specific transition probability divided by the amount of micro time steps. If $\xi$ happens to be larger than the relevant transition probability, the trader stays in his current strategy. However, if it is smaller than the randomly generated number, the strategy change occurred.

In order to avoid the absorbing states (ii) and (iii), as well as the critical value of chartists $z_{\text{max}}$, the total amount of traders is checked for these boundaries. For example, if the number of optimistic chartists minus the sum of chartists switching to fundamentalistic or pessimistic strategies is smaller than $n_{\text{min}}$, the portion of outgoing agents is adjusted, so that there are at least $n_{\text{min}}$ optimistic traders. The same principle applies to all strategies. On the other hand, to prevent the simulation from producing unstable equilibria of type (i), the maximum amount of chartists is checked. If during the simulation the relation $z < z_{\text{max}}$ is violated, optimistic and
pessimistic traders are adjusted in order to meet this requirement again.
In the next step, all relevant variables are updated. This includes the chartist index
$z$, the opinion index $x$ as well as the specific demand variables and thus also the
price adjustments. Because the auctioneer is modelled in a way where he adjusts
the price $\Delta p = -0.01, 0$ or $+0.01$, the price change $\dot{p}$ is restricted to this very same
set of possibilities. This is avoided by calculating $\dot{p}$ with respect to the price changes
during the interval $[t - 0.2, t)$. The price change connotes the final step of the inner
loop. The above discussed procedure is repeated for the predefined number of micro
time steps. After that the log returns are calculated by $ret_t = \log \left( \frac{p_t}{p_{t-1}} \right)$ and the
outer loop starts again. After all loops have finished the price development, the
returns, the chartist index as well as the opinion index are collected and visualized
in graphs. The simulation ends with this step.

2.3 The Results

This section deals with the analysis of the results of their simulated market and
explains their importance. As mentioned in Section 2.1 on page 3, the authors
aimed to provide possible explanations for several stylized facts in financial time
series data, such as that (i) prices follow a random walk, (ii) returns are stationary
and show no autocorrelation, (iii) the distribution of returns shows fat tails and (iv)
returns exhibit volatility clustering.

The random walk property of the simulated asset prices is checked with the Dickey-
Fuller test in the form of:

$$H_0 : \rho = 1 \ vs. \ H_1 : \rho < 1 \ in \ the \ regression: \ p_t = \beta_0 + \rho p_{t-1} + \epsilon_t \ \ (2.8)$$
That means if $\rho$ happens to be close to or equals to 1, one is unable to reject the null hypothesis of a unit root. The time series is therefore non-stationary and follows a random walk. This is usually the case for real data sets and stock prices. The authors simulated 20,000 observations and split them in 40 subgroups of 500 observations each. Furthermore they applied the Dickey-Fuller test on 4 different parameter sets, resulting in a total of 160 tests. None of these one-sided tests resulted in a rejection of the unit root in the 95% confidence interval. The results are robust to sample size and also yield the same results when using the Augmented-Dickey-Fuller test. It is therefore save to say, that the simulated asset prices follow a random walk. This is an important result, because it goes along with the stylized fact (i) and the characteristics of real life time series data.

In Figure 2.2 on the next page some of the remaining stylized facts (ii)-(iv) can be already visually inspected. The upper part of the graph shows the simulated return series of the model. It can be seen that returns fluctuate around zero for most of the time, indicating a stationary time series. However, these periods are interrupted by large deviations where returns reach unusual values of more than 15% which appear to happen in clusters. This can be especially seen in the periods from $t = 3100$ to $t = 3400$ and suggests the existence of volatility clusters in this generated return series.

The authors then compare the simulated return graph with the development of the chartist index in order to see if the destabilization of the system is brought about by the number of chartists ($z = \frac{n_c}{N}$), as is suggested by the theoretical results in Equation (2.5) and Equation (2.6) on page 9. In the lower part of Figure 2.2, the threshold of destabilization $z_{\text{max}}$ is indicated by a broken line.
CHAPTER 2. THE LUX-MARCHESI-MODEL

Figure 2.2: The upper part depicts the simulated return series. The lower part illustrates the development of the proportion of chartists and fundamentalists by drawing $z = \frac{N_c}{N}$ (taken from Lux and Marchesi, 2000).

Whenever the fraction of chartists $z$ gets close to this line, the price changes start to fluctuate heavily and therefore generate large deviations in returns. On the other
hand, if \( z \) stays away from \( z_{\text{max}} \), returns float around zero. The large deviations can be explained by the behaviour of chartists. The divergence from the equilibrium becomes self-reinforcing through the herding behaviour and following the price trend, once there is a certain prevalence of technical traders. However, these phases are stabilized by the large excess returns of the fundamentalistic strategy, causing the chartists to switch their behaviour and to calm the market again. The authors conclude, that the volatility clusters stem from the interaction of traders and their behaviour, therefore proving stylized fact (iv) to be evident in this computer generated time series. This argument is further strengthened by analysing the the autocorrelation functions of raw returns, squared returns and absolute returns. Figure 2.3a illustrates a typical autocorrelation function for the simulated returns. While raw returns float around zero, indicating no long time memory in returns and therefore no predictability for the future, squared and absolute returns behave differently. Even for 300 lags, the autocorrelation function of squared and absolute

---

\(^1\)Daily returns from 17.10.1997 to 30.08.2017, resulting in 5000 daily return observations. Data retrieved from https://finance.yahoo.com/quote/%5Egspc
returns only slowly decays to zero. Absolute returns even exhibit significant autocorrelation for lags larger than 300, indicating the presence of volatility clusters. This is in line with real life observations as can be seen in Figure 2.3b on the preceding page where the autocorrelation function of 5000 daily return observations of the S&P 500 are plotted. The graph shows the same characteristics, with raw returns exhibiting no significant autocorrelation, while squared and absolute returns do and thus prove stylized fact (iv) to be evident in the simulation.

At last, the authors investigate whether returns do have more probability mass in their tails to address stylized fact (iii). Again, the authors examine their simulated data set of 20,000 observations from 4 different parameter sets and compare their distribution parameters. In all cases the kurtosis is much larger than 15, compared to the fourth moment of a normal distribution, which is 3. To confirm the existence of fat tails in the returns, tail index estimates are calculated. The Hill index estimation calculates the frequency of extreme events, such as extraordinarily high or low returns, in a predefined region of the tails. Again, the computer generated time series yields similar results with tail index estimates in the range of 2 to 5, which are also commonly found in empirically observable data sets. The simulated return series is therefore leptokurtic, which means that extreme events, far left or far right of the distribution, are more likely to happen compared to the normal distribution, and hence confirms stylized fact (iii).

In conclusion, Lux and Marchesi account the emergence of these stylized facts to the interaction of technical and fundamental traders. They provide an alternative explanation for this phenomena and developed a framework for future research.
3 Modifications of Lux-Marchesi

This chapter will be the main part of this master thesis. It deals with the modifications of the original model of Lux and Marchesi in order to analyse the research question aforementioned in Chapter 1 on page 1. The original model was rebuilt in Java and provides a graphical user interface where the most important data and graphs are updated dynamically. The program can be used on every computer that has the Java Runtime Environment (JRE)\(^1\) installed. In Section 3.1 the results of the original study are recreated in order to validate the findings. Following this in Section 3.2 on page 30 are the contributions the author of the work at hand has made to the model.

3.1 Validation

In order to ensure that the rebuilt model is correctly programmed, the results of the original model are replicated. The only noteworthy change to the original model is the amount of micro time steps in the inner loop. Lux and Marchesi restricted themselves to a time increment of 0.01 and only decreased it to 0.002 during high volatile sequences. Since computers nowadays are much more potent than they were 20 years ago, the micro time steps are permanently set to 0.002. This means that there are 500 iterations through the inner loop for every observation. A simulation

\(^1\)The Java software is downloadable for free under: https://java.com/download/
CHAPTER 3. MODIFICATIONS OF LUX-MARCHESI

with 20,000 observations therefore produces 10,000,000 repetitions which approximately takes 7-8 minutes to simulate.

The rest of the program should be identical and therefore produce similar results. The next step is to feed the Java program with one of the parameter sets, analyse the output and compare it to the results of Lux and Marchesi. The following results are independent of the chosen parameter set, therefore Parameter set I is arbitrarily chosen for this purpose.

3.1.1 Unit root property of Prices

The first part in replicating the results of Lux and Marchesi deals with the unit root property of asset prices. This property backs the conjecture of the famous Efficient Markets Hypothesis, first brought forth by Fama (1970), which is, to some extent widely accepted in the financial literature. In its most efficient form, the Efficient Market Hypothesis suggests that every information is reflected in the price of a stock. Therefore the price changes of any future date will only reflect the news at that certain point of time. As the name already implies, news are unpredictable and therefore price changes in the future are also unpredictable and random. This summarizes the Random Walk Hypothesis that derives from the Efficient Market Hypothesis. Expressed in a more mathematical form, the price at time \( p_t \) follows an autoregressive process with random increments \( e_t \). The price development can be expressed with the following equation:

\[
p_t = p_{t-1} + e_t
\]  

(3.1)

There is a strand of literature testing this hypothesis with mixed results. A study by Borges (2009) for example analyses the stock markets of Germany, France, United
Kingdom, Spain, Portugal and Greece, and shows that there is indeed support for the Random Walk Hypothesis. However, an earlier study by Malkiel (2003) summarizes some effects that render the Efficient Market Hypothesis invalid and therefore provides feasible arbitrage possibilities which implies that stock prices are partially predictable. Sewell (2011) provides a good summary of the most notable literature dealing with the divergence of opinions and results related to the Efficient Market Hypothesis.

The author of the work at hand follows Lux and Marchesi on their procedure to analyse the simulated time series, i.e. 20,000 observations are simulated and 40 subsamples of 500 observations each are created. As already explained in Section 2.3 on page 14, the unit root property of asset prices can be checked with the Dickey-Fuller test and Equation (2.8) on page 14.

On the first sight the graphs in Figure 3.1 on the following page look similar.
In their paper Lux and Marchesi claim that by visual inspection it is hard to decide whether the time series is stationary or non-stationary. Although it is agreed on
CHAPTER 3. MODIFICATIONS OF LUX-MARCHESI

that it is hard to justify only by visual inspection whether a time series has the
unit root property or not, from examining the graph, however, one could jump to
the conclusion that there are mean-reverting tendencies visible. This can be seen in
Figure 3.2. If one imagines a constant line at the price level of 10 during the whole
time period, it can be seen that the price fluctuates around this line.

![Figure 3.2: Time series with a constant line at the price level of 10.](image)

Prices fluctuating around a constant mean would therefore disconfirm the findings
made by Lux and Marchesi. Table 3.1 on the following page reports the results of
the unit root tests made for all parameter sets. In all 160 cases (4 parameter sets
á 40 subsamples), Lux and Marchesi find no rejection of the unit root property of
prices, independent of the choice of sample size or using logs of prices instead of
levels.
CHAPTER 3. MODIFICATIONS OF LUX-MARCHESI

<table>
<thead>
<tr>
<th>Parameters</th>
<th>range of $\hat{\rho}$</th>
<th>No. of rejections for one-sided test at 95% level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter set I</td>
<td>0.999819–1.000022</td>
<td>0</td>
</tr>
<tr>
<td>Parameter set II</td>
<td>0.999977–1.000021</td>
<td>0</td>
</tr>
<tr>
<td>Parameter set III</td>
<td>0.999957–1.000030</td>
<td>0</td>
</tr>
<tr>
<td>Parameter set IV</td>
<td>0.999972–1.000014</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3.1: Results of the unit root tests reported by Lux and Marchesi.

In Table 3.2 the results of the Dickey-Fuller test from the replicated model are reported. It can be seen that the results differ significantly as the test now rejects the null hypothesis of a unit root in all 40 subsamples of Parameter set I. The range of values for $\hat{\rho}$ are significantly smaller than 1. Again, this holds for all parameter sets. Also when using the Augmented-Dickey-Fuller test given in Equation (3.2) with a lag of 1, all 40 subsamples led to a rejection of the unit root property.

$$H_0 : \gamma = (\rho - 1) = 0 \text{ vs. } H_1 : \gamma = (\rho - 1) < 0$$

in the regression:

$$\Delta p_t = \beta_0 + \gamma p_{t-1} - \sum_{j=1}^{k} \alpha_j \Delta p_{t-j} + \epsilon_t$$  \hspace{1cm} (3.2)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>range of $\hat{\rho}$</th>
<th>No. of rejections for one-sided test at 95% level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter set I (DF)</td>
<td>0.441865 – 0.851797</td>
<td>40</td>
</tr>
<tr>
<td>Parameter set I (ADF)</td>
<td>-0.670146 – -0.123019</td>
<td>40</td>
</tr>
</tbody>
</table>

Table 3.2: Results of the unit root test from the replicated model and using Equation (2.8) and Equation (3.2).
It is therefore save to say that the simulation produces a time series that is in fact, despite the reportings of Lux and Marchesi, stationary, has no unit root and is for that reason not a random walk. Considering the specifications of the original model this finding is very plausible after all because the asset price has a constant fundamental value of 10 throughout the simulation period. Since fundamental traders are always perfectly aware of the constant fundamental price of the asset, they drive prices up (down) when the trading price of the asset is low (high). This explains the mean-reverting pattern of Figure 3.2 on page 23.

However, this is only a minor mistake. If the fundamental price would be modelled as a stochastic process with exogenous news arrival e.g. every trading day rather than being constant over time, the stylized fact of non-stationary prices could be observed again (Lux and Marchesi, 1999). This is shown in Chapter 4 on page 34.

The possible cause of the faulty results of Lux and Marchesi could be due to an incorrect use of the Dickey-Fuller test. Table 3.3 reports the results when regressing without a constant. The values are now close to 1 and 0 for the Dickey-Fuller test, respectively, for the Augmented-Dickey-Fuller test with a lag of 1 and resemble the results obtained by Lux and Marchesi. It seems that the authors erroneously performed their test without a constant, when in fact, there is a constant significantly different to 0, suggested by the constant fundamental price of 10.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>range of $\hat{\rho}$</th>
<th>No. of rejections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter set I (DF)</td>
<td>0.999803 – 1.00006</td>
<td>0</td>
</tr>
<tr>
<td>Parameter set I (ADF)</td>
<td>-0.000190 – 0.000102</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 3.3: Results of the unit root test from the replicated model and using Equation (2.8) and Equation (3.2) but without the constant $\beta_0$. 
3.1.2 Fat-tail distribution of Returns

The next step is to check the results of the replicated model for the fat-tail distribution of returns. A fat-tailed distribution implies that there is more probability mass in the tails compared to the standard Normal distribution. This means that extreme events (excessively positive or negative returns) happen more frequently than the standard Normal distribution would suggest (LeBaron, 2008). Lux and Marchesi report two indicators for this property: the Kurtosis of the return series and the Hill tail index estimator. The results of which are printed in Table 3.4 on the following page. The Hill tail index estimator, introduced in Hill (1975), is a possible solution to the assumption of tails following a power law, approximated by a Pareto distribution of the form $F(x) = 1 - ax^{-\alpha}$. The estimator gives information about the prevalence of anomalously high or low returns. Hills estimator for $\alpha$ is obtained as given in Equation (3.3).

$$\alpha_H = \frac{1}{\frac{1}{k} \sum_{j=1}^{k} (\ln x_{n-i+1} - \ln x_{n-k})}$$

To make use of the aforementioned formula the procedure of Lux and Marchesi in splitting the data into 10 subsamples of 2,000 observations each is followed. Returns are then taken with their absolute value to merge both tail regions. The values are sorted in descending order and the cut-off values $k$ are identical to the ones used in Lux and Marchesi.
<table>
<thead>
<tr>
<th>Parameters</th>
<th>Kurtosis</th>
<th>2.5% tail</th>
<th>5% tail</th>
<th>10% tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter set I</td>
<td>135.73</td>
<td>2.04</td>
<td>2.11</td>
<td>1.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.61-4.50)</td>
<td>(1.51-2.64)</td>
<td>(1.26-2.44)</td>
</tr>
<tr>
<td>Parameter set II</td>
<td>16.1</td>
<td>2.82</td>
<td>2.52</td>
<td>2.18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.28-3.73)</td>
<td>(2.00-3.17)</td>
<td>(1.55-2.36)</td>
</tr>
<tr>
<td>Parameter set III</td>
<td>27.11</td>
<td>4.63</td>
<td>3.48</td>
<td>2.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.41-6.82)</td>
<td>(2.33-8.60)</td>
<td>(1.80-4.84)</td>
</tr>
<tr>
<td>Parameter set IV</td>
<td>37.74</td>
<td>3.08</td>
<td>2.46</td>
<td>1.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.11-4.06)</td>
<td>(2.13-7.86)</td>
<td>(1.65-3.18)</td>
</tr>
</tbody>
</table>

Table 3.4: Fat-tail properties reported by Lux and Marchesi. Kurtosis statistics are reported for all 4 parameter sets with 20,000 observations each. Median $\alpha_H$ tail index estimates are reported from 10 samples of 2,000 observations each (range of estimates in parentheses).

As can be seen in Table 3.5 on the following page, the results of the recreated model do not exhibit the abnormally large fourth moment that is reported by Lux and Marchesi. However, with a kurtosis of 18.95 it is well above the standard normal value of 3, indicating a leptokurtic distribution with fat tails. This result is confirmed by looking at the $\alpha_H$ estimates which are in the range of 1.54 - 4.25, a range which is often found in empirical studies (see for example Lux (1996), LeBaron (2008) or Echaust (2014)). Note that smaller values of $\alpha_H$ indicate fatter tails (see Kirchler and Huber, 2007). Furthermore, the occurrence of fat tails can be visually inspected in a Normal Quantile-Quantile Plot given in Figure 3.3 on the next page.
Table 3.5: Fat-tail properties of the recreated model. Kurtosis is reported for the complete set of 20,000 observations. Median $\alpha_H$ tail index estimates are reported from 10 samples of 2,000 observations each (range of estimates in parentheses).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Kurtosis</th>
<th>2.5% tail</th>
<th>5% tail</th>
<th>10% tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter set I</td>
<td>18.95</td>
<td>3.15</td>
<td>2.65</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.96–4.25)</td>
<td>(1.74–3.22)</td>
<td>(1.54–2.67)</td>
</tr>
</tbody>
</table>

This figure indicates that the simulated return series does not follow a normal distribution. The deviation from the plotted line in the left and right region of the plot are indicators of fat tails. The observed values are either too small or too large to conform to the normal distribution.

Figure 3.3: Normal Q-Q plot of the simulated returns. If returns were normally distributed, the dots would move along the solid line.
3.1.3 Presence of Volatility Clusters

Lastly the presence of volatility clusters in return series are checked. This phenomenon was first discovered by Mandelbrot (1967) and has been discussed in various empirical studies. Volatility clusters are defined as phases were large changes in prices or returns are followed by large changes again and small deviations also by small deviations over time. This results in the variance changing over time (Jacobsen and Dannenburg, 2003). It has been found that while raw returns usually show no correlation, squared returns and absolute returns feature significantly positive and slowly decaying autocorrelation. This is illustrated in Figure 3.4 were the autocorrelation function of raw returns, squared and absolute returns are plotted over 150 lags.

![Autocorrelation of raw, squared and absolute returns](image)

Figure 3.4: Autocorrelation function for raw, squared and absolute returns plotted over 150 lags.

It is evident that raw returns fluctuate around zero and therefore show no significant autocorrelation. This picture changes once squared and absolute returns are inspected. They show significantly higher autocorrelation, with absolute returns
having the highest degree of long term dependency. The autocorrelation for these
two series only decay very slowly to zero and still show positive coefficients even
when looking at 150 lags. As was pointed out earlier in Section 2.3 on page 14 this
confirms to real life observations.

3.2 Adjustments

It was shown that some of the results obtained by Lux and Marchesi were faulty due
to some minor mistakes made in the analysis of the price series or are not as extreme
as they found in their generated return series. Overall, however, the recreated model
developed in Java by the author of the work at hand provides satisfying and similar
results compared to the original model by Lux and Marchesi. This is an important
requirement and a solid starting point in order to make changes to the model itself.
The results obtained by the original model will serve as a benchmark in the following
sections. The development and deployment of the software should also nourish future
research.

3.2.1 Individualization of Traders

Since there is no evidence on how the authors of the original model initially developed
their simulation program, it was assumed that they purely looked at the agents at
the group level. The assumption is backed by the use of the term “multi-agent
framework” or “multi-agent system” in their papers (see Lux and Marchesi, 1999,
2000). This means that there is no room for individualization of the traders in the
model. For example, every agent in the chartist group behaves like any other agent in
the same subgroup. They all share the same characteristics and behaviour. However,
one of the many features of an agent-based model is the ability to look at an agent at
a more detailed level. In the recreated version of the original model an emphasis was put on the individualization capabilities of traders. This allows for more adjustments and a more detailed investigation in the agents decision making. The traders are truly separated unique entities with their own set of parameters. This is a hotbed for future research and builds the basis for the subsequent alterations.

### 3.2.2 Implementation of Wealth

One way to distinguish between successful and unsuccessful traders is by looking at the development of their wealth. This is a feature that was not considered by Lux and Marchesi in their original model. However, it plays an important role in the implementation of the next section. It is assumed that all optimistic chartists and all pessimistic chartists as well as all traders following a fundamental strategy follow the same wealth calculation in their respective subgroup. The wealth calculation is explained in Equation (3.4) in case of a buy order of a chartist.

\[
\begin{align*}
  t = 0 : & \quad -t_c \cdot p_0 \\
  t = 1 : & \quad -t_c \cdot p_0 \cdot (1 + r) + t_c \cdot (p_1 + d) \\
  & = t_c \cdot (\Delta p + d - r) \text{ where: } \Delta p = p_1 - p_0
\end{align*}
\]

The equation can be explained as follows. A buy order (sell order) of chartists is triggered whenever their excess demand \( ED_c = (n_{co} - n_{cp})t_c \) is positive (negative). In cases where the population of optimistic and pessimistic chartists is in balance, no trades are made. In case of a positive excess demand all chartists buy the same amount of shares \( t_c \) and pay the same price \( p_t \). They therefore pay at \( t = 0 \) the total amount of \( t_c \cdot p_0 \). The traders are now in possession of the asset. In the next period, however, they close this position. This can be thought of as precautionary measures since the traders do not want to be exposed to potential contingency risk.
At time $t = 1$ the initial payment for the asset has risen to $t_c \cdot p_0 \cdot (1 + r)$. This term is synonymous with opportunity costs. Nevertheless, the trader is entitled to the proceeds of the asset sale plus the dividend payments and therefore earns $t_c \cdot (p_1 + d)$. The sum of these two terms calculates the net wealth of the agent. Note that this is an example of a chartist buying the asset. For short-selling purposes only the algebraic signs would change and for fundamentalists the amount of shares they buy would change to $\gamma$ instead of $t_c$. This process happens for every micro-time step in the inner loop of the model and for every observation. Put into a more comprehensible frame, the wealth calculation is done every virtual minute of the likewise virtual trading day.

### 3.2.3 Entering and Leaving the Market

One of the goals of this master thesis is to examine the effects of traders leaving and entering the market. The original model by Lux and Marchesi is unrealistic in this regard as it allows traders to accumulate potentially infinite negative balances. Since the model is now able to track the wealth development of the individual trader it allows to rank the agents according to their capital. Furthermore a simple rule is added to the market that resembles a Darwinian process. This process takes care of the inflow and outflow dynamics of agents. Therefore, an arbitrarily chosen threshold $t = 30$ is implemented into the model at which the traders balances are evaluated and ordered from best to worst. On the cutoff date the worst 20% of all traders are removed. Note that this not only eliminates traders with a negative account but rather also sorts out agents with a positive balance. This allows for a more vivid mixing of traders and supports the idea of the survival of the fittest. In addition it is assumed that the worst traders are replaced by some new agents. As was explained earlier in Section 2.2 on page 10 equilibria of type (ii) and (iii)
are avoided by the additional constraint of having at least \( n_{\text{min}} = 4 \) agents in each trading strategy as well as having a maximum of \( N \cdot z_{\text{max}} \) chartists in total. By removing and adding agents from and to the system, it is possible that these conditions are violated. Therefore the outflow and inflow of the traders is handled as follows.

There are no restrictions on how many agents of one group leave the system. It would be possible for example that all technical traders are among the worst 20%. However, when newly populating the system with incoming agents, the lower boundary of \( n_{\text{min}} \) and the upper limit of \( N \cdot z_{\text{max}} \) are adhered to. The repopulation follows the same procedure introduced in Section 2.2 on page 10, i.e. agents are randomly assigned to the three groups and can be mathematically expressed as in Equation (3.5).

\[
\begin{align*}
  n_{\text{clb}} &= \max(2 \cdot n_{\text{min}} - n_{cs}, 0) \quad n_{\text{cub}} = \min[N \cdot x, N \cdot z_{\text{max}} - n_{cs}] \\
  n_{\text{cin}} &= U([n_{\text{clb}}, n_{\text{cub}}]) \quad n_{\text{fin}} = N \cdot x - n_c \\
  n_{\text{olb}} &= \max(n_{\text{min}} - n_{os}, 0) \quad n_{\text{oub}} = \max[n_{\text{cin}} - \max(n_{\text{min}} - n_{ps}, 0), 0] \\
  n_{\text{oin}} &= U([n_{\text{olb}}, n_{\text{oub}}]) \quad n_{\text{pin}} = n_{\text{cin}} - n_{\text{oin}}
\end{align*}
\]

The inflow of chartists \( (n_{\text{cin}}) \) is a uniformly distributed random variable on the interval \([n_{\text{clb}}, n_{\text{cub}}]\) where the lower bound of chartists assures that there are at least \( 2 \cdot n_{\text{min}} \) technical traders in the environment and the upper bound assures that there are not more than \( N \cdot z_{\text{max}} \). The variable \( n_{cs} \) represents the sum of optimistic \( (n_{os}) \) and pessimistic \( (n_{ps}) \) traders after the elimination of the worst \( x = 20\% \). On the other hand the inflow of fundamentalists \( (n_{\text{fin}}) \) is the difference between the total amount of incoming agents \( (N \cdot x) \) and the inflow of chartists \( (n_{\text{cin}}) \). Calculating the lower \( (n_{\text{olb}}) \) and upper bound \( (n_{\text{oub}}) \) assures that there are at least \( n_{\text{min}} \) agents in each technical trading strategy.
4 Results and Interpretation

After the adjustments to the model were made, an equally sized dataset as we find in Lux and Marchesi was generated, i.e. 20,000 observations split into 40 subsamples of 500 observations each. However, this time the simulation included the adaptations and the threshold that was presented in Section 3.2 on page 30.

4.1 Random Walk

As it was suggested earlier in Section 3.1.1 on page 20, it is shown that the model is indeed able to generate price series that contain the unit-root property. In the original model, the fundamental value was constantly set to 10 throughout the simulation. With fundamental traders perfectly being aware of the fundamental price at all times, this led to a mean-reverting price series fluctuating around that constant value. It was shown that Lux and Marchesi probably performed a flawed Dickey-Fuller test for the unit-root property, as they presumably excluded the constant by mistake in their test. This lead to wrong estimates of the coefficient $\hat{\rho}$ and the illegitimate acceptance of $H_0$.

In order to restore the models ability to generate prices that follow a random walk process, a simple change to the fundamental value has to be made. If one assumes that the fundamental value follows a stochastic process in the form of $p_t = p_{t-1} + \epsilon_t$ where $\epsilon_t$ is a normal distributed disturbance term, simulating the incoming positive
or negative news, the model indeed produces an integrated time series of order 1 (I(1)). Considering that in the real world companies or stock indices are always subject to news, coming from micro- or macroeconomic sources, it is also a reasonable addition to the model to make it more realistic. Figure 4.1 shows the pathway of the asset price for the complete set of 20,000 observations. Starting at a price of 10, the stock quotation then takes a random development, almost crashes but recovers and closes above the initial starting point. Now, one is unable to reasonable judge only by visual inspection whether the series is non-stationary or not and therefore has to rely on the statistical results of a unit-root test.

Figure 4.1: Development of the price when the fundamental value follows a stochastic process using a normal distributed news term $\epsilon_t = N(0, 0.1)$.

Performing the Augmented Dickey-Fuller test in its simplest form, i.e. no trend and no lagged differences, given in Equation (3.2) on page 24, results in the output
presented in Table 4.1. When looking at the complete set of 20,000 observations, one is not able to reject the null of a unit-root at all three confidence levels.

<table>
<thead>
<tr>
<th>Dickey-Fuller test for unit root</th>
<th>Number of obs = 19999</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Interpolated Dickey-Fuller</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test</td>
</tr>
<tr>
<td>Statistic</td>
</tr>
<tr>
<td>Z(t)</td>
</tr>
</tbody>
</table>

MacKinnon approximate p-value for $Z(t) = 0.1039$

| D.price | Coef. | Std. Err. | t | P>|t| | [95% Conf. Interval] |
|---------|-------|-----------|---|-----|----------------------|
| L1.price | -0.000659 | 0.000258 | -2.55 | 0.011 | -0.001165 | -0.000152 |
| _cons | 0.006074 | 0.002524 | 2.41 | 0.016 | 0.001126 | 0.011021 |

Table 4.1: Results of the Augmented Dickey-Fuller test performed in Stata. The null hypothesis of a unit-root is not rejected.

However, when splitting the data into subsamples of 500 observations each, the results change. Table 4.2 on the next page provides a summary of the 40 Augmented Dickey-Fuller tests. It can be seen that in 34 of 40 cases, the presence of a unit-root is not rejected at the 5% significance level. This number increases to 38 of 40 cases when using a more stringent confidence level of 99%. However, it also inevitably raises the probability of accepting a false null.

The results therefore join the ranks of mixed findings on this topic. Even if stock prices are predictable to some extent, this does not necessarily mean that there are feasible arbitrage opportunities. Exploiting these anomalies could be difficult and cumbersome or even impossible due to certain circumstances such as transaction
costs or taxes (Dupernex, 2007).

<table>
<thead>
<tr>
<th>Parameters</th>
<th>range of $\hat{\rho}$</th>
<th>No. of rejections for one-sided test at 95% level</th>
<th>No. of rejections for one-sided test at 99% level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter set I</td>
<td>-0.094559 – -0.000412</td>
<td>6</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4.2: Results of the Augmented Dickey-Fuller test where the fundamental value follows a stochastic process.

In the next sections the concept of the fundamental value following a random walk process is discarded and set to its constant value again in order to ease the comparison between the original results and the findings that are due to the adaptations that were made.

## 4.2 Aggregated Wealth

One of the most important changes to the original model is the individualization of the traders and thus transforming the multi-agent system by Lux and Marchesi into a single-agent model. However, sometimes, it is reasonable to look at the aggregate level to draw inferences about a single entity. Furthermore, following Lux and Marchesi, every single agent in his specific trading strategy behaves the same way, e.g. parameters for fundamentalists are the same for all traders in this strategy. The reason for this is simply due to the easier comparison between both models. That being said, agents are still distinguishable from each other by their personal wealth that they accumulate over time. The development of the aggregated wealth of chartists and fundamentalists is an indicator for the profitability of the two trading strategies. It gives information about whether one of the strategies performs better than the other.
Figure 4.2 shows the aggregated wealth of fundamentalists and chartists over a period of 5,000 observations. The wealth of traders is calculated in the same way as was presented in Section 3.2.2 on page 31 and is accumulated for each strategy over the chosen time horizon. This can be thought of as if traders never changed their strategy during the simulation and strictly followed one trading rule.

The picture gives rise to the assumption that fundamentalists outperform chartists and consequently drive them out of the market.

Two publications by Bloembergen (2015) and Bloembergen et al. (2015) support this claim under certain circumstances. They developed a multi-agent-based model closely related to the model of Tóth and Scalas (2008) and the experimental market studied by Huber et al. (2008) in order to investigate the dynamics of trading strategies and the value of information. They find that fundamentalists consistently outperform other market participants (chartists and zero-intelligence traders) when
the acquiring of fundamental information is cost-free. However, when information costs are taken into account, traders of the other subgroups are able to survive. This creates a market equilibrium where each trading strategy is present (see Tóth and Scalas, 2008). The excess profits of fundamentalists in the model of Lux and Marchesi could be diminished by the addition of costs for gathering information about the true fundamental value. Furthermore, since traders are allowed to switch their behaviour during the trading period, potential losses caused by following technical trading could be prevented by switching their strategy to the fundamental approach. In addition to that, it is reasonable to assume that in real life, markets are not static but rather dynamic in terms of traders leaving and entering the market. It is possible that chartists are replaced by chartists again. This will be explained in Section 4.3. Therefore, for the reasons mentioned above, it can be concluded that fundamentalists not necessarily cut out technical traders in the model developed by Lux and Marchesi. This also justifies the inclusion of the minimum amount $n_{\text{min}}$ of traders in each trading strategy.

4.3 Dynamic Market Population

It was shown in Section 4.2 on page 37 that chartists are potentially driven out of the market by fundamental traders because they outperform the technical strategy. In the original model of Lux and Marchesi this lead to a static model where agents could incur infinite negative balances without being punished. This is an unrealistic assumption because in the real world investors would have to meet their margin calls. If they are not able to do so they would have to file for bankruptcy and are therefore removed from the market. As was pointed out earlier in Section 3.2.3 on page 32 it is assumed that after a certain cut-off date ($t = 30$, simulating a threshold
date of 1 month), the worst 20% of the whole population is removed from the market and randomly replaced by newly incoming agents of both types.

In the following the effects of this dynamic population are investigated. An emphasis is put on the statistical differences of the characteristics of the price series, the return series and the composition of the market compared to the initial model.

### 4.3.1 Properties of the Price Series

The main interest of this thesis is to gather insights in the effects of a dynamic trader population on the market. Therefore two datasets are generated with 20,000 observations each in order to be in line with the evaluation of Lux and Marchesi. The first dataset, referred to as original, generates a time series of prices according to the rules outlined in Chapter 2 on page 3. The second dataset is simulated by the very same rules but include the modifications presented in Section 3.2 on page 30. This allows to study the effects of traders leaving and entering the market. Descriptive statistics are given in Table 4.3. Furthermore, the two price series are depicted in Figure 4.3 on the following page.

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{\text{original}}$</td>
<td>20000</td>
<td>10.002080</td>
<td>0.107076</td>
<td>9.01</td>
<td>10.89</td>
</tr>
<tr>
<td>$D_{\text{dynamic}}$</td>
<td>20000</td>
<td>10.001660</td>
<td>0.177060</td>
<td>8.46</td>
<td>11.28</td>
</tr>
</tbody>
</table>

Table 4.3: Summary statistics of the simulated price series.
CHAPTER 4. RESULTS AND INTERPRETATION

(a) Time series of prices of the original model.

(b) Time series of prices with dynamic population.

Figure 4.3: Comparison between the simulated price series of the original and the replicated model with the addition of the dynamic population.

By visual inspection, one can clearly see that Figure 4.3b is different to Figure 4.3a. Both prices fluctuate around the constant fundamental value of 10. However, it
CHAPTER 4. RESULTS AND INTERPRETATION

<table>
<thead>
<tr>
<th>Observations</th>
<th>Ratio $= \frac{sd_1}{sd_2}$</th>
<th>No. of rejections in favour of $R &gt; 1$</th>
<th>No. of rejections in favour of $R &lt; 1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>0.303356 – 1.169418</td>
<td>1</td>
<td>36</td>
</tr>
</tbody>
</table>

Table 4.4: Results of the variance-homogeneity test.

seems that prices in the second picture fluctuate more heavily around this base. This is also shown in the descriptive statistics where the standard deviation of the second dataset is larger than the standard deviation of the original dataset.

This view is confirmed by the statistical analysis of the series. Both datasets were split into 40 subsamples of 500 observations each. Following this, the subsamples are compared with each other and tested for homogeneity in their variances. Table 4.4 presents the results of all 40 tests. The variances were compared with the Stata inbuilt function $sdtest$ explained in Armitage et al. (2002, p. 149-153). This tests whether the ratio of the standard deviations of two datasets is equals to 1. It can be seen that out of 40 subsamples, 36 reported a significantly lower ratio ($< 1$) indicating a larger standard deviation in the second dataset. A larger standard deviation is mirrored in more extreme negative and positive prices, as can be seen in Table 4.3 on page 40.

A statistical analysis of the price level is not necessary because the price series is stationary around a constant mean. This is supported by the mean of the whole sample size given in Table 4.3 on page 40. The values for both datasets are very close to the fundamental value of 10. Performing an unpaired two-sample t-test on the mean of the subsamples would yield statistically significant differences, however, those differences are neglectable as they are too small to give a meaningful interpretation. Therefore the choice of the statistical test for comparing the variance came
down to the Stata built in function `sdtest` instead of the usual Kolmogorov-Smirnov test, which would compare the whole distribution.

### 4.3.2 Properties of the Return Series

In addition to the price series, the return series is also checked for any noticeable differences. Table 4.5 lists the descriptive statistics of both sets of data. It is noticeable that the mean of both series is effectively zero. This can be explained by the stationary fundamental value and the mean of the price series being very close to that value. Performing a t-test on the equality of means of the two series yields no significant differences for all subsamples. However, inspecting Figure 4.4 on the next page gives again rise to the assumption that there is a difference between the datasets.

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_{\text{original}}$</td>
<td>20000</td>
<td>0.000000</td>
<td>0.008556</td>
<td>-10.22%</td>
<td>10.37%</td>
</tr>
<tr>
<td>$D_{\text{dynamic}}$</td>
<td>20000</td>
<td>0.000000</td>
<td>0.017536</td>
<td>-18.54%</td>
<td>15.42%</td>
</tr>
</tbody>
</table>

Table 4.5: Summary statistics of the simulated return series.
CHAPTER 4. RESULTS AND INTERPRETATION

(a) Time series of returns of the original model.

(b) Time series of returns with dynamic population.

Figure 4.4: Comparison between the simulated return series of the original and the replicated model with the addition of the dynamic population.

Similarly to the price series, the returns are tested for the homogeneity in their variances. Therefore the same test as previously mentioned is applied to the dataset.
The results are presented in Table 4.6 and yield the same outcome as it was the case for the prices. The dataset with the dynamic population is significantly more volatile, leading to more extreme negative and positive returns.

<table>
<thead>
<tr>
<th>Observations</th>
<th>Ratio = \frac{sd_1}{sd_2}</th>
<th>No. of rejections in favour of R &gt; 1</th>
<th>No. of rejections in favour of R &lt; 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>0.179508 – 1.363158</td>
<td>1</td>
<td>36</td>
</tr>
</tbody>
</table>

Table 4.6: Results of the variance-homogeneity test for returns.

Below are the results of the Hill estimates, conducted in the same way as was explained in Section 3.1.2 on page 26. The results for the second dataset are similar to the original model, i.e. kurtosis is significantly larger than 3 for the case of a normal distribution. Furthermore the data exhibits fat-tails as indicated by the Hill estimators in Table 4.7. However, for the second series the results are less prominent but still in the range that is observed in real financial markets (see Voit, 2010, p. 149f).

<table>
<thead>
<tr>
<th></th>
<th>Kurtosis</th>
<th>2.5% tail</th>
<th>5% tail</th>
<th>10% tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>D_{original}</td>
<td>18.95</td>
<td>3.15</td>
<td>2.65</td>
<td>2.23</td>
</tr>
<tr>
<td></td>
<td>(1.96–4.25)</td>
<td>(1.74–3.22)</td>
<td>(1.54–2.67)</td>
<td></td>
</tr>
<tr>
<td>D_{dynamic}</td>
<td>11.95</td>
<td>3.45</td>
<td>3.01</td>
<td>2.42</td>
</tr>
<tr>
<td></td>
<td>(2.49–3.81)</td>
<td>(2.14–3.55)</td>
<td>(2.09–2.99)</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.7: Fat-tail properties of the original model and the dynamic model. Kurtosis is reported for the complete set of 20,000 observations. Median \( \alpha_H \) tail index estimates are reported from 10 samples of 2,000 observations each (range of estimates in parentheses).
CHAPTER 4. RESULTS AND INTERPRETATION

Figure 4.5: Autocorrelation function for raw, squared and absolute returns plotted over 150 lags after the adjustments.

Additionally, the second data series is checked for long range dependency and volatility clusters. Figure 4.5 plots the Autocorrelation function of raw, squared and absolute returns in the same way as was shown earlier in Section 3.1.3 on page 29. We can see that the autocorrelation function for squared and absolute returns are again slowly decaying. However, after the adjustments were made this seems to happen a bit faster than in the original setting. Nevertheless, the returns show significant autocorrelation in squared and absolute returns and they therefore conform with the stylized fact observed in real financial data.

4.3.3 Market Composition

Even though we have evidence that the dynamic population induces excessive volatility, it is still unclear where the excessive volatility in the price and return series of dataset 2 is coming from. Therefore the market composition is studied. As already explained in Chapter 2 on page 3, the market consists of three kinds of traders, fundamentalists and chartists, subdivided into optimistic and pessimistic ones. Fig-
Figure 4.6 on the following page depicts the chartist index \( z = \frac{n_{co} + n_{cp}}{N} \), i.e. the proportion of chartists compared to the total amount of traders. Descriptive statistics are given in Table 4.8. Note that both datasets share the same minimum and maximum value. This is due to the boundaries that were implemented to guarantee that the model stays within equilibrium (i) as was explained in Section 2.1 on page 3. Multiplying the minimum value of 0.016 with the total amount of traders \( (N = 500) \), yields \( 2 \cdot n_{min} \), respectively, multiplying the maximum results in \( z_{max} \cdot N \).

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D_{\text{original}} )</td>
<td>20000</td>
<td>0.138919</td>
<td>0.120766</td>
<td>0.016</td>
<td>0.66</td>
</tr>
<tr>
<td>( D_{\text{dynamic}} )</td>
<td>20000</td>
<td>0.291975</td>
<td>0.150155</td>
<td>0.016</td>
<td>0.66</td>
</tr>
</tbody>
</table>

Table 4.8: Summary statistics for the simulated chartist indices.

Visual inspection shows that the maximum value \( z_{max} \) is more frequently reached in Figure 4.6b than in Figure 4.6a. This should result in a positively shifted mean in the second dataset. The indicated assumption is supported by the descriptive statistic given above.
CHAPTER 4. RESULTS AND INTERPRETATION

(a) Development of the chartist index for the original model.

(b) Development of the chartist index with dynamic population.

Figure 4.6: Comparison of the market composition between the simulated return series of the original and the replicated model with the addition of the dynamic population.
Since the data is not normally distributed, the use of the two independent samples t-test is not allowed. Instead the non-parametric Wilcoxon-Mann-Whitney test is used to compare 40 subsets of 500 observations each. Out of 40 tests, 37 resulted in ranks being higher in dataset 2.

<table>
<thead>
<tr>
<th>Observations</th>
<th>No. of equal ranks</th>
<th>No. of rejections in favour of R1 = R2</th>
<th>No. of rejections in favour of R1 &lt; R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>2</td>
<td>1</td>
<td>37</td>
</tr>
</tbody>
</table>

Table 4.9: Results of the Wilcoxon-Mann-Whitney test for the chartist index.

Figure 4.7: Enlarged sample of the chartist index. The vertical lines represent the cut-off dates.

Figure 4.7 depicts a close-up shot of the chartist index. The vertical lines represent the cut-off dates where traders leave and enter the market. It can be seen that after the cut-off date the composition of the market tends to shift more rapidly and thus
contribute to the volatility increase in the market.

In the previous sections evidence was reported for a volatility increase in prices and returns. Furthermore it was shown that by allowing agents to leave and enter the market a new equilibrium is created where chartists are more prominent than in the original model. This can be explained by the nature of the chartists themselves. Since they are strictly trend following and base their expectations on the price change as well as on the opinion index, the arrival of new chartists in the model triggers a rally. This leads to a self-fulfilling prophecy where chartists drive prices up. Consequently more traders change their trading strategy and use a technical approach until the deviations from the fundamental value are large enough to switch to the fundamental strategy again. This leads to more volatile prices and returns.
5 Conclusion

In this master thesis an existing agent-based model by Lux and Marchesi (2000) was recreated in order to investigate the effects of traders leaving and entering the artificial stock market. The agent-based approach was chosen due to the inabilities of traditional models in explaining several stylized facts and the emergence of bubbles, e.g. the Chinese Warrants Bubble, where inexperienced traders entered the market and contributed to the excessive volatility in the market, a behaviour that could not be explained with a representative agent.

Therefore the original model was explained in detail in Chapter 2 on page 3. This chapter covered the basic idea of the model, the results as well as the importance of their findings. The results of the two authors are replicated in Chapter 3 on page 19 and validated. A minor mistake in the original study was found where the authors probably applied a wrong unit root test to the generated dataset. This resulted in the incorrect reporting of a non-stationary time series, when in fact, the series was stationary. Furthermore, the modifications to the original model were explained in this chapter. One of the most important changes was the transformation of the multi-agent system of Lux and Marchesi to a single-agent model. This allowed the implementation of a personal bank account for every agent in the system which later on serves as the decisive factor of agents leaving and entering the market. A cut-off date was implemented where the worst performing traders are removed from the
market. This allowed new traders to enter the market.

In Chapter 4 on page 34 it was shown that by removing and replacing traders in the artificial stock market, the price and return volatility significantly increased. The main driver of this volatility increase was in consequence of chartists entering the market and triggering a price rally. This led to a self-fulfilling prophecy where more traders changed to a technical approach and drove the price further away from its fundamental price. It was shown that the stylized facts of autocorrelation in returns, fat tails and volatility clustering are still present in the market.

The results contribute to the existing literature by giving a possible explanation for the volatility increase in e.g. the Chinese Warrants Bubble in 2005-2008. Furthermore, since the developed software is publicly available now, further research can be conducted. Even though traders are allowed to leave and enter the market, the total amount of agents is still fixed at 500. One possible extension could be the removal of the limitation of a static amount of traders in the model without running into the problem of the finite-size effect put forward by Egenter et al. (1999). Another research direction could go towards the implementation of information costs for fundamental traders in order to make the model more realistic.
Bibliography


Mandelbrot, Benoit (1967). “The Variation of Some Other Speculative Prices”. In: The Journal of Business 40.4, p. 393. ISSN: 0021-9398. DOI: 10.1086/295006.


## List of Figures

2.1 Simulation flowchart of Lux and Marchesi .......................... 11
2.2 Link between the chartist index and volatility clusters .......... 16
2.3 Comparison between the simulated autocorrelation and the autocorrelation of the S&P 500 index .......................... 17

3.1 Comparison between the simulated price series of the original and the replicated model. ........................................... 22
3.2 Time series with a constant line at the price level of 10. ............ 23
3.3 Normal Q-Q plot of the simulated returns ............................... 28
3.4 Autocorrelation function for raw, squared and absolute returns .... 29

4.1 Development of the price when the fundamental value follows a stochastic process ..................................................... 35
4.2 Aggregated wealth of both trading strategies ............................ 38
4.3 Comparison between the simulated price series of the original and the replicated model with the addition of the dynamic population. ... 41
4.4 Comparison between the simulated return series of the original and the replicated model with the addition of the dynamic population. ... 44
4.5 Autocorrelation function for raw, squared and absolute returns after adjustments ....................................................... 46
4.6 Comparison of the market composition between the simulated return series of the original and the replicated model with the addition of the dynamic population. .......................... 48
4.7 Enlarged sample of the chartist index ................................. 49
List of Tables

2.1 Original parameter set of Lux and Marchesi .......................... 10
3.1 Results of the unit root tests reported by Lux and Marchesi. ....... 24
3.2 Results of the unit root test from the replicated model ............... 24
3.3 Results of the unit root test from the replicated model but without a constant ......................................................... 25
3.4 Fat-tail properties reported by Lux and Marchesi .................... 27
3.5 Fat-tail properties of the recreated model ............................... 28
4.1 Results of the Augmented Dickey-Fuller test where the fundamental value follows a stochastic process ............................ 36
4.2 Results of the Augmented Dickey-Fuller test in the subsamples, where the fundamental value follows a stochastic process ............ 37
4.3 Summary statistics of the simulated price series. ........................ 40
4.4 Results of the variance-homogeneity test ................................. 42
4.5 Summary statistics of the simulated return series ........................ 43
4.6 Results of the variance-homogeneity test for returns .................. 45
4.7 Fat-tail properties of the original model and the dynamic model ... 45
4.8 Summary statistics for the simulated chartist indices .................. 47
4.9 Results of the Wilcoxon-Mann-Whitney test for the chartist index... 49