CHALMERS

Reproducing the stylized facts of financial returns

An investigation of the parameter space of a stochastic discrete-time model

Master's thesis in Complex Adaptive Systems

RASMUS EINARSSON

Department of Energy and Environment
Division of Physical Resource Theory
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2013
Master's thesis 2013:2

MASTER'S THESIS IN COMPLEX ADAPTIVE SYSTEMS

Reproducing the stylized facts of financial returns

An investigation of the parameter space of a stochastic discrete-time model

RASMUS EINARSSON

Department of Energy and Environment
Division of Physical Resource Theory
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2013

Reproducing the stylized facts of financial returns
An investigation of the parameter space of a stochastic discrete-time model
RASMUS EINARSSON

Master's thesis 2013:2
Department of Energy and Environment
Division of Physical Resource Theory
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: +46 (0)31-772 1000

Chalmers Reproservice
Göteborg, Sweden 2013

Reproducing the stylized facts of financial returns
An investigation of the parameter space of a stochastic discrete-time model

RASMUS EINARSSON
Department of Energy and Environment
Chalmers University of Technology

Abstract

Statistical analysis of price fluctuations in various financial markets has revealed a number of statistical regularities that are persistent across different markets, asset types, and time periods. These regularities have been studied extensively during the last decades, and by now a number of them are widely acknowledged as so-called "stylized facts" of financial price series. We present a subset of the most well-established stylized facts of financial price series, and mention a few mathematical modeling approaches that have been used in attempts to reproduce the stylized facts.

The main part of this thesis is devoted to one such mathematical model proposed in the literature, which has been demonstrated using numerical simulation to reproduce a few of the stylized facts. We begin by simplifying the original model by making a slight approximation and a variable change. The resulting simplified model is a discrete-time stochastic process with three parameters. We derive the requirements for convergence of the theoretical moments of the stochastic process, and find analytical expressions for some quantities relating to the stylized facts. Numerical simulations are combined with analytical techniques to explore the parameter space of the model. Using the restrictions required for the existence of moments of the stochastic process, and the restrictions required to match some stylized facts, we find a region in the parameter space where the model is in reasonable agreement with the stylized facts.

The final chapter contains a few general thoughts on the purpose, difficulties and potential benefits of developing mathematical models to match the stylized facts.

Keywords: stylized facts, finance, financial, returns, stochastic, discrete, model, parameter, moments

Acknowledgements

First I would like to thank my supervisor, professor Kristian Lindgren. Working on this thesis has been a rewarding experience, not least thanks to your encouraging and open-minded attitude.

During my random walks in the literature and in the exploration of various models, I have also benefited from the interactions with the group of researchers working on land use change and agricultural commodity prices at the Division for Physical Resource Theory. Kristian, David, Liv, Emma, Ville and Martin, thank you for allowing me participate in the discussions.

Contents

Chapter 1

Introduction

Traditional economics and finance are disciplines often accused of horrible things. The fundamental flaw, according to some, is that the basic assumptions about perfect competition, rational agents, perfect information and market equilibrium have been applied too widely and without regard to empirical facts.

In an essay in Nature (2008) entitled *Economics needs a scientific revolution*, physics professor Jean-Philippe Bouchaud wrote:

> "The supposed omniscience and perfect efficacy of a free market stems from economic work done in the 1950s and 1960s, which with hindsight looks more like propaganda against communism than plausible science. In reality, markets are not efficient, humans tend to be over-focused in the short-term and blind in the long-term, and errors get amplified, ultimately leading to collective irrationality, panic and crashes. Free markets are wild markets."

Bouchaud argues that decisions in "wild markets" based on such optimistic theories may have disastrous effects on the real economy. For example, regarding the Black-Scholes model which was designed to price financial options, he wrote that

> "[the model] is still used extensively. But it assumes that the probability of extreme price changes is negligible, when in reality, stock prices are much jerkier than this. Twenty years ago, unwarranted use of the model spiralled into the worldwide October 1987 crash [. . .] ."

And similarly, a partial cause to the financial crisis of 2007/2008, he claims, is

> "the development of structured financial products that packaged sub-prime risk into seemingly respectable high-yield investments. The models used to price them were fundamentally flawed: they underestimated the probability that multiple borrowers would default on their loans simultaneously. These models again neglected the very possibility of a global crisis, even as they contributed to triggering one."

A lengthier and more balanced criticism of traditional economics by Farmer and Geanakoplos (2009) points out some situations where equilibrium models can make useful predictions, and some situations where they can never make useful predictions. But in any case, the authors stress, it "shouldn't be a question of dogma, and should be resolved empirically."

All three, Bouchaud, Farmer and Geanakoplos, have contributed significantly in the last two decades to a set of empirical and theoretical developments called "econophysics", with the aim to provide a complementary view on economics, not least financial economics. The term econophysics was coined by merging the words economics and physics, since the research has drawn some inspiration from the field of statistical physics.

A significant piece of this research has focused on identifying and explaining a set of so-called "stylized facts", statistical regularities that seem to be common to most financial markets. For example, as noted by Bouchaud, real-world stock prices are much more volatile than the Black-Scholes model predicts. Specifically, it has been established as a stylized fact that the price increments of financial instruments have distributions with fat tails, i.e., much higher probability of large jumps than would be predicted by a normal distribution.

A whole range of such stylized facts of financial markets are now widely acknowledged, and there have been several attempts to reproduce them with mathematical models. A model that successfully reproduces some of the stylized facts could be a way of improving our understanding of financial markets: What are the drivers of volatility? Why are there such large price jumps? To the extent that market volatility has negative effects on the real economy, how could the market be stabilized? The aim of this thesis is not to answer any of these hard questions, but to take a closer look at one example of a model from the literature.

1.1 Scope of this thesis

The aim of this thesis project has been to improve the understanding of a stochastic, discrete-time model designed to reproduce some statistical regularities of financial price series, often called stylized facts. The model, proposed by Westerhoff and Franke (2012), was demonstrated using simulation results to match a few stylized facts relatively well. This report presents an effort to map the parameter space of the model, using a combination of analytical methods and computer simulation.

1.2 Organization of the report

The remainder of the thesis is organized as follows.

Chapter 2 presents some mathematical concepts used in the thesis, reviews a small collection of well-established stylized facts of financial return series, and finally mentions a few different modeling approaches that have been used to reproduce the stylized facts.

Chapter 3 presents the analysis method, results and a discussion around the model proposed by Westerhoff and Franke.

Chapter 4 concludes with a general discussion about the purpose, difficulties and potential benefits of developing mathematical models to match the stylized facts.

Chapter 2

Background

This chapter summarizes some topics that are fundamental to the method and the results of our work, which is presented in the next chapter. In section 2.1, we begin by going through some notation and mathematical concepts that will be used. Then in sections 2.2–2.4, we cover some basic concepts in financial time series and introduce some of the so-called "stylized facts". Finally in section 2.5, we mention some of the mathematical modeling approaches used so far to reproduce the stylized facts.

This background chapter should be sufficient to follow the work we present in the following chapters, but it is far from a complete introduction to the topics mentioned above. If the reader is interested to learn more than the very basics, it is highly recommended to take a look outside this meager introduction. The sources referenced in this chapter provide a richer picture of the matters.

2.1 Mathematical tools and notation

2.1.1 Moments and central moments

For a stochastic variable X, the kth moment for a positive integer k is the expectation $\mathbb{E}\left[X^k\right]$, if the expectation exists. The expectation is said to exist for a continuous random variable X with probability density function $f(x)$ if (Rice, 2007, p. 22)

$$\int_{-\infty}^{\infty} f(x)\left|x^k\right| \,\mathrm{d}x < \infty.$$

The first moment is simply the mean $\mathbb{E}[X]$. The kth *central* moment $\mathbb{E}\left[(X-\mathbb{E}[X])^k\right]$ is often used for integers $k \geq 2$. For example, the second central moment is also known as variance. The third central moment is the skewness, which measures the asymmetry of the distribution of the random variable. The fourth central moment is often normalized and called kurtosis,

$$\kappa = \frac{\mathbb{E}\left[(X-\mathbb{E}[X])^4\right]}{\sigma^4},$$

where σ is the standard deviation $\sigma = \sqrt{\mathbb{E}\left[(X-\mathbb{E}[X])^2\right]}$.

2.1.2 Stochastic processes and stationarity

A stochastic process, loosely speaking, is a set of random variables, $\{X_t : t \in T\}$, where T is an ordered set interpreted as "time". In this text, we exclusively deal with one-dimensional, discrete-time stochastic processes, i.e., sequences of real stochastic variables such as $X_1, X_2, \ldots$. We will simply refer to such a stochastic process as $\{X_t\}$. See, e.g., (Najim, Ikonen, and Daoud, 2004) for a rigorous introduction.

We will say that a stochastic process $\{X_t\}$ is strictly stationary if the joint distribution of any subset of variables $\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\}$ is not affected by a shift in time, i.e., the two subsets

$$\{X_{t_1}, X_{t_2}, \ldots, X_{t_n}\} \text{ and } \{X_{t_1+\Delta t}, X_{t_2+\Delta t}, \ldots, X_{t_n+\Delta t}\}$$

have the same joint distribution for any Δt.

Strict stationarity is so called because we can also come up with weaker forms of stationarity requirements. For example, the requirement that the expectation $\mathbb{E}[X_t]$, variance $\mathrm{Var}(X_t)$ and covariance $\mathrm{Cov}(X_t, X_{t+s})$ all exist and are independent of t, is sometimes called covariance stationarity or wide sense stationarity.

Later in this text, we will work with the requirement that a certain process is stationary to the nth order, in the sense that all moments $\mathbb{E}\left[X_t^k\right]$ exist for $k = 1, 2, \ldots, n$, and that they are independent of t.

2.1.3 The autocorrelation coefficient

The correlation coefficient of two stochastic variables X and Y is

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}},$$

which should be interpreted as the covariance of X and Y, normalized such that $-1 \le \rho \le 1$. The value $\rho = 1$ indicates perfect covariance (one goes up, the other one surely does too), $\rho = -1$ perfect anti-covariance (one goes up, the other one goes down), and $\rho = 0$ if and only if X and Y are statistically independent.

Assuming that a stochastic process $\{Z_t\}$ is at least weakly stationary, the autocorrelation coefficient at time lag τ, between Z_t and $Z_{t+\tau}$, is independent of t, and it is denoted

$$C_\tau\left(Z_t\right) = \frac{\mathrm{Cov}(Z_t, Z_{t+\tau})}{\mathrm{Var}(Z_t)}.$$

2.1.4 Hill's tail estimator

Hill (1975) proposed a simple and general approach for statistical inference about the shape of the tail of a distribution. Assuming that the distribution is of the Zipf type, i.e., the cumulative density function follows $F(x) = 1 - Cx^{-\alpha}$ for large x, a maximum likelihood estimator for the exponent α was derived. No assumption is needed about the general shape of the distribution, only that the distribution tail follows a power law.

The upper tail, i.e., for $x \to +\infty$, is estimated by taking a sample $X_1, \ldots, X_k$ from the distribution, which is assumed to follow $F(x) = 1 - Cx^{-\alpha}$ for $x \ge D$ for some known D. Take the largest r observations $x_i, i = 1, 2, \ldots, r$, ordered such that $x_r \le y_{r-1} \le \cdots \le x_1$ and where $x_r \ge D$. The maximum likelihood estimator, from here on referred to as Hill's tail estimator, is then

$$\hat{\alpha} = \frac{r}{\sum_{i=1}^{r-1} \ln x_i - (r-1)\ln x_r}.$$

In practice, it is an arbitrary choice how many of the largest observations to pick out. We generally take all observations above a certain sample quantile, by taking fraction f_{H} largest samples. For example, to take everything above the 95% sample percentile, we choose where $f_{\mathrm{H}} = 0.05$ and take

$$\lfloor r \rfloor = 0.05k,$$

where $\lfloor \cdot \rfloor$ is the floor function.

The tail of the distribution puts an upper limit on how many finite moments the distribution can have. Specifically, a distribution can never have a finite kth moment if $\alpha < k$. This can be seen by noting that the kth moment exists only if $\int \left|x^k\right| f(x)\,\mathrm{d}x < \infty$, and assuming that the distribution tail follows a power law above some finite D, i.e., $f(x) = Cx^{-(1+\alpha)}$ for $x > D$,

$$\int_{-\infty}^{\infty} \left|x^k\right| f(x)\,\mathrm{d}x \;\leq\; 2\int_D^{\infty} x^k C x^{-(1+\alpha)}\,\mathrm{d}x = 2C\int_D^{\infty} x^{k-1-\alpha}\,\mathrm{d}x.$$

This integral is infinite if $k \geq \alpha$, so the tail index α is an upper limit for the highest finite moment.

2.2 Financial time series

We begin with a little bit of notation for financial time series. Let $P(t)$ be the price of an asset, sampled by taking the price used in the latest transaction before time t. What really matters to most traders of financial instruments is not absolute prices, but the relative price change, the return, over a time period Δt. The return is

$$R_t(\Delta t) = \frac{P(t) - P(t - \Delta t)}{P(t - \Delta t)},$$

where the notation $R_t(\Delta t)$ is used to indicate that returns depend on the time scale. In practice, what is more often studied is the log return, defined as

$$r_t(\Delta t) = \ln\left(1 + R_t(\Delta t)\right) = \ln\left(\frac{P(t)}{P(t - \Delta t)}\right) = \ln P(t) - \ln P(t - \Delta t).$$

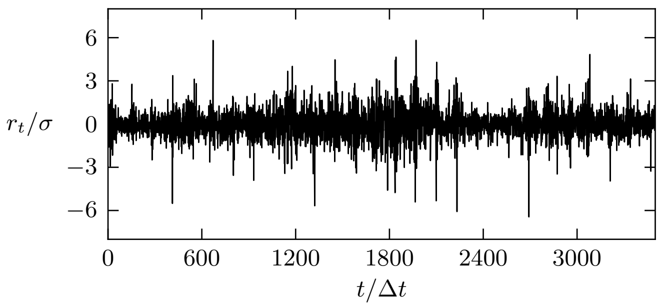
For simplicity, we define the log price $p_t = \ln P(t)$, so that

$$r_t(\Delta t) = p_t - p_{t-\Delta t}.$$

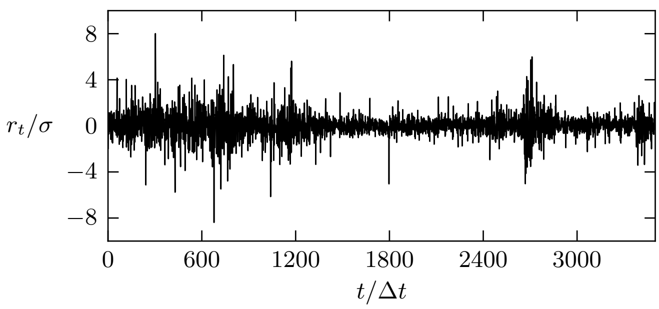
Two things should be noted at this point. First, we will only consider log returns in the remainder of this report. Note that for small returns $R_t(\Delta t) \ll 1$, returns and log returns are approximately equal, since $\ln\left(1 + R_t(\Delta t)\right) \approx R_t(\Delta t)$. From here on, we write "returns" or "log returns" interchangeably, but the strict meaning is always log returns. Similarly, we may write "prices" or "log prices", which always should be interpreted as log prices. Second, we will eventually measure time in units of Δt, so that $r_t = p_t - p_{t-1}$, but for now we will be explicit about the time scale while reviewing a few empirical facts.

2.3 The stylized facts

The idea of stylized facts was introduced by the macroeconomist Kaldor, who argued that scientists should be free to start "with a stylized view of the facts". Kaldor identified a number of statistical facts concerning macroeconomic growth and used these as a starting point for his theoretical modeling (Kaldor, 1961; Chakraborti et al., 2011a).



(a) $\Delta t = 3$ minutes



(b) $\Delta t = 1$ day

Figure 2.1: Log return series of the IBM stock at time scales of 3 minutes (top) and 24 hours (bottom), normalized by the sample standard deviation of returns during the two periods. Visual inspection of the data shows that the return distribution has fat tails at least compared to a normal distribution: returns of large magnitude are much more common than in a normal distribution, where the probability of observing values outside, for example $\pm 5\sigma$ is less than 10^{-6}. The figure also illustrates that the volatility (the magnitude of fluctuations) varies over time at both time scales. In other words, the volatility is relatively low for a while, then higher for a while, etc. The 3-minute returns series (top) covers almost two calendar weeks in late 2007, and the 1 day returns series (bottom) covers about 14 calendar years 1998–2011.

The concept was later adopted to describe some statistical regularities in return series of financial assets that are persistent across different markets, asset types, time periods, return time scales, etc. By now a number of stylized facts are widely acknowledged, not least because the data availability has increased enormously since financial exchanges have moved from physical trading floors to electronic exchanges, where all transactions are recorded to a database.

To get an intuitive feel for a couple of these stylized facts, it may be interesting to start with a technique called visual inspection or, colloquially, eyeballing the data. Figure 2.1 shows log returns of the IBM stock at two different time scales and illustrates two things rather clearly. First, the distribution of returns is much more fat-tailed than a normal distribution. The plots both show several returns of more than five standard deviations in 3500 samples, whereas the normal distribution would take such extreme values with probability below 10^{-6}. Second, it is obvious that the volatility of the returns series, often measured as the standard deviation over a shorter period, is not constant. At both time scales, it seems like volatility is "sticky": it is high for a while, then drops gradually to a lower level for a while, then rises again, etc. This property is often referred to as clustered volatility.

In the following pages, we will briefly review some of the most widely acknowledged stylized facts of financial return series. For a more complete treatment, see, e.g., the reviews by Cont (2001), Chakraborti et al. (2011a), and Guillaume et al. (1997). The following material draws heavily on those sources.

2.3.1 No autocorrelation in returns

The first and perhaps most obvious stylized fact of financial time series is that returns are not significantly autocorrelated. Assuming that returns were significantly autocorrelated, one could make correct predictions of returns at least on average, which could be used by traders to make profitable investment decisions. When autocorrelations are exploited, market prices should equilibrate in such a way that the autocorrelations disappear.

Nevertheless, there is a whole discipline in quantitative finance called technical analysis, which aims at forecasting the direction of price changes by studying past price data and finding patterns that can be used as signals to buy or sell. Consider popular book titles like *Technical Analysis: The Complete Resource for Financial Market Technicians* (Kirkpatrick and Dahlquist, 2010) or *Chart Your Way to Profits: The Online Trader's Guide to Technical Analysis* (Knight, 2007).

To be fair, technical trading strategies not only include looking at past price movements, but also finding correlations across different markets and assets, and whatever other information could be valuable. Still, the concept of technical trading is incompatible with one of the theoretical cornerstones of finance called the efficient market hypothesis (EMH). The EMH states that financial markets makes efficient use of all available information, thereby instantly incorporating any "hidden value" of an asset into its price. If the EMH is strictly true, there is no chance of making money on technical analysis: the future returns of exchange-traded financial assets are always unpredictable.

This apparent inconsistency in the existence of technical trading strategies on one hand, and the stylized fact of absence of autocorrelation on the other, is perhaps not so important. The fact is only a stylized one, which means that it should hold qualitatively, but perhaps not in every case.

In practice, the information professional traders have about price movements is more or less instantaneous, and it does seem like they utilize the information rather efficiently. See figure 2.2 for an indication that the autocorrelation in returns, $C_\tau(r_t(\Delta t))$ seems to vanish on average within a few minutes for the IBM stock.

2.3.2 Long autocorrelations in squared returns

The next stylized fact concerns the autocorrelation of the squared returns process $\{r_t^2\}$. Already in an influential paper from 1963, Mandelbrot wrote that "large price changes are not isolated between periods of slow change" but "large changes tend to be followed by large changes – of either sign – and small changes tend to be followed by small changes" (Mandelbrot, 1963, p. 418). In other words, the autocorrelation of absolute or squared returns is positive. This is sometimes referred to as a long memory effect or clustered volatility.

This statistical fact has since been verified in many financial markets, and it has been quantitatively refined. According to Cont (2001), several authors have remarked that the autocorrelation in squared returns decays like a power law,

$$C_\tau(r_t^2(\Delta t)) \sim \tau^{-\beta},$$

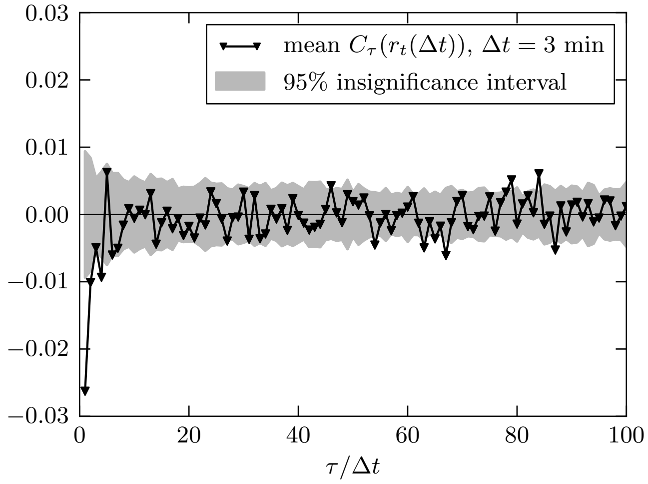


Figure 2.2: Sample autocorrelation of 3-minute returns of the IBM stock and the insignificance interval which sample values fall inside with probability 95% under the null hypothesis $H_0 : C_\tau(r_t) = 0$. The first autocorrelation at $\tau = 3$ minutes is highly significantly different from zero ($p < 10^{-6}$). The hypothesis testing was done with a bootstrap technique, sampling 50 random blocks of $5 \cdot 10^3$ consecutive values in the 3-minute returns series of the IBM stock between January 1998 and February 2013 (total number of data points approximately $5 \cdot 10^5$). Each block sample was used to estimate the autocorrelation at all different values of τ, and the approximate insignificance interval was estimated by assuming that the mean value of sample autocorrelations for each lag τ is normally distributed with mean 0 and standard deviation equal to sample standard deviation.

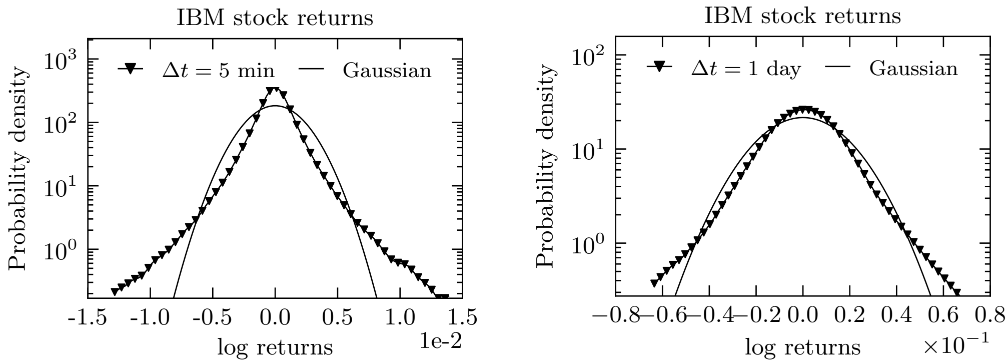


Figure 2.3: Empirical distribution of log returns $r_t(\Delta t)$ of the IBM stock, compared to a normal distribution with the same standard deviation. It is clear that the return distribution has much fatter tails than a normal distribution, both at $\Delta t = 5$ minutes and $\Delta t = 24$ hours, although the shape is less pronounced for longer time scales. The figures are based on data from January 1998 to February 2013. The empirical distributions are Epanechikov kernel estimates using a bandwidth choice suggested by Davidson and MacKinnon (2004, p. 681, equation (15.63)).

with a coefficient $\beta \in [0.2, 0.4]$. The quantitative results may be dependent on the time step Δt, but we will not go deeper into the topic at this point.

This stylized fact is important because it says that price increments are not independently distributed: the magnitude of future returns can be predicted to some extent based on past returns, even though the sign of returns cannot.

2.3.3 Fat tails in the return distribution

Another widely acknowledged stylized fact is that returns are not normally distributed. This is illustrated with an example in figure 2.3, which clearly shows that the return distribution has a narrower central part and fatter tails than the normal distribution. This empirical regularity was noted already by Mandelbrot (1963) and since then, a whole range of different distributions have been suggested, but there is no consensus on the exact form of the tails (Chakraborti et al., 2011a).

The difference can be quantified in at least two different ways. First, empirical return distributions typically have a large normalized fourth moment, also known as kurtosis (see section 2.1.1 on page 3). The kurtosis of the normal distribution is $\kappa = 3$, and a distribution with $\kappa > 3$ is called leptokurtic. The sample kurtosis of financial returns falls roughly in the range $5 < \kappa < 100$ (Cont, 2001; Lux, 2009). No single value should be accepted as universally correct, but it is clear that all return series are significantly leptokurtic.

A second way of quantifying the distribution shape is to compute its tail index using the Hill estimator (see section 2.1.4 on page 4). Malevergne, Pisarenko, and Sornette (2005) found Hill estimator values roughly in the region $2.7 \leq \hat{\alpha} \leq 4.0$ using return series from three different stock indices with return intervals from 5 minutes to 24 hours, and with the Hill estimator based on empirical returns above different quantiles from the 95% to the 99% quantile. Similar results have been reported for daily returns in some stock indices by Lux (2009).

2.3.4 Skewness

Finally we briefly mention the skewness of empirical return distributions. Cont (2001) reports normalized skewness values between -0.4 and -0.1 for five-minute returns of three different financial instruments. Cont describes it as a "gain/loss asymmetry", because "one observes large drawdowns in stock prices and stock index values but not equally large upward movements" (Cont, 2001, p. 224).

However, it has been pointed out by Kim and White (2004) and Bonato (2011) that standard estimators of skewness may be likely to overestimate the magnitude of skewness when dealing with return series. We expand briefly on this topic in section 2.4.2 below. It seems like a sufficiently conservative statement that returns skewness should probably be below zero, and probably not much smaller than -0.5.

2.4 Limitations of the stylized facts

The stylized facts mentioned above are among the least controversial: they have been demonstrated in different asset types, at different time scales, and in different time periods (Cont, 2001). Nonetheless we want to mention a couple of caveats at this point.

2.4.1 Time scales

The first complication is the time scale. Many of the stylized facts have been demonstrated in return series at different time scales, from seconds and minutes up to days or even weeks, i.e., ranging over five or six orders of magnitude. But things do change with the time scale. For example, according to Cont (2001) the return distribution looks more and more like a normal distribution as the time step Δt is increased. This effect is illustrated in figure 2.3 above.

More generally, one can find quite different results by sampling return series using different "clocks". One method is to sample transaction prices with a constant time interval, as we have assumed so far. This "calendar time" is the most common (Chakraborti et al., 2011a). But one could instead sample at every placement and cancellation of orders in the electronic order book to get a series in "event time", or at every transaction to get a series in "trade time", or sample only when the price changes to get a series in "tick time" (Chakraborti et al., 2011a).

Without digging deeper at this point, we note that revisiting the stylized facts with a new clock can give additional insight into return series. This could be an essential consideration in designing mathematical models for explaining the stylized facts.

2.4.2 Existence of moments

Another caveat is that the stylized facts have been established using statistical estimation methods which may not accurately describe the "true" underlying values. Specifically, assuming that the data generating process for a random variable X does not have a kth moment, i.e., the integral $\int \left|x^k\right| f(x)\mathrm{d}x$ diverges, the sample estimator based on empirical data is perfectly computable but will supposedly exhibit erratic behavior (Bonato, 2011).

The behavior of conventional estimators of skewness and kurtosis when the data generator process is a distribution which does not possess second or third or fourth moment has been investigated by Kim and White (2004) and Bonato (2011). One result was that for a symmetric, fat-tailed distribution sample, skewness is not a valid indicator of the presence of asymmetry.

This point is highly pertinent to our work since the return distributions are fat tailed according to a basic stylized fact. Recall that the tail index of a distribution sets an upper limit on highest finite moment. Since the tail index in empirical distributions are estimated at $\alpha \approx 3$, it seems likely that the fourth moment is infinite, and perhaps the third moment too. Hence, any estimate of skewness (third moment) or kurtosis or autocorrelation of squared returns (both depend on the fourth moment), or higher moments should also be interpreted with great care. However, according to Cont (2001), a systematic analysis of a wide range of US and French stocks indicates that typical returns series at least have finite variance (second central moment).

2.5 Reproducing the stylized facts: modeling approaches

A number of different modeling approaches have been used to replicate or explain various stylized facts about financial markets. We do not attempt to make an overview here, but merely point to a few examples.

2.5.1 The ARCH family

A large literature in econometrics uses variants of a model proposed by Engle (1982), originally designed to reproduce the time-varying volatility of macroeconomic variables. Engle's model class, called ARCH (autoregressive conditional heteroskedasticity), was later generalized into a class now known as GARCH (a textbook introduction is given by Davidson and MacKinnon (2004, pp. 587–595)).

The GARCH class is merely a mathematical structure and the models provide no explanation whatsoever to why the stylized facts arise. In any case, it is an interesting class because there are plenty of theoretical results in the literature, and the GARCH model, and different variants of it, has later been explicitly applied to financial return series with some success, e.g., by Ding and Granger (1996).

A fundamental shortcoming of the GARCH models is that they produce exponentially decaying squared autocorrelations, i.e., much faster decaying than the empirically observed power law behavior.

2.5.2 Agent-based order book dynamics

There is also a large number of more concrete, descriptive models. Some of them focus on order book dynamics, i.e., the placement and cancellation of buy and sell orders in an electronic order book which affect the market price. Order book dynamics are not considered elsewhere in this thesis, but it should be noted that several interesting statistical regularities about order flow have been identified (Chakraborti et al., 2011a), which inform the design of order book models.

A simple example was presented by Challet and Stinchcombe (2001), who modeled the order book with a "zero-intelligence" order flow similar to a physical deposition/evaporation process. A more elaborated model by Mike and Farmer (2008) takes a more concrete approach to trader behavior as a function of the current market conditions. The model was calibrated against empirical data for one stock and rather successfully validated against 24 other stocks on the same market. Several other order book models can be found in the review by Chakraborti et al. (2011b) and references therein.

2.5.3 Heterogeneous agents in aggregated market models

Finally, we mention a rather specific class of agent-based models which is interesting because there are several similar such models in the literature which have successfully reproduced some of the stylized facts, such as clustered volatility and fat tails of the return distribution. Our work, presented in chapter 3, contains results for one model of this class. These models build on the notion of different trading strategies that agents employ to determine their demand for an asset. The market is described in a more aggregated, abstract way than in the order book models, using a discrete-time difference equation for the price increments,

$$p_{t+1} = p_t + f(D_t),$$

where p_t is the price at time t, $f(\cdot)$ is a function with positive slope and $f(0) = 0$ (many models use $f(x) = kx$ for a constant k), and D_t is the net demand at time t. Hence, the price increases (decreases) if the net demand is positive (negative).

The interesting part of these models is the function D_t, which is in general a nonlinear, stochastic function of the previous prices $p_t, p_{t-1}, \ldots$. Various functional forms may be motivated in various ways, but a common factor to several models is the concept of combining "fundamentalist" agents trading on the belief that prices will eventually return to a certain "correct" price level motivated by the fundamental value of the asset, and technical traders or "chartists" betting that prices follow some form of trend. The concept of combining fundamentalist and technical traders is consistent with questionnaire data on trading strategies collected among financial professionals, reviewed, e.g., by Hommes (2006, pp. 1118–1122). According to these surveys, professional traders act both on the long-run expectation that prices return to some fundamental level, and on the short-run expectation that prices can be predicted based on recent trends.

Such models with different forms of the demand function have been proposed, e.g., by Day and Huang (1990); Farmer and Joshi (2002); Tramontana, Westerhoff, and Gardini (2010), and Westerhoff and Franke (2012), all similar in some sense, but exhibiting different behaviors according the functional form chosen for D_t. The models just mentioned have constant populations of traders following different strategies, while other models allow agents to switch their investment rules (e.g., Grauwe and Dewachter, 1993; Brock and Hommes, 1998; Chiarella, He, and Hommes, 2006; Alfi et al., 2009). A couple of the models (e.g., Grauwe and Dewachter, 1993; Brock and Hommes, 1998) are deterministic while most others have stochastic elements.

All these models have been demonstrated by the authors to reproduce some stylized facts of financial markets, specifically lack of returns autocorrelation, fat tails of the return distribution, and volatility clustering. However, the matching of stylized facts is in many cases only qualitative, e.g., squared or absolute returns are autocorrelated, but not necessarily according to a power law; the return distribution has fat tails, but not necessarily with a tail index close to empirical observations. Among the most quantitative attempts at matching results we have encountered were presented by Westerhoff and Franke (2012) in the model we analyze in chapter 3. Westerhoff and Franke tried to quantitatively match the tail index and absolute returns autocorrelations against empirical values with some success.

Chapter 3

Analysis of a stochastic model

This chapter contains an investigation of a stochastic asset pricing model proposed by Westerhoff and Franke (2012). They demonstrated using computer simulations that the model can roughly match some empirical observations in daily returns series from a stock market and a foreign exchange market (the S&P 500 index and the USD-DEM exchange rate), specifically autocorrelations in returns and the Hill estimator of tail index. Their simulation results also have the first 100 autocorrelations in absolute returns in the right order of magnitude, although it is not clear from their simulation results how well the model matches the power law decay we expect based on the stylized facts. Figure 3.1 shows simulation results corresponding to the parameters in the stock market scenario from the original paper (Westerhoff and Franke, 2012, p. 428).

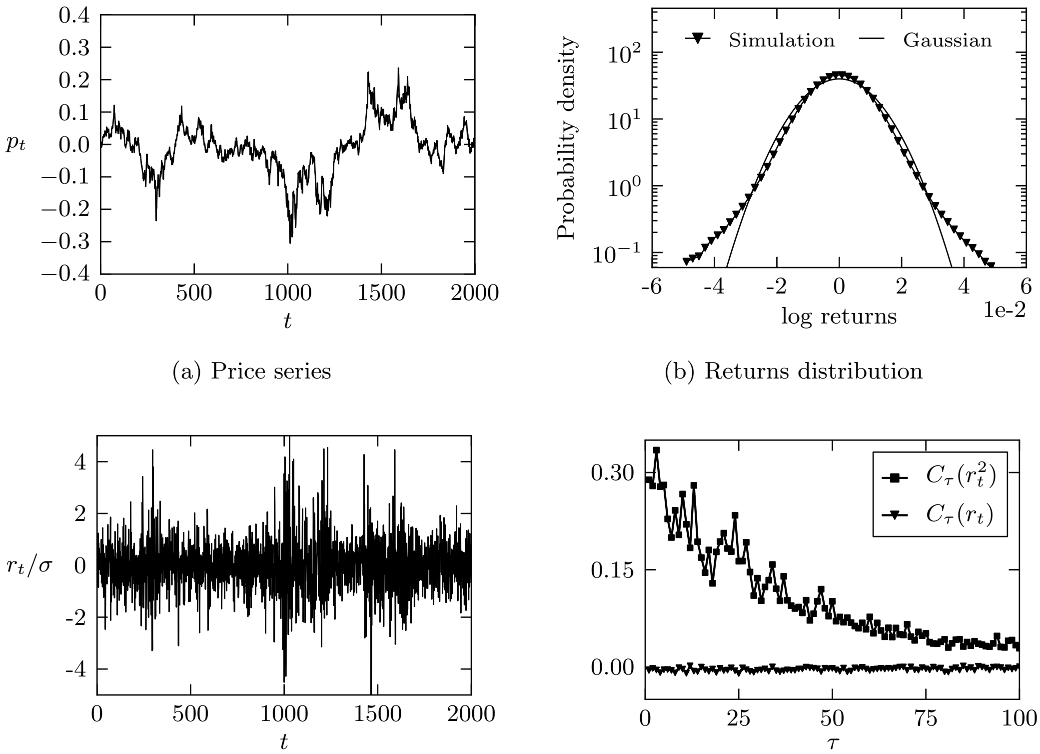
However, the results presented by Westerhoff and Franke are only simulation results from two different points in the parameter space. This naturally raises questions about parameter sensitivity. How does a change in parameter values affect the results? In which parts of the parameter space does the model reproduce the stylized facts?

This thesis provides partial answers to those questions using a combination of theoretical analysis and numerical simulation. First, by putting some stationarity requirements on the process $\{p_t\}$, some constraints on the model parameters are derived theoretically, which restricts the analysis to a smaller part of the parameter space. Second, analytical expressions are derived for returns skewness and kurtosis, autocorrelation of returns, and autocorrelation of squared returns. Matching these results against the stylized facts, even further constraints can be put on the parameters. Finally, computer simulations are used to analyze how the Hill tail estimator varies in the relevant parameter space.

The remainder of this chapter is organized as follows. In sections 3.1 and 3.2, the model by Westerhoff and Franke is presented and simplified. Section 3.3 provides the theoretical basis for computing moments of the stochastic process, and presents derivations of some quantities. Section 3.4 briefly discusses numerical simulation of the model. Finally, the results are summarized and discussed in section 3.5.

3.1 The original model formulation

The model proposed by Westerhoff and Franke is a stochastic difference equation, motivated by the notion of three trader types. First, the fundamentalists who think that the asset has a constant fundamental value p^* and whose demand at time t is proportional to the perceived mispricing $(p^* - p_t)$. A fundamentalist who trades at



(a) Price series

(b) Returns distribution

(c) Returns series

(d) Autocorrelations

Figure 3.1: Simulation results of the stock market scenario in the model proposed by (Westerhoff and Franke, 2012, p. 428). The price and returns series displayed is only a short subsample at the end of a long simulation. The return distribution and autocorrelation functions were estimated based on a single simulation of 10^6 time steps, discarding the first $5 \cdot 10^5$ to eliminate any transient effects. The parameters transformed to our simplified model are $a = 0.989, \sigma_X = 0.101, b = 0$. The empirical distributions are Epanechikov kernel estimates using a bandwidth choice suggested by Davidson and MacKinnon (2004, p. 681, equation (15.63)).

time t makes the order

$$d_t^{\mathrm{F}} = \phi(p^* - p_t),$$

where ϕ is a positive constant.

Second, there are chartists who bet on the opposite, that any mispricing will increase. A chartist who trades at time t makes the order

$$d_t^{\mathrm{C}} = \chi(p_t - p^*),$$

where χ is a positive constant.

Third and last, investors are traders who act independently of any price signals. An investor is either a buyer or seller, who if active at time t makes the order, respectively,

$$\begin{aligned} d_t^{\mathrm{B}} &= \kappa, \\ d_t^{\mathrm{S}} &= -\kappa, \end{aligned}$$

for a positive κ.

The stochastic elements of the model are motivated by each trader type having a finite population where each individual decides to trade with a certain probability each time step. For example, the population of fundamentalists is N_{F} and they all independently decide with probability π_{F} to trade in each time step. Hence, the number of active traders of each type at time t is binomially distributed,

$$\begin{aligned} F_t &\sim \mathrm{Bin}(N_{\mathrm{F}}, \pi_{\mathrm{F}}), \\ C_t &\sim \mathrm{Bin}(N_{\mathrm{C}}, \pi_{\mathrm{C}}), \\ B_t &\sim \mathrm{Bin}(N_{\mathrm{B}}, \pi_{\mathrm{B}}), \\ S_t &\sim \mathrm{Bin}(N_{\mathrm{S}}, \pi_{\mathrm{S}}), \end{aligned}$$

where F denotes fundamentalists, C chartists, B buyers and S sellers.

Finally, it is assumed that the price change between time t and $t+1$ is proportional to the total demand with a constant factor μ. The complete model is written

$$p_{t+1} = p_t + \mu\left[(\chi C_t - \phi F_t)(p_t - p^*) + \kappa(B_t - S_t)\right] \tag{3.1}$$

3.2 The simplified model

First, an approximation will be used to simplify the model. The binomially distributed variables F_t, C_t, B_t, S_t can be approximated by normally distributed variables. The approximation $\mathrm{Bin}(n, \pi) \approx \mathcal{N}(n\pi, n\pi(1-\pi))$ is exact in the limit $n \to \infty$ and it is quite good already for small n if π is not too close to 0 or 1 (J. S. Hunter, W. G. Hunter, and Box, 2005). In fact, the same approximation was used in the simulations presented in the the original paper by Westerhoff and Franke (2012).

Using the normal approximation, it is easily seen that

$$\begin{aligned} \mu(\chi C_t - \phi F_t) &\text{ is approximately distributed as } && \mathcal{N}\left(\mu_X, \sigma_X^2\right) \\ \kappa(B_t - S_t) &\text{ is approximately distributed as } && \mathcal{N}\left(\mu_Y, \sigma_Y^2\right) \end{aligned}$$

where

$$\mu_X = \mu\left(\pi_C \chi N_C - \pi_F \phi N_F\right),$$
$$\sigma_X^2 = \mu^2 \left(\left(1-\pi_C\right)\pi_C \chi^2 N_C + \left(1-\pi_F\right)\pi_F \phi^2 N_F\right),$$
$$\mu_Y = \kappa\mu\left(\pi_B N_B - \pi_S N_S\right),$$
$$\sigma_Y^2 = \kappa^2\mu^2\left(\left(1-\pi_B\right)\pi_B N_B + \left(1-\pi_S\right)\pi_S N_S\right).$$

This way, equation (3.1) can be approximated by

$$p_{t+1} = p_t + \left(\sigma_X X_t + \mu_X\right)\left(p_t - p^*\right) + \left(\sigma_Y Y_t + \mu_Y\right),$$

where $X_t \sim \mathcal{N}(0,1)$ and $Y_t \sim \mathcal{N}(0,1)$.

Second, two of the parameters can be canceled through a suitable variable change. Simply subtract p^* from both sides of equation (3.1) and then divide by σ_Y to find

$$\frac{p_{t+1} - p^*}{\sigma_Y} = \frac{p_t - p^*}{\sigma_Y} + (\sigma_X X_t + \mu_X)\frac{p_t - p^*}{\sigma_Y} + Y_t + \frac{\mu_Y}{\sigma_Y},$$

or, with $p_t' = (p_t - p^*)/\sigma_Y$,

$$p_{t+1}' = p_t' + (\sigma_X X_t + \mu_X)p_t' + Y_t + \frac{\mu_Y}{\sigma_Y},$$

The dynamics in p_t' have only been shifted by the constant term $-p^*$ and scaled by $1/\sigma_Y$, which does not matter for any of the results in this thesis, since the results presented below only concern normalized central moments of the return distribution, autocorrelations in returns and squared returns, and the Hill tail estimator of the return distribution. All of these are independent of price scaling and translation, as is easily verified using the definitions given in section 2.1.

The simplified model is finally obtained by dropping the apostrophe on p_t', collecting terms and setting $a = 1 + \mu_X$ and $b = \mu_Y/\sigma_Y$. Then,

$$p_{t+1} = p_t\left(a + \sigma_X X_t\right) + Y_t + b, \tag{3.2}$$

where X_t and Y_t are random variables of the standard normal distribution $\mathcal{N}(0,1)$, a and b are real-valued constants, and σ_X is a non-negative real constant. (We could allow σ_X to be negative, but since X_t is symmetrically distributed, this does not change anything.)

3.3 Theoretical analysis of the simplified model

In this section, we derive a few theoretical results for the simplified model in equation (3.2) above. Most notably, we derive the skewness and kurtosis of the return distribution, and the expected autocorrelation of returns and squared returns.

We begin with a few theoretical considerations in sections 3.3.1 – 3.3.3, then derive the actual results in sections 3.3.4 – 3.3.4. The results are summarized and discussed in section 3.5.

3.3.1 The general principle

The following calculations are essentially motivated by two facts about expected values. First, expectation is linear, and second, the expectation is multiplicative for independent random variables.

$$\mathbb{E}[A+B] = \mathbb{E}[A] + \mathbb{E}[B] \text{ for all } A, B,$$
$$\mathbb{E}[AB] = \mathbb{E}[A]\,\mathbb{E}[B] \text{ if } A \text{ and } B \text{ are independent.}$$

Now, some of the quantities we are interested in, such as the autocorrelation of returns, and skewness and kurtosis of returns, can be written as a linear combination of expressions of the form

$$\mathbb{E}\left[\prod_{i=0}^{m} p_{t+i}^{\nu_i}\right], \tag{3.3}$$

where ν_i are non-negative integer constants. In fact, any moment (or central moment) of prices or returns, or powers of those (including mixed moments) can be written in this form.

Furthermore, any expression of the form (3.3) can be rewritten using the model equation (3.2) repeatedly until only powers of p_t remain. Using linearity of expectation and multiplicativity of independent expectations, the general result is

$$\mathbb{E}\left[\prod_{i=0}^{m} p_{t+i}^{\nu_i}\right] = \sum_{n=0}^{N} \gamma_n \mathbb{E}[p_t^n],$$

where γ_n are constants and the largest possible exponent is $N = \sum_{i=0}^{m} \nu_i$. (For a proof, see appendix A.1.) Hence, the expectation (3.3) exists, and can be computed, if also $\mathbb{E}[p_t^n]$ exists and can be computed for all $n = 1, \ldots, N$.

3.3.2 Existence and derivation of moments

As previously discussed in section 2.4.2, any talk about moments of the return distribution should be careful, because it is not obvious that moments exist, much less that they are constant.

In our model, the moments $\mathbb{E}[p_t^n]$ for positive integers n do not necessarily exist, and if they exist, they generally depend on t. However, in slightly vague terms it can be said that moments are indeed constant and finite assuming that the process has been going on for a long time, and that the parameter values follow some restrictions.

To be specific, a theoretical investigation by Vervaat (1979) shows that the price process p_t in the model does converge in distribution as $t \to \infty$ under sufficiently strict parameter restrictions, and more importantly that the moments $\mathbb{E}[p_t^n]$ tend to finite constants as $t \to \infty$ under less strict restrictions on the parameters. A less formal explanation can be made as follows.

From here on, we will assume that the price process p_t is defined for non-negative integers $t \in \{0, 1, 2, \ldots\}$. The expectations $\mathbb{E}[p_t^n]$ for positive integers n can then be computed through a general algorithm. We show in appendix A.3 that for a positive integer t, the expectations $\mathbb{E}[p_t^n]$ and $\mathbb{E}[p_0^n]$ always have a simple relationship if they exist, namely

$$\mathbb{E}[p_t^n] = C_n^t \mathbb{E}[p_0^n] + D_n \sum_{\nu=0}^{t-1} C_n^{\nu},$$

where C_n and D_n are constants that depend on the parameter values and the power n. Specifically, $C_n = C_n(a, \sigma_X)$ and $D_n = D_n(a, b, \sigma_X)$. Assuming $\mathbb{E}[p_0^n] < \infty$ and requiring that $\mathbb{E}[p_t^n] < \infty$ leads to the restrictions $|C_n| < 1$ and $D_n < \infty$. Under those conditions, $\mathbb{E}[p_t^n]$ converges exponentially in absolute value to a constant,

$$\lim_{t\to\infty} \mathbb{E}[p_t^n] = \frac{D_n}{1 - C_n}. \tag{3.4}$$

Since the expectation $\mathbb{E}[p_t^n]$ converges, we can loosely say that it is constant if the price process has been going on forever. From here on, we will assume that the process p_t has moments independent of t for all positive integers n whenever the limit $\lim_{t\to\infty} \mathbb{E}[p_t^n]$ exists, if nothing else is stated.

For a complete description of the steps taken to compute the moments $\mathbb{E}\left[p_t^k\right]$, refer to appendix A.3.

3.3.3 Parameter restrictions for stationarity

In conclusion, we require $|C_n| < 1$ and $D_n < \infty$. We show in appendix A.3 that the constant D_n is finite if and only if all $|C_k| < 1$ for $k = 1, 2, \ldots, (n-1)$. Hence, to compute the expectation

$$\mathbb{E}\left[\prod_{i=0}^{m} p_{t+i}^{\nu_i}\right],$$

where the powers ν_i sum to $N = \sum_i \nu_i$, it is required that

$$|C_n| < 1 \text{ for } n = 1, 2, \ldots, N.$$

The constants C_n depend on a and σ_X, so we have to impose some restrictions on a and σ_X. The restrictions (derived in appendix A.1) are

$$|C_n| = \left| a^n + \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{2k} a^{n-2k} \sigma_X^{2k} (2k-1)!! \right| < 1 \text{ for } n = 1, 2, \ldots, N. \tag{3.5}$$

Note that σ_X^{2k} is always positive and a^{n-2k} has the same sign for all integers n and k. Therefore all terms in this sum have the same sign, so it is equivalent to require

$$|C_n| = |a|^n + \sum_{k=1}^{\lfloor n/2 \rfloor} \binom{n}{2k} |a|^{n-2k} \sigma_X^{2k} (2k-1)!! < 1 \text{ for } n = 1, 2, \ldots, N.$$

The first four C_n are

$$\begin{aligned}
C_1 &= a, \\
C_2 &= a^2 + \sigma_X^2, \\
C_3 &= a^3 + 3a\sigma_X^2, \\
C_4 &= a^4 + 6a^2\sigma_X^2 + 3\sigma_X^4.
\end{aligned}$$

3.3.4 Calculation of some quantities

Following the general principles established in previous sections, it is possible to calculate closed-form expressions for the theoretical mean of several quantities associated with the stylized facts. In the following pages, we present calculations of the expected returns, skewness and kurtosis of the return distribution, and autocorrelation in returns and squared returns.

Expected price and expected returns

From sections 3.3.2 and 3.3.3 above, it follows that the expected returns $\mathbb{E}[r_t] = \mathbb{E}[p_t - p_{t-1}]$ exists if $|a| < 1$. In that case $\mathbb{E}[p_t]$ is a constant, independent of t, and

$$\mathbb{E}[r_t] = \mathbb{E}[p_t - p_{t-1}] = \mathbb{E}[p_t] - \mathbb{E}[p_{t-1}] = 0 \qquad |a| < 1.$$

In other words, the expected price is constant for $|a| < 1$ and hence expected returns are always zero. For completeness, we still derive the expected price $\mathbb{E}[p_t]$ as follows.

First, we note using linearity of expectation and multiplicativity of independent expectations, and $\mathbb{E}[X_t] = \mathbb{E}[Y_t] = 0$, that

$$\begin{aligned}\mathbb{E}[p_{t+1}] &= \mathbb{E}[p_t(a + \sigma_X X_t) + Y_t + b] \\ &= a\mathbb{E}[p_t] + \sigma_X \mathbb{E}[p_t]\,\mathbb{E}[X_t] + \mathbb{E}[Y_t] + b \\ &= a\mathbb{E}[p_t] + b.\end{aligned}$$

Using the same relation again, we see

$$\mathbb{E}[p_{t+2}] = a^2\mathbb{E}[p_t] + ab + b,$$

and so on. Iterating N times yields

$$\mathbb{E}[p_{t+N}] = a^N\mathbb{E}[p_t] + b\sum_{\nu=0}^{N-1} a^\nu.$$

In the limit $N \to \infty$, this expression is only convergent for $|a| < 1$, and then

$$\lim_{N\to\infty} \mathbb{E}[p_{t+N}] = \lim_{N\to\infty} a^N\mathbb{E}[p_t] + b\frac{1-a^N}{1-a} = \frac{b}{1-a}, \qquad |a| < 1. \tag{3.6}$$

For $|a| \geq 1$, the limit $\lim_{N\to\infty} \mathbb{E}[p_{t+N}]$ does not exist, so there is no unconditional expectation $\mathbb{E}[p_t]$.

Expected squared returns

Similarly, we will show that the expected squared returns $\mathbb{E}\left[r_t^2\right] = \mathbb{E}\left[(p_t - p_{t-1})^2\right]$ exists if $|a| < 1$ and $\left|a^2 + \sigma_X^2\right| < 1$.[1] To compute expected squared returns, we first use linearity of expectation to find

$$\mathbb{E}\left[r_t^2\right] = \mathbb{E}\left[(p_t - p_{t-1})^2\right] = \mathbb{E}\left[p_t^2\right] - 2\mathbb{E}[p_t p_{t-1}] + \mathbb{E}\left[p_{t-1}^2\right],$$

where $\mathbb{E}[p_t p_{t-1}]$ can be rewritten using the model equation so that

$$\begin{aligned}\mathbb{E}\left[r_t^2\right] &= \mathbb{E}\left[p_t^2\right] - 2\mathbb{E}\left[p_{t-1}^2(a + \sigma_X X_t) + p_{t-1}(b + Y_t)\right] + \mathbb{E}\left[p_{t-1}^2\right] \\ &= \mathbb{E}\left[p_t^2\right] - 2a\mathbb{E}\left[p_{t-1}^2\right] + b\mathbb{E}[p_{t-1}] + \mathbb{E}\left[p_{t-1}^2\right].\end{aligned}$$

We know that $\mathbb{E}\left[p_t^2\right]$, if it exists, is independent of t, so $\mathbb{E}\left[p_{t-1}^2\right] = \mathbb{E}\left[p_t^2\right]$. Using this, and the expectation $\mathbb{E}[p_t]$ from equation (3.6) (assuming $|a| < 1$),

$$\mathbb{E}\left[r_t^2\right] = 2\left((1-a)\mathbb{E}\left[p_t^2\right] - b\mathbb{E}[p_{t-1}]\right) = \frac{2b^2}{a-1} - 2(a-1)\mathbb{E}\left[p_t^2\right] \tag{3.7}$$

Now, we just have to find $\mathbb{E}\left[p_t^2\right]$. Applying the model equation once, using additivity and multiplicativity, and $\mathbb{E}[X_t] = \mathbb{E}[Y_t] = 0$, we find

$$\begin{aligned}\mathbb{E}\left[p_{t+1}^2\right] &= \mathbb{E}\left[(p_t(a + \sigma_X X_t) + Y_t + b)^2\right] \\ &= 1 + b^2 + 2ab\mathbb{E}[p_t] + \left(a^2 + \sigma_X^2\mathbb{E}\left[X_t^2\right]\right)\mathbb{E}\left[p_t^2\right].\end{aligned}$$

[1] Of course, $\left|a^2 + \sigma_X^2\right| < 1$ implies $|a| < 1$, but in principle it has to be confirmed that all $|C_k| < 1$ for $k = 1, 2, \ldots, n$ to guarantee the existence of $\mathbb{E}[p_t^n]$.

Since $X_t \sim \mathcal{N}(0,1)$, the random variable X_t^2 is χ^2-distributed with one degree of freedom, so $\mathbb{E}\left[X_t^2\right] = 1$. More generally, it is shown in appendix A.2 that

$$\mathbb{E}\left[X_t^k\right] = \mathbb{E}\left[Y_t^k\right] = \begin{cases} (k-1)!! \text{ for even } k, \\ 0 \text{ for odd } k. \end{cases} \tag{3.8}$$

Using equation (3.6) and equation (3.8), we find

$$\mathbb{E}\left[p_{t+1}^2\right] = 1 + b^2 \frac{1+a}{1-a} + \left(a^2 + \sigma_X^2\right) \mathbb{E}\left[p_t^2\right].$$

Applying this relation iteratively, as in the derivation of $\mathbb{E}[p_t]$, yields

$$\mathbb{E}\left[p_{t+N}^2\right] = \left(a^2 + \sigma_X^2\right)^N \mathbb{E}\left[p_t^2\right] + \left(1 + b^2 \frac{1+a}{1-a}\right) \sum_{\nu=0}^{N-1} \left(a^2 + \sigma_X^2\right)^\nu.$$

With the same logic as before, it is easily seen that the expectation $\mathbb{E}\left[p_t^2\right]$ exists only if $\left|a^2 + \sigma_X^2\right| < 1$ and it is

$$\mathbb{E}\left[p_t^2\right] = \frac{a\left(b^2 - 1\right) + b^2 + 1}{(a-1)\left(a^2 + \sigma_X^2 - 1\right)}, \qquad \left|a^2 + \sigma_X^2\right| < 1. \tag{3.9}$$

The final result is obtained by inserting equation (3.9) in equation (3.7),

$$\mathbb{E}\left[r_{t+1}^2\right] = \frac{2\left((a-1)^2 + b^2 \sigma_X^2\right)}{(a-1)\left(a^2 + \sigma_X^2 - 1\right)}, \qquad \left|a^2 + \sigma_X^2\right| < 1. \tag{3.10}$$

Returns skewness and kurtosis

The returns skewness and kurtosis can be calculated using the same sort of procedure used for expected returns and expected squared returns in sections 3.3.4 and 3.3.4 above.

However, the exercise is a bit more laborious since returns skewness and kurtosis require the third and fourth moments $\mathbb{E}\left[p_t^3\right]$ and $\mathbb{E}\left[p_t^4\right]$ respectively. Expanding $\mathbb{E}\left[p_{t+1}^4\right] = \mathbb{E}\left[\left(p_t(a + \sigma_X X_t - 1) + Y_t + b\right)^4\right]$ obviously yields a relatively large number of terms, which would be uncomfortable to handle manually. The results presented below have instead been calculated using the symbolic mathematics software Wolfram Mathematica, by implementing the rules for linearity of expectation, multiplicativity of expectation of independent variables, expected values $\mathbb{E}\left[X_t^k\right]$, etc.

The resulting expression for the normalized skewness of the return distribution is

$$\frac{\mathbb{E}\left[r_t^3\right]}{\mathbb{E}\left[r_t^2\right]^{3/2}} = \frac{3(2a+1)b\sigma_X^2}{\sqrt{2}\left(a^3 + 3a\sigma_X^2 - 1\right)\sqrt{\frac{(a-1)^2 + b^2\sigma_X^2}{(a-1)\left(a^2 + \sigma_X^2 - 1\right)}}}$$

The returns kurtosis is given in appendix A.5.

Autocorrelation of returns

By definition, autocorrelation of returns is

$$C_\tau(r_t) = \frac{\mathbb{E}\left[\left(r_{t+\tau} - \mathbb{E}[r_{t+\tau}]\right)\left(r_t - \mathbb{E}[r_t]\right)\right]}{\sqrt{\mathbb{E}\left[\left(r_{t+\tau} - \mathbb{E}[r_{t+\tau}]\right)^2\right] \mathbb{E}\left[\left(r_t - \mathbb{E}[r_t]\right)^2\right]}},$$

which immediately can be simplified since the moments of r_t are independent of t, and $\mathbb{E}[r_t] = 0$. A simpler expression is

$$C_\tau(r_t) = \frac{\mathbb{E}[r_{t+\tau} r_t]}{\left|\mathbb{E}\left[r_t^2\right]\right|}.$$

Setting $\tau = 1$, the calculation of this quantity is straightforward using the principles established in sections 3.3.1 and 3.3.3. The result is

$$C_1(r_t) = -\frac{1-a}{2} \qquad \text{if } \left|a^2 + \sigma_X^2\right| < 1.$$

Now, the returns autocorrelation at any lag $\tau \geq 1$ can be found by noting that

$$\begin{aligned}
\frac{C_{\tau+1}(r_t)}{C_\tau(r_t)} &= \frac{\mathbb{E}[r_{t+\tau+1} r_t]}{\mathbb{E}[r_{t+\tau} r_t]} = \frac{\mathbb{E}[(p_{t+\tau+1} - p_{t+\tau}) r_t]}{\mathbb{E}[r_{t+\tau} r_t]} \\
&= \frac{\mathbb{E}[(p_{t+\tau}(a + \sigma_X X_{t+\tau}) + Y_{t+\tau} - p_{t+\tau-1}(a + \sigma_X X_{t+\tau-1}) - Y_{t+\tau-1}) r_t]}{\mathbb{E}[r_{t+\tau} r_t]} \\
&= \frac{a\mathbb{E}[r_{t+\tau} r_t] + \mathbb{E}[(\sigma_X (X_{t+\tau} - X_{t+\tau-1}) + Y_{t+\tau} - Y_{t+\tau-1})\, r_t]}{\mathbb{E}[r_{t+\tau} r_t]} \\
&= a,
\end{aligned}$$

where the last equality holds since $X_{t+\tau-1}$ and $Y_{t+\tau-1}$ are independent of r_t when $\tau \geq 1$. In conclusion,

$$C_{\tau+1}(r_t) = aC_\tau(r_t), \qquad \tau \geq 1,$$

so the general formula for autocorrelations at lags $\tau = 1, 2, 3, \ldots$ is

$$C_\tau(r_t) = -a^{\tau-1}\frac{1-a}{2}, \qquad \left|a^2 + \sigma_X^2\right| < 1. \tag{3.11}$$

Autocorrelation of squared returns

The autocorrelation of squared returns,

$$C_\tau(r_t^2) = \frac{\mathbb{E}\left[\left(r_{t+\tau}^2 - \mathbb{E}\left[r_{t+\tau}^2\right]\right)\left(r_t^2 - \mathbb{E}\left[r_t^2\right]\right)\right]}{\sqrt{\mathbb{E}\left[\left(r_{t+\tau}^2 - \mathbb{E}\left[r_{t+\tau}^2\right]\right)^2\right]\mathbb{E}\left[\left(r_t^2 - \mathbb{E}\left[r_t^2\right]\right)^2\right]}},$$

is calculated the same way as the quantities found in previous sections. The autocorrelation of squared returns only exists if all the first four price moments up to $\mathbb{E}\left[p_t^4\right]$ exist.

An expression for the first lag autocorrelation $C_1(r_t^2)$ is found in appendix A.5.

When it comes to higher lags $\tau \geq 2$, things get slightly more complicated. It can be shown that choosing the parameter $b = 0$ yields an exponential decay similar to the autocorrelation of returns,

$$C_{\tau+1}(r_t^2) = (a^2 + \sigma_X^2) C_\tau(r_t^2), \qquad b = 0, \quad \tau \geq 1.$$

However, when $b \neq 0$, the decay is generally not exponential, but

$$C_\tau(r_t^2) = \left(K_1 a^{\tau-1} + K_2 \left(a^2 + \sigma_X^2\right)^{\tau-1}\right) C_1(r_t^2), \qquad \tau \geq 1. \tag{3.12}$$

with constants K_1 and K_2 given in appendix A.5.

Note that if $a = a^2 + \sigma_X^2$, the autocorrelation decay is exponential. In the other case ($a \neq a^2 + \sigma_X^2$), the decay of the squared returns autocorrelation also tends to an exponential (at least in absolute value) as $\tau \to \infty$, since one of the terms in the sum $K_1 a^{\tau-1} + K_2 \left(a^2 + \sigma_X^2\right)^{\tau-1}$ vanishes faster than the other.

The calculations behind these results are found in appendix A.4, and a complete expression for $C_\tau(r_t^2)$ is given in appendix A.5.

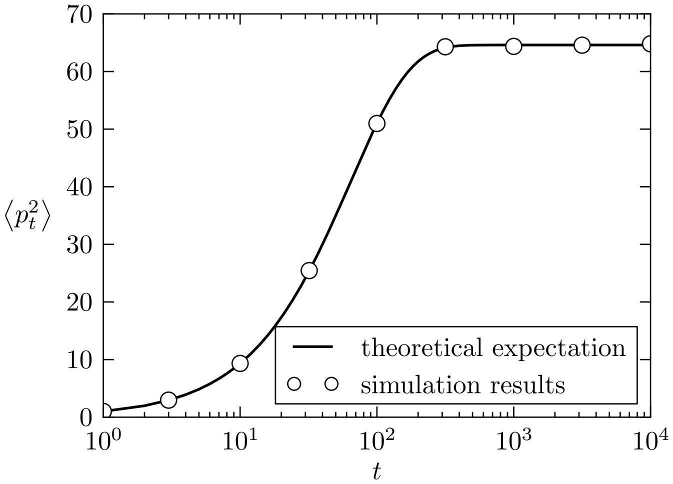


Figure 3.2: In numerical simulations, the moments of the price process converge predictably, if they are finite. The figure shows the theoretical convergence of the second price moment $\mathbb{E}\left[p_t^2\right]$ towards its stationary value, confirmed by simulation results. Simulation results are averages of p_t^2 over 10^6 independent runs, started with $p_0 = 0$. Parameter values: $a = 0.989$, $\sigma_X = 0.08$, $b = 0$.

3.4 Numerical simulation

Numerical simulation of the price process is straightforward. After choosing an initial price p_0, obtaining a finite series $p_1, p_2, \ldots, p_N$ is simply a matter of iterating the model equation N times with random numbers X_t and Y_t drawn independently from the standard normal distribution.

When interpreting the simulation results, however, there are two important points to note. First, to ensure that moments of interest have converged, a suitable number of time steps should be discarded from the start of the simulation. Second, it should be noted that estimators of moments may have infinite variance. These two points are elaborated in sections 3.4.1 and 3.4.2 below.

3.4.1 Convergence of moments

It is known from equation (3.4) that the price moments, if they exist, converge exponentially toward constant values. A practical implication is that numerical simulations generally need to run for some time before a given moment can be reliably estimated from simulation results.

For simplicity, numerical simulations were always started with the first price value $p_0 = b/(1-a)$, i.e. equal to the theoretical stationary moment $\lim_{t\to\infty} \mathbb{E}[p_t]$. From the calculations above it is then clear that for all times $t = 1, 2, \ldots, N$, the first moment is always equal to its theoretical value.

However, since the first value p_0 is a constant, all higher moments are initially equal to some constant which is generally not the theoretical moment, i.e. $\mathbb{E}\left[p_0^k\right] = p_0^k$ for $k \geq 2$, and therefore the moments $\mathbb{E}\left[p_t^k\right]$ are generally not independent of t. The dependence of the moments $\mathbb{E}\left[p_t^k\right]$ is explicitly given in equation (3.4), and from there it is easily computed how many time steps are needed to obtain a given relative error of the finite-time moment $\mathbb{E}\left[p_t^k\right]$ compared to the infinite-time limit. This principle is illustrated with the second moment as an example in figure 3.2.

3.4.2 Variance of estimators

An estimate of the kth moment of the price at time t can be computed by simulating N independent price series to get a set of realizations of p_t, i.e. $\{p_t(i)\}_{i=1}^N$. The estimate is then computed as

$$\widehat{\mathbb{E}\left[p_t^k\right]} = \frac{1}{N}\sum_{i=1}^{N}\left(p_t(i)\right)^k .$$

The variance of the estimator can be computed by noting that all realizations $p_t(i)$ are identically distributed so that $\mathbb{E}[p_t(i)] = \mathbb{E}[p_t]$, and also independently distributed, so all the mixed terms simplify as $\mathbb{E}\left[p_t(i)^k p_t(j)^k\right] = \mathbb{E}\left[p_t(i)^k\right]^2$. Keeping this in mind, it is seen that

$$\begin{aligned}
\mathrm{Var}\left(\widehat{\mathbb{E}\left[p_t^k\right]}\right) &= \mathbb{E}\left[\left(\widehat{\mathbb{E}\left[p_t^k\right]} - \mathbb{E}\left[\widehat{\mathbb{E}\left[p_t^k\right]}\right]\right)^2\right] = \mathbb{E}\left[\widehat{\mathbb{E}\left[p_t^k\right]}^2\right] - \underbrace{\mathbb{E}\left[\widehat{\mathbb{E}\left[p_t^k\right]}\right]^2}_{=\mathbb{E}\left[p_t^k\right]^2} \\
&= \frac{1}{N^2}\mathbb{E}\left[\left(\sum_{i=1}^{N}\left(p_t(i)\right)^k\right)^2\right] - \mathbb{E}\left[p_t^k\right]^2 \\
&= \frac{1}{N^2}\sum_{i=1}^{N}\mathbb{E}\left[\left(\left(p_t(i)\right)^k\right)^2\right] + \frac{N(N-1)}{N^2}\mathbb{E}\left[\left(p_t(i)\right)^k\right]^2 - \mathbb{E}\left[p_t^k\right]^2 \\
&= \frac{1}{N}\mathbb{E}\left[p_t^{2k}\right] + \left(1 - \frac{1}{N}\right)\mathbb{E}\left[p_t^k\right] - \mathbb{E}\left[p_t^k\right]^2 \\
&= \frac{1}{N}\left(\mathbb{E}\left[p_t^{2k}\right] - \mathbb{E}\left[p_t^k\right]^2\right).
\end{aligned}$$

The point of this exercise is to show that the variance of the estimator of the kth moment depends on the $(2k)$th moment, e.g., the estimator of the second moment depends on the fourth moment. Clearly, if the fourth moment diverges, so does the variance of the estimator of the second moment. But if the $(2k)$th moment does exist, the variance of the estimator decreases like $1/N$ as we are used to.

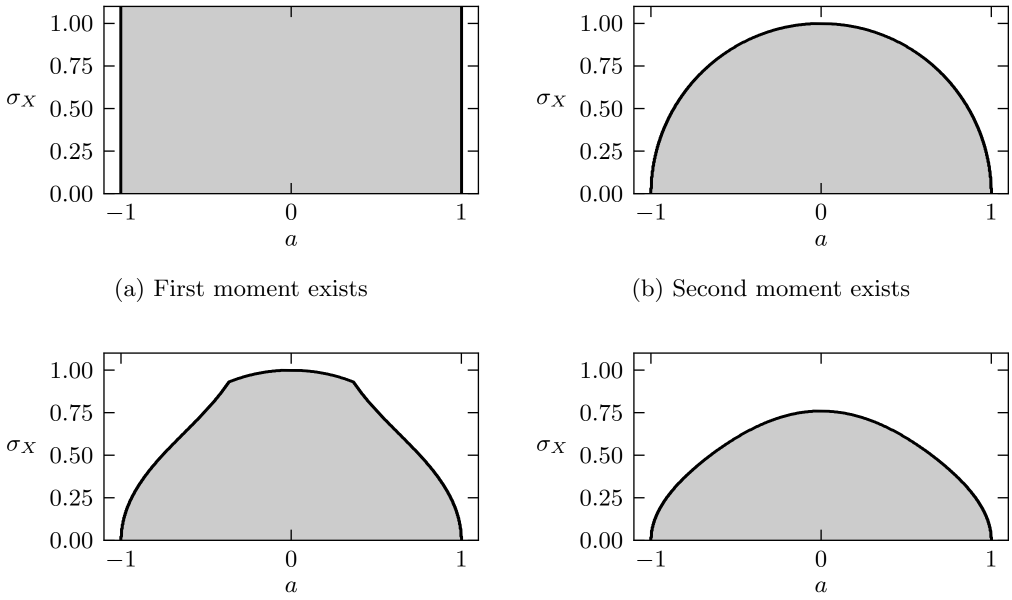
3.5 Results and conclusions

The theoretical results derived in section 3.3 improves our understanding of the model significantly, in two distinct ways described below.

3.5.1 Parameter restrictions for stationarity

First, the calculation of the various quantities leads to a number of mathematical restrictions on the values of the parameters a and σ_X. For example, requiring that the autocorrelation of squared returns or the kurtosis of the return distribution exists, also implies that all the moments up to $\mathbb{E}\left[p_t^4\right]$ must exist. This, in turn, restricts a and σ_X as described in section 3.3.3. The regions where the first four moments of the return distribution exist are illustrated in figure 3.3.

In the final chapter, we discuss a few issues which pertain more generally to models of financial markets. One of them, mentioned also by Cont (2001), is that we perhaps should not require the theoretical existence of higher moments such as skewness and kurtosis. At this point, it can be noted that the simulation results by



(a) First moment exists

(b) Second moment exists

(c) Third moment exists

(d) Fourth moment exists

Figure 3.3: The regions in the (a, σ_X)-plane where the first four moments of the return distribution exist. Note that changing sign on a makes no difference for the existence of moments, since equation (3.5) is symmetric in a. We require $\sigma_X \geq 0$ because X_t is symmetrically distributed and changing signs of σ_X makes no difference whatsoever for the model.

Westerhoff and Franke (2012) were made in a parameter region where the theoretical second and third moments exist, but the fourth moment diverges. This is illustrated in figure 3.4.

3.5.2 Parameter restrictions for agreement with stylized facts

Second, even more restrictions on the parameter values can be found if we require that the quantities we have found expressions for agree with the stylized facts to some extent.

Autocorrelation in returns

Perhaps the most basic stylized fact is that there should be no significant autocorrelations in returns. Set any limit $|C_{1,\max}|$ on the maximal absolute value of the first-lag autocorrelation. Since the first autocorrelation of returns is $(1-a)/2$, it is required that $|1-a| < 2\,|C_{1,\max}|$, and since $|a| < 1$ if the first moment exists,

$$a > 1 - 2\,|C_{1,\max}|\,.$$

Furthermore, the autocorrelation of returns only exists if $a^2 + \sigma_X^2 < 1$, so this creates a closed region in the (a, σ_X)-plane where all autocorrelations are smaller than the limit $|C_{1,\max}|$. This is illustrated in figure 3.5 for the autocorrelation limits 10^{-2} and 10^{-3}.

Fat tails of the return distribution

The kurtosis of the return distribution is always $\kappa \geq 3$ (the expression is found in appendix A.5), and $\kappa = 3$ only when $\sigma_X = 0$. In other words, the return distribution

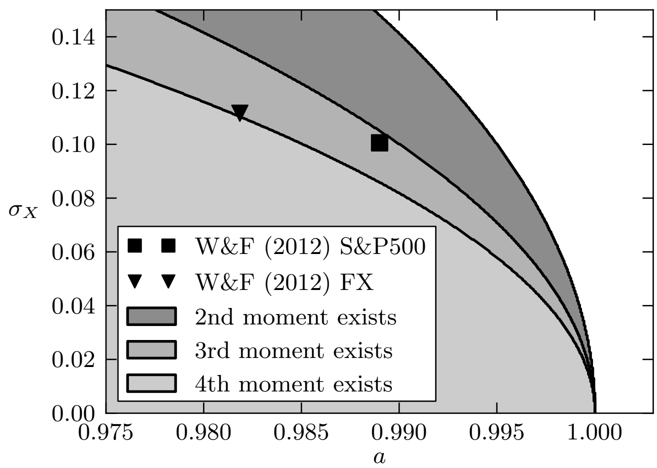


Figure 3.4: Convergence regions of the second, third and fourth price moments are marked in shaded areas. The two scenarios simulated by Westerhoff and Franke (2012) are marked in the region where the three first moments exist, but not the fourth. The two scenarios simulated by Westerhoff and Franke both have the last parameter $b = 0$.

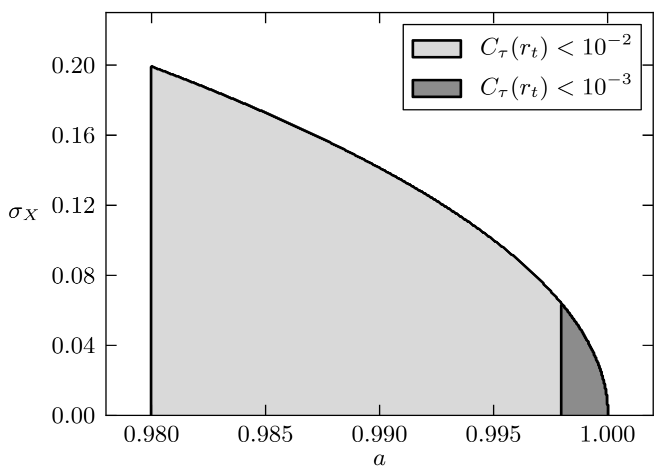
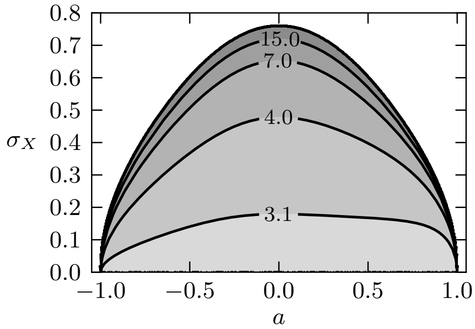
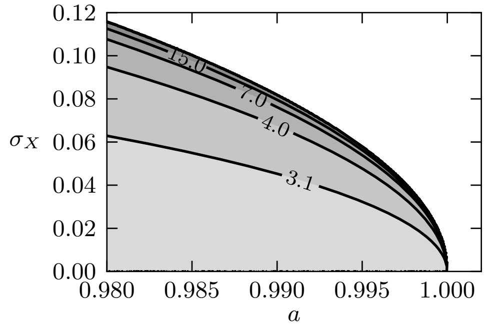


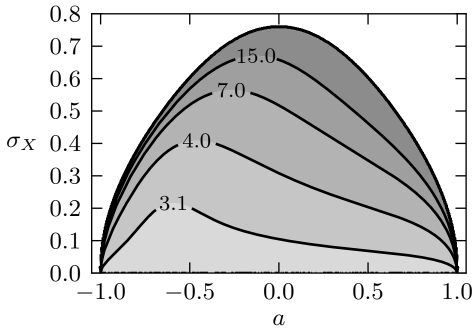
Figure 3.5: The regions in the (a, σ_X)-plane where the autocorrelation of returns exists and is smaller than 10^{-2} or 10^{-3} in absolute value for all lags $\tau \geq 1$.



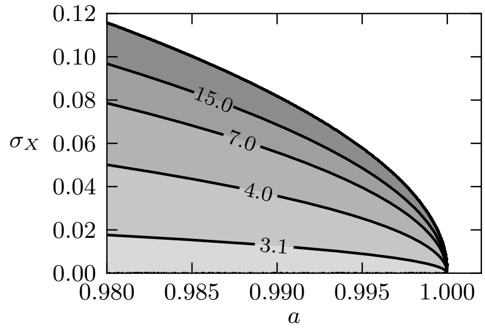
(a) $b = 0$



(b) $a > 0.98$, $b = 0$



(c) limit $b \to \infty$



(d) $a > 0.98$, limit $b \to \infty$

Figure 3.6: Contour plots showing kurtosis of the return distribution in the (a, σ_X)-plane with $b = 0$ (upper row) and in the limit $b \to \infty$ (lower row). The return distribution is leptokurtic ($\kappa > 3$) in the whole shaded area. Darker colors indicate higher kurtosis. The fourth moment of the return distribution does not converge in the white region (so the kurtosis does not exist). The kurtosis is an even function with respect to b, and increases monotonically with the absolute value of b, so the two cases $b = 0$ and $b \to \infty$ minimize and maximize the kurtosis, respectively, for any combination of a and σ_X.

strictly speaking is leptokurtic for all $\sigma_X > 0$. However, to find kurtosis values of roughly $5 < \kappa < 100$ as reported by Cont (2001) and Lux (2009), the model parameters are much more restricted, as illustrated in figure 3.6.

Figure 3.7 shows how the shape of the estimated return distribution varies with b and a fixed pair (a, σ_X) corresponding to the stock market scenario in Westerhoff and Franke's paper.

Skewness

The normalized skewness of the return distribution,

$$\frac{\mathbb{E}\left[r_t^3\right]}{\mathbb{E}\left[r_t^2\right]^{3/2}} = \frac{3(2a+1)b\sigma_X^2}{\sqrt{2}\left(a^3 + 3a\sigma_X^2 - 1\right)\sqrt{\frac{(a-1)^2 + b^2\sigma_X^2}{(a-1)\left(a^2+\sigma_X^2-1\right)}}},$$

is obviously odd with respect to b, and it is easily verified that it increases monotonically with b. The stylized facts about returns skewness are not quantitatively certain, but based on the values reported by Cont (2001), we choose at least to require that skewness is non-positive, if it exists, which implies $b \geq 0$.

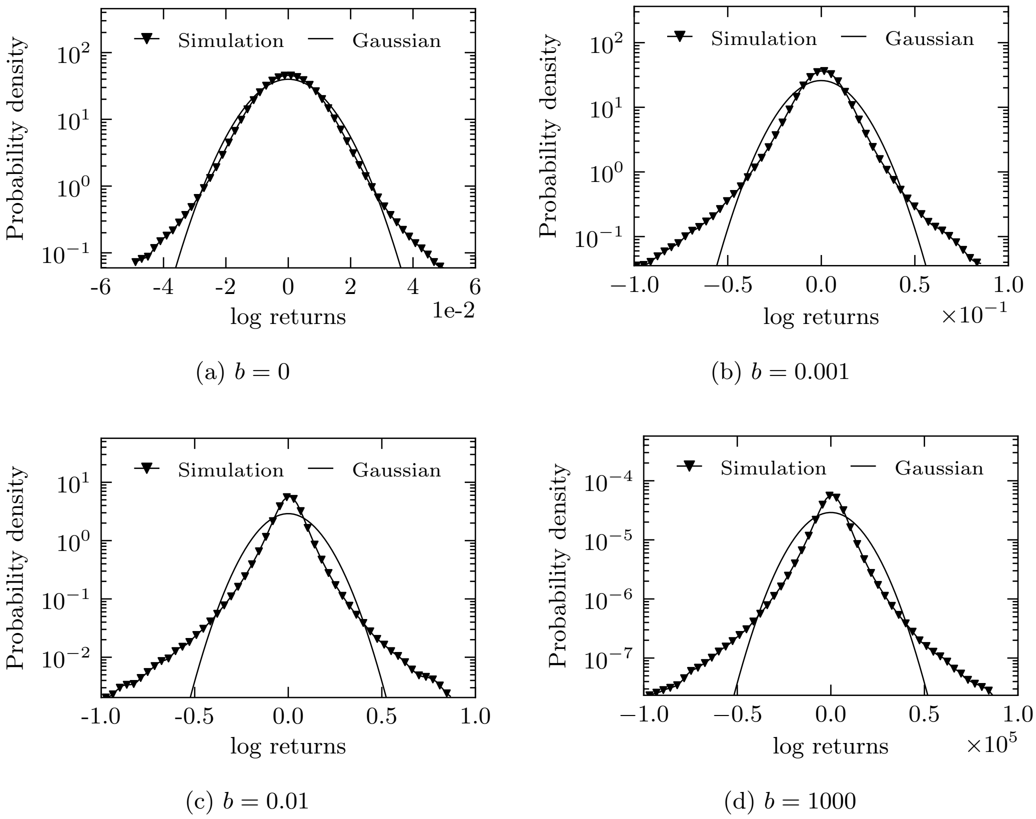


Figure 3.7: Variation of the shape of the simulated return distribution with different values of b, and a and σ_X fixed at the parameters adapted from Westerhoff and Franke's stock market scenario. The results show that a change from $b = 0$ to $b \geq 10^{-3}$ significantly changes the shape of the return distribution, introducing a slight visible skewness, but also a clear narrowing of the center of the distribution. The return distributions are Epanechikov kernel estimates using a bandwidth choice suggested by Davidson and MacKinnon (2004, p. 681, equation (15.63)), based on simulations of 10^6 time steps and discarding the first $5 \cdot 10^5$ to eliminate any transient effects. The parameters of Westerhoff and Franke's stock market scenario transformed to our simplified model are $a = 0.989, \sigma_X = 0.101, b = 0$.

The normalized skewness converges as $b \to \infty$,

$$\lim_{b\to\infty} \frac{\mathbb{E}\left[r_t^3\right]}{\mathbb{E}\left[r_t^2\right]^{3/2}} = -\frac{3(2a+1)\sigma_X^2}{\sqrt{2}\sqrt{\frac{\sigma_X^2}{(a-1)\left(a^2+\sigma_X^2-1\right)}}\left(1-a^3+3a\sigma_X^2\right)},$$

so for each point (a, σ_X), we can be sure that the normalized skewness is never more negative than this quantity. This minimal skewness function in illustrated in figure 3.8.

With the requirement that autocorrelations in returns are small in absolute value, say $< 10^{-2}$, the relevant parameter region is $a > 0.98$. An illustration of the skewness values obtained in this region is found in figure 3.9.

Figure 3.7 shows the return distribution in simulations of the stock market scenario in Westerhoff and Franke's paper, and variations with increasing b.

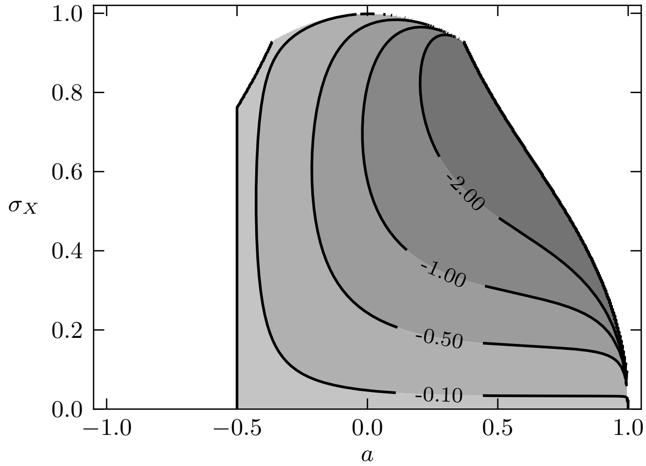
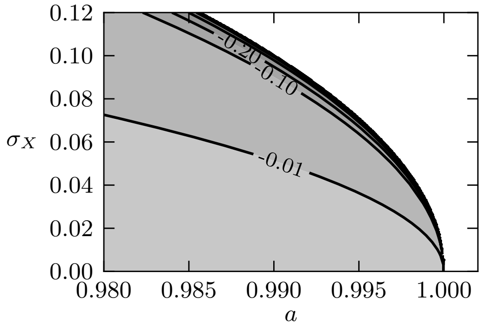
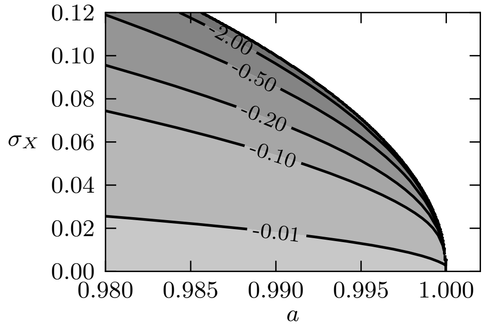


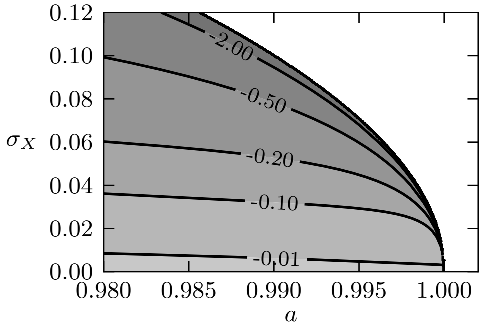
Figure 3.8: Contour plot in the (a, σ_X)-plane of the minimal (most negative) skewness of the return distribution, obtained in the limit $b \to \infty$. Darker colors indicate more negative skewness. In the white regions, skewness is positive or does not exist.



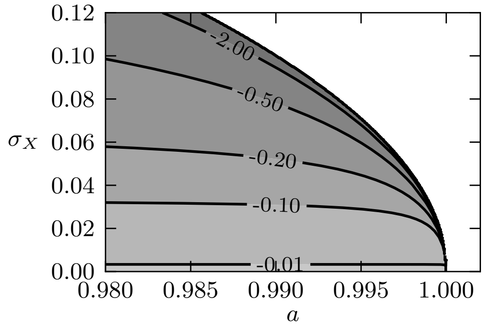
(a) $b = 0.01$



(b) $b = 0.1$



(c) $b = 1$



(d) limit $b \to \infty$

Figure 3.9: Contour plots showing skewness of the return distribution for different values of b. The skewness is negative for all $b > 0$ in this part of the (a, σ_X)-plane. Darker colors indicate more negative skewness. The third moment does not exist in the white region. Not much change can be seen between the cases $b = 1$ and $b \to \infty$, which means that skewness almost reaches its minimal value already at $b = 1$ in this region of the (a, σ_X)-plane.

Autocorrelation in squared returns

The central stylized fact about autocorrelation in squared returns is that it decays slowly with increasing lags, perhaps like a power law. This stylized fact is never matched by the model, since the autocorrelation in squared returns always falls approximately exponentially. This is a fundamental shortcoming of the model in reproducing the stylized facts.

Since the shape of the autocorrelation function of squared returns does not match empirical data, it is perhaps not a worthwhile exercise to compare the function values either. With this caveat, we present some rough results anyway.

To keep skewness non-positive, we require $b \geq 0$. Differentiating the expression for the first-lag autocorrelation of squared returns $C_1(r_t^2)$ (see appendix A.5), it can be verified that $C_1(r_t^2)$ increases monotonically with b if $\sigma_X > 0$. As b approaches infinity, the first autocorrelation converges to an expression that only depends on a and σ_X. In conclusion, we know that if $b \geq 0$, the minimal autocorrelation in squared returns is obtained for $b = 0$, and the maximal autocorrelation is obtained in the limit $b \to \infty$.

The first-lag autocorrelation $C_1(r_t^2)$ is illustrated in figure 3.10 for the two cases $b = 0$ and $b \to \infty$ in the whole region of the (a, σ_X)-plane where the fourth moment $\mathbb{E}\left[r_t^4\right]$ exists.

In the region where autocorrelation of returns is small, i.e., where $1 - a$ is small, the first-lag autocorrelation is largest along the border where the fourth moment $\mathbb{E}\left[r_t^4\right]$ diverges. This is illustrated in figure 3.11 for some different values of b. The decay of the autocorrelation $C_\tau(r_t^2)$ with the lag τ is also slower close to this border (see equation (3.12)), so choosing a point (a, σ_X) in a darker region in figure 3.11 both gives relatively large autocorrelation in squared returns and a relatively slow decay.

Hill tail estimator

The simulation results of the Hill tail estimator obtained by Westerhoff and Franke (2012) were on average $\hat{\alpha} \approx 3.5$ in their stock market scenario. Our results show that this is a very typical value of the Hill estimator in the parameter region where the other stylized facts are well reproduced. Contour plots of simulation results in the relevant part of the parameter space is shown in figure 3.12.

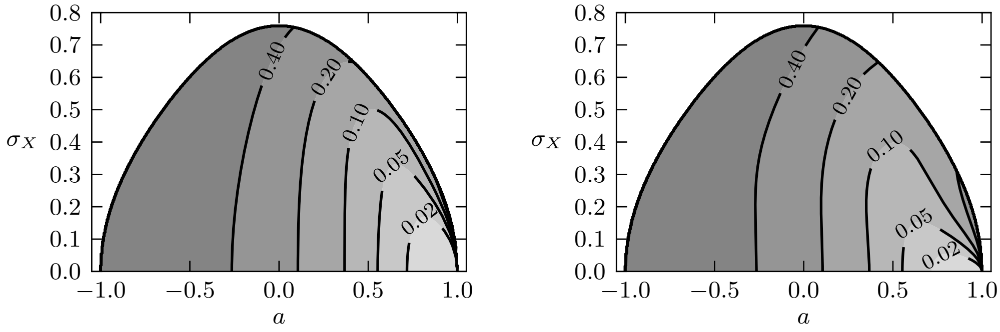
It is reassuring to note that the Hill estimator values do not vary much if the utilized fraction f_H of the distribution changed. Westerhoff and Franke calculated the tail index based on the top 5% of the returns, i.e., $f_\mathrm{H} = 0.05$, and so did we for the results shown in figure 3.12. But additional simulations (not illustrated in the report) show that very similar results are obtained for $0.001 < f_\mathrm{H} < 0.05$. This indicates that the tail really has characteristics similar to a power law.

3.5.3 Conclusions

The most basic requirement on the model is perhaps that the autocorrelation of returns is small in absolute value, which restricts the parameters a and σ_X significantly, as shown in figure 3.5.

Requiring that the skewness of the return distribution is zero implies $b = 0$. Allowing also negative skewness values implies $b \geq 0$.

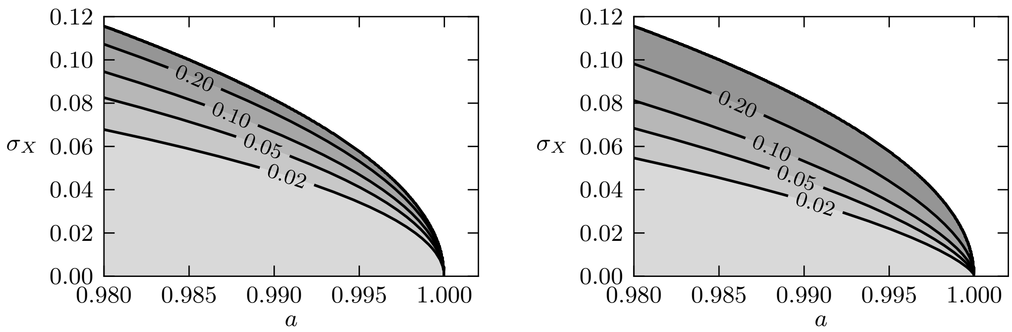
Requiring that the return distribution has a finite fourth moment, but is clearly leptokurtic (say $5 \lesssim \kappa$) restricts a and σ_X even further to a relatively narrow band just below the curve $1 = a^4 + 6a^2\sigma_X^2 + 3\sigma_X^4$, where the fourth moment diverges.



(a) $b = 0$

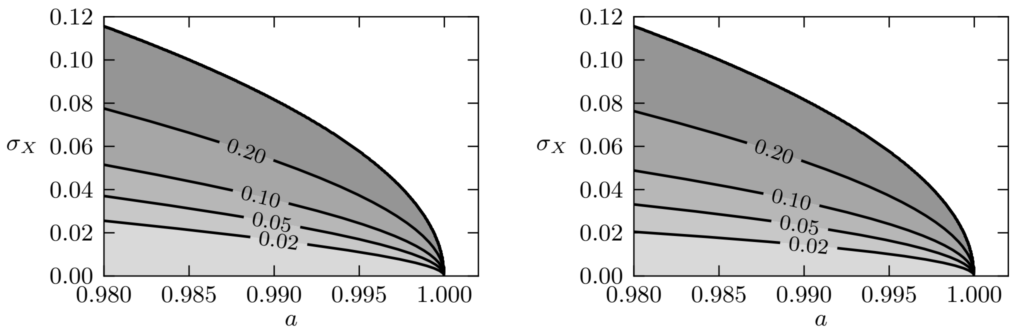
(b) limit $b \to \infty$

Figure 3.10: Contour plot in the (a, σ_X)-plane of first-lag autocorrelation of squared returns, $C_1(r_t^2)$. The autocorrelation is always positive. Darker colors indicate larger values. In the white regions, the fourth moment of the return distribution does not exist.



(a) $b = 0$

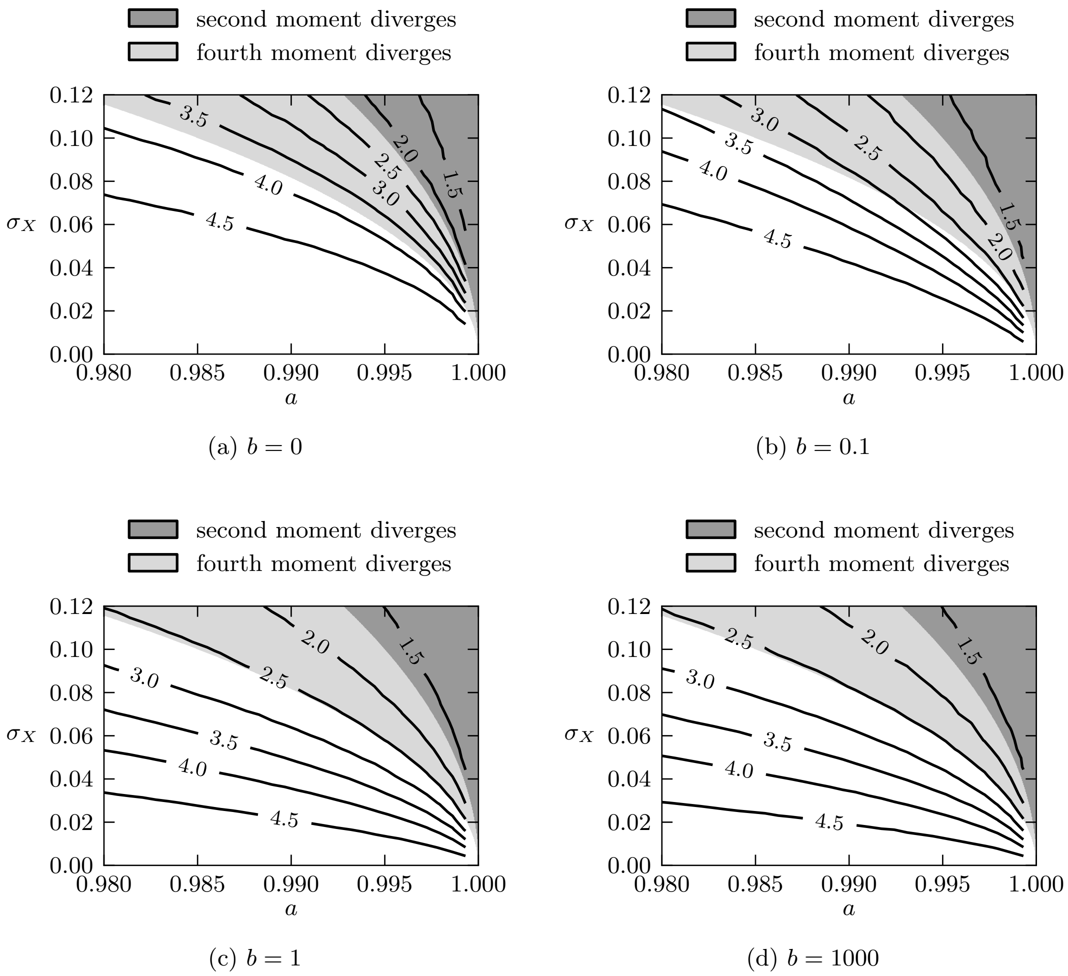
(b) $b = 0.1$



(c) $b = 1$

(d) limit $b \to \infty$

Figure 3.11: Contour plots showing first-lag autocorrelation of squared returns, $C_1(r_t^2)$ for different values of b. The autocorrelation is always positive for $b > 0$. Darker colors indicate larger values. The fourth moment does not exist in the white region. Not much change can be seen between the cases $b = 1$ and $b \to \infty$, which means that autocorrelation almost reaches its maximal value already at $b = 1$ in this region of the (a, σ_X)-plane.



(a) $b = 0$

(b) $b = 0.1$

(c) $b = 1$

(d) $b = 1000$

Figure 3.12: Contour plots of the Hill tail estimator for different values of b. The shaded regions show where the theoretical second and fourth moments of the return distribution diverge. The return distribution was estimated by taking 10^6 time steps, starting with the price $p_0 = b/(1-a)$ and discarding the first 10^3 values. The Hill estimator was computed based on the upper 5% of the simulated values, and finally averaged over 30 independent runs.

This band is less narrow for large b. Kurtosis results are illustrated in figure 3.6.

Figure 3.7 shows simulated return distributions for different values of b and with "realistic" choices of a and σ (corresponding to the stock market scenario in Westerhoff and Franke's paper). The simulation results show that a change from $b = 0$ to $b \geq 10^{-3}$ significantly changes the shape of the return distribution, introducing a slight visible skewness, but also a clear narrowing of the center of the distribution.

The contour plots for returns kurtosis (figure 3.6) and the first-lag autocorrelation of squared returns (figure 3.11) are quite similar in the relevant region. The two figures show that strong leptokurtosis is obtained in the same region as large and relatively slowly decaying autocorrelations in squared returns.

The long-range autocorrelations in squared returns is never reproduced by this model, in the region where the fourth moment is finite. The autocorrelations decay exponentially as a function of the lag τ, which is far from matching the stylized fact of a power law decay.

Numerical simulation results indicate that the Hill tail estimator of the return distribution always takes values roughly in the range $2.5 \lesssim \hat{\alpha} \lesssim 4$, in the region of (a, b, σ_X)-space where returns autocorrelation, kurtosis and skewness reasonably match the stylized facts.

In summary, to the extent that the model can reproduce the stylized facts about autocorrelations of returns and squared returns, returns skewness and kurtosis, and the values of the Hill tail estimator, they all do so simultaneously in a relatively narrow region of the parameter space. We can only speculate at this point whether these results are a coincidence, or if they actually indicate that model provides an explanation of some regularities seen in financial markets.

The choice of parameter values a, b and σ_X seems impossible to make from first principles, since the model is a rather abstract description of reality. Our work only asserts that the model produces interesting results in a region of the parameter space, but provides no clue as to why the parameters should be chosen in that region. We note that the simulation results presented by Westerhoff and Franke (2012) were produced in the region where the three first price moments exist, but the fourth moment diverges (see figure 3.4 on page 25).

Chapter 4

Discussion and outlook

In this final chapter, we touch upon a few discussion points that are relevant for the model presented in the previous chapter. We also believe that these last points are relevant points to consider for all the heterogeneous agent models mentioned in section 2.5.3 (page 12).

4.1 Existence of moments of the returns process

The foundation of our model analysis has been to derive the theoretical moments of the returns process. Insisting that a sufficient number of moments exist, i.e., that they do not diverge as $t \to \infty$, we derived expressions for the autocorrelation of the returns process, the skewness and kurtosis of the return distribution, and the autocorrelation function of squared returns.

However, it is not clear at all that we should impose such restrictions on the parameter choices. As we discussed in section 2.4.2 (page 10), it seems like empirical returns series have a tail index around 3, which means at least that the fourth moment does not exist. If the tail index is below 3, the skewness does not exist, either. But the possible nonexistence of higher moments does not really stop us from estimating them, as we have done with the autocorrelation of squared returns in simulation results from a region where the fourth moment does not theoretically exist (see figure 3.1d on page 14). But the point should be taken seriously: if the fourth moment of the return distribution doesn't exist, there is no such thing as a autocorrelation in squared returns, and those who decide to use the estimator anyway should keep this in mind.

4.2 What do models explain?

The model presented in chapter 3 has a couple of strong advantages speaking for it. First, it successfully reproduces some of the stylized facts in a certain region of the parameter space, which was a primary purpose. Second, the model is simple enough to allow some analytical treatment, which significantly improves our understanding of it.

The main disadvantage of the model is its limited explanatory power. Westerhoff and Franke did provide a motivation for the different terms in the model equation, but the motivation is only conceptual, and it seems difficult to determine any of the parameter values in the original formulation based on empirical data.

One difficulty in mathematical modeling is to choose a mathematical representation of the things that are known (or thought to be known) about the real world, because equations and explanations do not have a one-to-one correspondence. In other words, a single explanation could be modeled mathematically in different ways, and a single mathematical formulation could match any number of different explanations.

This brings us to a general point about choosing parameter values in a model. If the parameter values of a model can be estimated with reasonable accuracy, without considering their impact on the model, and if the model with those parameter values successfully reproduces some empirical facts, that is a sort of evidence in favor of the model structure, indicating that the model may actually have captured the actual mechanisms at work.

But if the parameter values going into a model are unknown, or too abstract to estimate, they usually have to be fitted to match model results against empirical results. In such cases, the validity of the model structure is much harder to assess. There is always a risk that the model through parameter choices can be made to reproduce a wide range of results. With such models it is preferable to make a very wide search if possible, to determine which region(s) of the parameter space produce interesting results. Unfortunately this is usually increasingly hard as the number of parameters grows, but it is also in models with many parameters that this point is most important.

4.3 Time scales and the source of randomness

We finally turn to an interesting point made by Westerhoff and Franke (2012). Their paper is entitled *Converse trading strategies, intrinsic noise and the stylized facts of financial markets*. The expression "intrinsic noise" refers to their motivation for the binomially distributed random variables they introduced, namely the random choice of N independent traders to participate in the market with probability π in each time step. In other words, the random fluctuations in the model are endogenously motivated.

This leads to a point of discussion we have largely swept under the carpet so far, namely the time step. The stock market scenario in Westerhoff and Franke's paper is calibrated against daily return series from the S&P500 stock index, so the time step is implicitly said to be $\Delta t = 1$ day. The problem is that if we increase the length of the time step, this source of randomness should gradually die out, because the distribution of the number of active traders becomes more narrow.

In terms of the model, increasing the time step from Δt to $m\Delta t$ is equivalent to increasing the number of independent trade decisions from N to mN, since the N traders have m independent chances of trading in a time period of $m\Delta t$. Letting $n(t, t + m\Delta t)$ denote the number of trades in the period t to $t + m\Delta t$, we note that the expected number of trades increases to $\mathbb{E}[n(t, t + m\Delta t)] = mN\pi$, while the standard deviation $\text{std}(n(t, t + m\Delta t)) = \sqrt{mN\pi(1 - \pi)}$. In other words, the expected number of trades (rather intuitively) increases linearly with the time step, while the standard deviation only increases as the square root.

Therefore, Westerhoff and Franke's "intrinsic noise" explanation only makes sense at short enough time scales, where the number of trade decisions is small enough that the outcome can fluctuate significantly. This does not invalidate the whole concept, but it is an important point to take into explicit consideration if the goal is to improve the explanatory value of the model.

Perhaps this last point is an argument in favor of the order book models introduced in section 2.5.2. Such models have a naturally defined time scale, and can certainly make use of the concept of "intrinsic noise".

4.4 Conclusion and outlook

The model we present and analyze in the previous chapter is relatively successful in reproducing some of the stylized facts of financial returns series, and several other models have been demonstrated to have similar performance in this respect. However, this model and several similar models are motivated in a relatively abstract way, which makes it hard to argue that any single model should be preferred above other models.

In one sense, it can be seen as positive that several simple models are relatively effective in reproducing the stylized facts. It means that some of these statistical regularities in financial markets are not so strange as they may seem. On the other hand, it also means that additional simple but abstract explanations have little value, because they do not necessarily increase our understanding of the financial markets in the real world. A model is not qualified for making concrete predictions or supporting policy advice, just because it reproduces some of the stylized facts.

A significant development one could hope for, that we have not encountered in the literature, is a model with parameters that could be estimated in other ways than by fitting model results to empirical data, and still reproduce some of the stylized facts. That could be a first step towards an improved understanding of the actual drivers of market behavior.

References

Alfi, V. et al. (2009). Minimal agent based model for financial markets I. *The European Physical Journal B* **67**.3, 385–397.

Bonato, Matteo (2011). Robust estimation of skewness and kurtosis in distributions with infinite higher moments. *Finance Research Letters* **8**.2, 77–87.

Bouchaud, JP (2008). Economics needs a scientific revolution. *Nature* **455**.October, 2008.

Brock, William A. and Cars H. Hommes (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control* **22**, 1235–1274.

Chakraborti, Anirban et al. (2011a). Econophysics review: I. Empirical facts. *Quantitative Finance* **11**.7, 991–1012.

– (2011b). Econophysics review: II. Agent-based models. *Quantitative Finance* **11**.7, 1013–1041.

Challet, Damien and Robin Stinchcombe (2001). Analyzing and modeling 1+1d markets. *Physica A: Statistical Mechanics and its Applications* **300**.1-2, 285–299.

Chiarella, Carl, Xue-Zhong He, and Cars H. Hommes (2006). A dynamic analysis of moving average rules. *Journal of Economic Dynamics and Control* **30**.9-10, 1729–1753.

Cont, Rama (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance* **1**.2, 223–236.

Davidson, Russel and James G. MacKinnon (2004). *Econometric Theory and Methods*. New York: Oxford University Press.

Day, R. H. and W Huang (1990). Bulls, bears and market sheep. *Journal of Economic Behavior & Organization* **14**, 299–329.

Ding, Z and CWJ Granger (1996). Modeling volatility persistence of speculative returns: a new approach. *Journal of econometrics* **73**, 185–215.

Engle, RF (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society* **50**.4, 987–1007.

Farmer, John Doyne and John Geanakoplos (2009). The virtues and vices of equilibrium and the future of financial economics. *Complexity* **14**.3, 11–38.

Farmer, John Doyne and Shareen Joshi (2002). The price dynamics of common trading strategies. *Journal of Economic Behavior & Organization* **49**.2, 149–171.

Grauwe, P De and Hans Dewachter (1993). A chaotic model of the exchange rate: the role of fundamentalists and chartists. *Open economies review* **4**.4, 351–379.

Guillaume, Dominique M. et al. (1997). From the bird's eye to the microscope: A survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics* **129**.1, 95–129.

Hill, BM (1975). A simple general approach to inference about the tail of a distribution. *The annals of statistics* **3**.5, 1163–1174.

Hommes, Cars H. (2006). “Heterogeneous Agent Models in Economics and Finance”. *Handbook of computational economics. Volume 2: Agent-based computational economics*. Ed. by Leigh Tesfatsion and Kenneth L. Judd. Amsterdam: North-Holland. Chap. Chapter 23.

Hunter, J. Stuart, Willam Gordon Hunter, and George E.P. Box (2005). *Statistics for experimenters*. Wiley-Interscience.

Kaldor, N (1961). “Capital accumulation and economic growth”. *The Theory of Capital*. Ed. by F. A. Lutz and D. C. Hague. London: McMillan, pp. 177–222.

Kim, Tae-Hwan and Halbert White (2004). On more robust estimation of skewness and kurtosis. *Finance Research Letters* **1**.1, 56–73.

Kirkpatrick, Charles D. and Julie R. Dahlquist (2010). *Technical Analysis: The Complete Resource for Financial Market Technicians*. 2nd ed. New Jersey: FT Press.

Knight, Tim (2007). *Chart Your Way to Profits: The Online Trader’s Guide to Technical Analysis*. 2nd ed. New Jersey: John Wiley & Sons.

Lux, Thomas (2009). “Stochastic Behavioral Asset-Pricing Models and the Stylized Facts”. *Handbook of Financial Markets: Dynamics and Evolution*. Ed. by Thorsten Hens and Klaus Reiner Schenk-Hoppé. Elsevier. Chap. 3, pp. 161–215.

Malevergne, Y., V. Pisarenko, and D. Sornette (2005). Empirical distributions of stock returns: between the stretched exponential and the power law? *Quantitative Finance* **5**.4, 379–401.

Mandelbrot, B (1963). The variation of certain speculative prices. *Journal of business* **36**.4, 394–419.

Mike, Szabolcs and John Doyne Farmer (2008). An empirical behavioral model of liquidity and volatility. *Journal of Economic Dynamics and Control* **32**.1, 200–234.

Najim, Kaddour, Enso Ikonen, and Ait-Kadi Daoud (2004). *Stochastic processes*. Oxford: Kogan Page Science.

Råde, Lennart and Bertil Westergren (2004). *Mathematics Handbook for Science and Engineering*. 5th ed. Lund: Studentlitteratur.

Rice, John A. (2007). *Mathematical Statistics and Data Analysis*. 3rd ed. Belmont, CA: Thomson Brooks/Cole.

Tramontana, Fabio, Frank Westerhoff, and Laura Gardini (2010). On the complicated price dynamics of a simple one-dimensional discontinuous financial market model with heterogeneous interacting traders. *Journal of Economic Behavior & Organization* **74**.3, 187–205.

Vervaat, W (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Advances in Applied Probability* **11**.4, 750–783.

Westerhoff, Frank and Reiner Franke (2012). Converse trading strategies, intrinsic noise and the stylized facts of financial markets. *Quantitative Finance* **12**.3, 425–436.

Appendix A

Supporting calculations

A.1 Reformulation of price product expressions

Any expression of the form

$$\mathbb{E}\left[\prod_{i=0}^{m} p_{t-i}^{\nu_i}\right], \tag{A.1}$$

with a non-negative integer m and non-negative integer exponents $\{\nu_i\}$, can be expressed as

$$\sum_{k=0}^{n} \gamma_k \mathbb{E}\left[p_{t-m}^{k}\right],$$

where $\{\gamma_k\}$ are constants and the largest exponent is

$$n = \sum_{i=0}^{m} \nu_i.$$

To see why, use the model equation to rewrite (A.1) as

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=0}^{m} p_{t-i}^{\nu_i}\right] &= \mathbb{E}\left[p_t^{\nu_0} \prod_{i=1}^{m} p_{t-i}^{\nu_i}\right] \\
&= \mathbb{E}\left[(p_{t-1}(a + \sigma_X X_{t-1}) + Y_{t-1} + b)^{\nu_0} \prod_{i=1}^{m} p_{t-i}^{\nu_i}\right] \\
&= \mathbb{E}\left[\sum_{\lambda_0=0}^{\nu_0} \binom{\nu_0}{\lambda_0} (p_{t-1}\,(a + \sigma_X X_{t-1}))^{\lambda_0}\, (Y_{t-1} + b)^{\nu_0-\lambda_0} \prod_{i=1}^{m} p_{t-i}^{\nu_i}\right] \\
&= \mathbb{E}\left[\sum_{\lambda_0=0}^{\nu_0} \binom{\nu_0}{\lambda_0} (a + \sigma_X X_{t-1})^{\lambda_0}\, (Y_{t-1} + b)^{\nu_0-\lambda_0} \prod_{i=1}^{m} p_{t-i}^{\nu_i+\delta_{i1}\lambda_0}\right],
\end{aligned}$$

where δ_{ij} is the Kronecker delta function ($\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise).

Observing that X_{t-1} and Y_{t-1} are independent of each other and independent of

p_{t-i} for $i \geq 1$, and using linearity of expectation, we find

$$\begin{aligned}
&\mathbb{E}\left[\prod_{i=0}^{m} p_{t-i}^{\nu_i}\right] \\
&= \sum_{\lambda_0=0}^{\nu_0} \binom{\nu_0}{\lambda_0} \mathbb{E}\left[(a + \sigma_X X_{t-1})^{\lambda_0}\right] \mathbb{E}\left[(Y_{t-1} + b)^{\nu_0 - \lambda_0}\right] \mathbb{E}\left[\prod_{i=1}^{m} p_{t-i}^{\nu_i + \delta_{i1}\lambda_0}\right] \\
&= \sum_{\lambda_0=0}^{\nu_0} \alpha_0(\lambda_0) \mathbb{E}\left[\prod_{i=1}^{m} p_{t-i}^{\nu_i + \delta_{i1}\lambda_0}\right], \qquad \text{(A.2)}
\end{aligned}$$

with constants $\alpha_0(\lambda_0)$ that can always be found since the expectation $\mathbb{E}\left[X_t^k\right] = \mathbb{E}\left[Y_t^k\right]$ is known (and finite) for all positive integers k.

Expression (A.2) is a linear combination of expressions of the form (A.1), but only containing products of $\{p_{t-1}, p_{t-2}, \ldots, p_{t-m}\}$ and not p_t. Each term in the linear combination (A.2) can then be rewritten in the same way, using products of $\{p_{t-2}, p_{t-3}, \ldots, p_{t-m}\}$,

$$\begin{aligned}
&\mathbb{E}\left[\prod_{i=0}^{m} p_{t-i}^{\nu_i}\right] \\
&= \sum_{\lambda_0=0}^{\nu_0} \alpha_0(\lambda_0) \mathbb{E}\left[\prod_{i=1}^{m} p_{t-i}^{\nu_i + \delta_{i1}\lambda_0}\right] \\
&= \sum_{\lambda_0=0}^{\nu_0} \sum_{\lambda_1=0}^{\nu_1+\lambda_0} \alpha_0(\lambda_0)\alpha_1(\lambda_1) \mathbb{E}\left[\prod_{i=2}^{m} p_{t-i}^{\nu_i + \delta_{i2}\lambda_1}\right],
\end{aligned}$$

and so on. After rewriting m times, the reduction is complete,

$$\begin{aligned}
&\mathbb{E}\left[\prod_{i=0}^{m} p_{t-i}^{\nu_i}\right] \\
&= \sum_{\lambda_0=0}^{\nu_0} \sum_{\lambda_1=0}^{\nu_1+\lambda_0} \sum_{\lambda_2=0}^{\nu_2+\lambda_1} \cdots \sum_{\lambda_m=0}^{\nu_m+\lambda_{m-1}} \prod_{i=0}^{m} \alpha_i(\lambda_i) \mathbb{E}\left[p_{t-m}^{\nu_m + \lambda_{m-1}}\right] \\
&= \sum_{k=0}^{n} \gamma_k \mathbb{E}\left[p_{t-m}^{k}\right].
\end{aligned}$$

The highest exponent on p_{t-m} that can appear in this final expression is clearly

$$\begin{aligned}
n &= \nu_m + \max \lambda_{m-1} \\
&= \nu_m + \nu_{m-1} + \max \lambda_{m-2} \\
&= \cdots = \sum_{i=0}^{m} \nu_i.
\end{aligned}$$

A.2 Moments of normally distributed variables

A standard normally distributed variable $X \sim \mathcal{N}(0,1)$ has the probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

Hence, the nth moment is

$$\mathbb{E}[X^n] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^n e^{-x^2/2}\,\mathrm{d}x.$$

If n is an odd integer, the integrand is an odd function, integrated over an even interval, so

$$\mathbb{E}[X^n] = 0, \qquad \text{for odd } n \geq 1.$$

For positive even integers $n = 2m$, the integral is

$$\mathbb{E}\left[X^{2m}\right] = \sqrt{\frac{2}{\pi}} \int_0^{\infty} x^{2m} e^{-x^2/2}\,\mathrm{d}x,$$

or with the variable change $t = x^2/2$,

$$\mathbb{E}\left[X^{2m}\right] = \frac{1}{\sqrt{\pi}} \int_0^{\infty} 2^m t^{m-\frac{1}{2}} e^{-t}\,\mathrm{d}t.$$

Consulting a standard handbook on mathematics, e.g. (Råde and Westergren, 2004, p. 287), this integral can be simplified by noting that the Gamma function is

$$\Gamma(z) = \int_0^{\infty} t^{z-1} e^{-t}\,\mathrm{d}t, \mathrm{Re} z > 0.$$

Hence, for a positive even integer $n = 2m$,

$$\mathbb{E}\left[X^{2m}\right] = \frac{1}{\sqrt{\pi}} 2^m \Gamma\left(m + \frac{1}{2}\right).$$

Another fact about the Gamma function is, for positive integers m,

$$\Gamma\left(m + 1/2\right) = \frac{(2m-1)!!\sqrt{\pi}}{2^m},$$

so in conclusion it is finally established for $X \sim \mathcal{N}(0,1)$,

$$\begin{aligned} \mathbb{E}[X^n] &= 0 && \text{for odd } n \geq 1, \\ \mathbb{E}[X^n] &= (n-1)!! && \text{for even } n \geq 2. \end{aligned}$$

A.3 Derivation of moments of prices

It has been rigorously shown by Vervaat (1979) that the moments of solutions to a certain class of stochastic difference equations converge to unique constants under some restrictions on the parameters and for almost any initial conditions. Our price process p_t belongs to this class of stochastic processes, and this section contains is a less formal calculation specific to the moments $\mathbb{E}[p_t^n]$ for $n > 0$. The main result is the following.

The nth moment $\mathbb{E}[p_t^n]$ for any integer $n \geq 1$ converges to a finite constant as $t \to \infty$, if the moments at the starting point, $\mathbb{E}[p_0^m]$, are finite constants for all $m = 1, 2, \ldots, n$, and

$$\left| a^m + \sum_{k=1}^{\lfloor m/2 \rfloor} \binom{m}{2k} a^{m-2k} \sigma_X^{2k} (2k-1)!! \right| < 1 \text{ for } m = 1, 2, \ldots, n, \tag{A.3}$$

where $\lfloor \cdot \rfloor$ denotes the floor function.

To prove this, we use the model equation, linearity of expectation and that p_t, X_t and Y_t are independent of each other to find

$$\begin{aligned}
\mathbb{E}\left[p_{t+1}^n\right] &= \mathbb{E}[(p_t\,(a+\sigma_X X_t) + Y_t + b)^n] \\
&= \mathbb{E}\left[\sum_{k=0}^{n} \binom{n}{k} (p_t\,(a+\sigma_X X_t))^k \,(Y_t+b)^{n-k}\right] \\
&= \sum_{k=0}^{n} \binom{n}{k} \mathbb{E}\left[(a+\sigma_X X_t)^k\right] \mathbb{E}\left[(Y_t+b)^{n-k}\right] \mathbb{E}\left[p_t^k\right] \\
&= \sum_{k=0}^{n} c_{nk} \mathbb{E}\left[p_t^k\right], \qquad \text{(A.4)}
\end{aligned}$$

where all the constants c_{nk} can be calculated exactly since $\mathbb{E}\left[X_t^k\right] = \mathbb{E}\left[Y_t^k\right]$ is known for all non-negative integers k.

Assuming that all the lower moments exist, i.e. that $\mathbb{E}\left[p_t^k\right]$ is finite and independent of t for all $k = 0, 1, \dots, (n-1)$, the result (A.4) can be written as

$$\mathbb{E}\left[p_{t+1}^n\right] = C_n \mathbb{E}[p_t^n] + D_n, \qquad \text{(A.5)}$$

where $C_n = c_{nn} = \mathbb{E}[(a+\sigma_X X_t)^n]$ and $D_n = \sum_{k=0}^{n-1} c_{nk}\mathbb{E}\left[p_t^k\right]$. It is easily verified that $D_n < \infty$ if a, b and σ_X are finite. Applying Equation (A.5) iteratively, it is seen that

$$\mathbb{E}\left[p_{t+N}^n\right] = C_n^N \mathbb{E}[p_t^n] + D_n \sum_{\nu=0}^{N-1} C_n^\nu.$$

Clearly, when $|C_n| \geq 1$, $\mathbb{E}\left[p_{t+N}^n\right]$ diverges as $N \to \infty$, but

$$\lim_{N\to\infty} \mathbb{E}\left[p_{t+N}^n\right] = \lim_{N\to\infty} C_n^N \mathbb{E}[p_t^n] + D_n \frac{1-C_n^N}{1-C_n} = D_n \frac{1}{1-C_n} \;\text{ for }\; |C_n| < 1.$$

In conclusion, the expectation $\mathbb{E}[p_t^n]$ converges to a finite constant as $t \to \infty$ if $|C_n| < 1$ and all $\mathbb{E}\left[p_t^k\right]$ are finite constants for $k = 1, 2, \dots, (n-1)$.

Make the assumption that the starting point p_0 has exactly those moments, i.e. $\mathbb{E}[p_0^n] = D_n/(1-C_n)$ whenever $|C_k| < 1$ for $k = 1, 2, \dots, n$. Then, the moment $\mathbb{E}[p_t^n]$ is obviously constant for all integers $t \geq 1$.

Beginning with $n = 1$, the expectation $\mathbb{E}[p_t]$ can be taken a finite constant if $|C_1| < 1$. Then the expectation $\mathbb{E}\left[p_t^2\right]$ can be taken a finite constant if $|C_2| < 1$ and $|C_1| < 1$, and so on. Generally, $\mathbb{E}[p_t^n]$ converges to a finite constant if

$$|C_m| < 1 \text{ for } m = 1, 2, \dots, n. \qquad \text{(A.6)}$$

A.4 Derivation of autocorrelations of squared returns

The autocorrelation of squared returns, $C_\tau(r_t^2)$ is derived in two steps, just like the autocorrelation of raw returns, $C_\tau(r_t)$ is derived in Section 3.3.4.

First, the autocorrelation of squared returns at lag $\tau = 1$ is computed using the general method using the definition of autocorrelation and the general method outlined in Section 3.3.

Second, the autocorrelation at higher lags $\tau \geq 2$ is computed by deriving a general formula for the quotient

$$\phi(\tau) = \frac{C_{\tau+1}(r_t^2)}{C_\tau(r_t^2)} = \frac{\mathrm{Cov}\big(r_{t+\tau+1}^2, r_t^2\big)}{\mathrm{Cov}\big(r_{t+\tau}^2, r_t^2\big)}$$

for $\tau \geq 1$.

Expanding the expression for $\phi(\tau)$ in terms of covariances yields an expression containing moments $\mathbb{E}\big[p_t^k\big]$ for $k = 1, 2, 3, 4$, which are known since before. But the expression also contains factors of the forms

$$\mathbb{E}\big[p_{t+n} r_t^2\big] \text{ and } \mathbb{E}\big[p_{t+n}^2 r_t^2\big],$$

for general integers $n \geq 1$. Given a closed-form expression for these expectations, $\phi(\tau)$ can also be expressed in closed form and the general autocorrelation $C_\tau(r_t^2)$ is thus known.

To find an expression for $\mathbb{E}\big[p_{t+n} r_t^2\big]$, where $n \geq 1$ note that

$$\begin{aligned}\mathbb{E}\big[p_{t+n} r_t^2\big] &= \mathbb{E}\big[(p_{t+n-1}(a + \sigma_X X_{t+n-1}) + b + Y_{t+n-1})\, r_t^2\big] \\ &= a\mathbb{E}\big[p_{t+n-1} r_t^2\big] + b\mathbb{E}\big[r_t^2\big].\end{aligned}$$

Since $\mathbb{E}\big[r_t^2\big]$ is a constant, this is a difference equation which can be iterated n times to find

$$\mathbb{E}\big[p_{t+n} r_t^2\big] = a^n \mathbb{E}\big[p_t r_t^2\big] + b\frac{1-a^n}{1-a} e[r_t^2],$$

which in turn is straightforward to evaluate using the techniques established in Section 3.3.

Similarly, $\mathbb{E}\big[p_{t+n}^2 r_t^2\big]$, where $n \geq 1$ can be expressed as

$$\mathbb{E}\big[p_{t+n}^2 r_t^2\big] = \cdots = (1 + b^2)\mathbb{E}\big[r_t^2\big] + 2ab\mathbb{E}\big[p_{t+n-1} r_t^2\big] + (a^2 + \sigma_X^2)\mathbb{E}\big[p_{t+n-1}^2 r_t^2\big],$$

which can be solved in a similar manner since $\mathbb{E}\big[p_{t+n-1} r_t^2\big]$ is now known.

The final result is listed in Appendix A.5.

A.5 Key results

In this section, we list some of the key results obtained using the techniques explained in Section 3.3.

A.5.1 First moment of returns

Exists if $|a| < 1$.

$$\mathbb{E}[r_t] = 0$$

A.5.2 Second moment of returns

Exists if $a^2 + \sigma_X^2 < 1$.

$$\mathbb{E}\big[r_t^2\big] = \frac{2\left((a-1)^2 + b^2\sigma_X^2\right)}{(a-1)\left(a^2 + \sigma_X^2 - 1\right)}$$

A.5.3 Third moment of returns

Exists if $\left|a^3 + 3\sigma_X^2 a\right| < 1$ and $a^2 + \sigma_X^2 < 1$.

$$\mathbb{E}\left[r_t^3\right] = \frac{6(2a+1)b\sigma_X^2\left((a-1)^2 + b^2\sigma_X^2\right)}{(a-1)\left(a^2+\sigma_X^2-1\right)\left(a^3+3a\sigma_X^2-1\right)}$$

$$\text{skewness } \frac{\mathbb{E}\left[r_t^3\right]}{\mathbb{E}\left[r_t^2\right]^{3/2}} = \frac{3(2a+1)b\sigma_X^2}{\sqrt{2}\left(a^3+3a\sigma_X^2-1\right)\sqrt{\frac{(a-1)^2+b^2\sigma_X^2}{(a-1)\left(a^2+\sigma_X^2-1\right)}}}$$

A.5.4 Fourth moment of returns

Exists if $a^4 + 6\sigma_X^2 a^2 + 3\sigma_X^4 < 1$ and $\left|a^3 + 3\sigma_X^2 a\right| < 1$.

$$\mathbb{E}\left[r_t^4\right] = \frac{k_1\left(k_2+k_3+k_4+k_5\right)}{(a-1)\left(a^2+\sigma_X^2-1\right)\left(a^3+3a\sigma_X^2-1\right)\left(a^4+6a^2\sigma_X^2+3\sigma_X^4-1\right)},$$

and the kurtosis is

$$\kappa = \frac{\mathbb{E}\left[r_t^4\right]}{\mathbb{E}\left[r_t^2\right]^2} = \frac{3(a-1)\left(k_2+k_3+k_4+k_5\right)\left(a^2+\sigma_X^2-1\right)}{\left(a^3+3a\sigma_X^2-1\right)\left(a^4+6a^2\sigma_X^2+3\sigma_X^4-1\right)\left((a-1)^2+b^2\sigma_X^2\right)},$$

where

$$\begin{aligned}
k_1 &= 12\left((a-1)^2 + b^2\sigma_X^2\right), \\
k_2 &= (a-1)^2\left(a^2+1\right)\left(a^2+a+1\right), \\
k_3 &= \left(2a^5 + 5a^4 - 5a^3 + \left(a^2+1\right)\left(a^2+a+1\right)b^2 + a^2 - 5a + 2\right)\sigma_X^2, \\
k_4 &= \left(6a\left(a^2+a-1\right) + (a(a(10a+9)+7)+4)b^2\right)\sigma_X^4, \\
k_5 &= 6(a+1)b^2\sigma_X^6.
\end{aligned}$$

In the case $b = 0$, kurtosis simplifies significantly to

$$\kappa = \frac{\mathbb{E}\left[r_t^4\right]}{\mathbb{E}\left[r_t^2\right]^2} = \frac{3\left(a^2+\sigma_X^2-1\right)\left(2\left(a^2+a-1\right)\sigma_X^2 + (a-1)\left(a^2+1\right)\right)}{(a-1)\left(a^4+6a^2\sigma_X^2+3\sigma_X^4-1\right)}, \quad \text{if } b = 0.$$

A.5.5 Autocorrelation of returns

Exists if $a^2 + \sigma_X^2 < 1$.

$$C_\tau(r_t) = -a^{\tau-1}\frac{1-a}{2}.$$

A.5.6 Autocorrelation of squared returns

Exists if $a^4 + 6\sigma_X^2 a^2 + 3\sigma_X^4 < 1$ and $\left|a^3 + 3\sigma_X^2 a\right| < 1$.

In the case $b = 0$, the decay is exponential. For $\tau \geq 2$,

$$C_\tau(r_t^2) = \left(a^2+\sigma_X^2\right)^{\tau-1} C_1(r_t^2), \quad \text{if } b = 0.$$

If $b \neq 0$, the autocorrelation of squared returns for $\tau \geq 2$ is

$$C_\tau(r_t^2) = \left(K_1 a^{\tau-1} + K_2\left(a^2+\sigma_X^2\right)^{\tau-1}\right) C_1(r_t^2),$$

where

$$K_1 = k_{23}/(k_{23}+k_{24}),$$
$$K_1 = k_{24}/(k_{23}+k_{24}),$$

where

$$k_{23} = -\frac{2\left(4a^2+a-2\right)b^2\sigma_X^4\left(a^2+\sigma_X^2-1\right)}{\left((a-1)a+\sigma_X^2\right)\left(a^3+3a\sigma_X^2-1\right)},$$

$$\begin{aligned}k_{24} = {} & \left((a-1)^2+\sigma_X^2\right)\left(\frac{2a\left(4a^2+a-2\right)b^2\sigma_X^2\left(a^2+\sigma_X^2-1\right)}{\left((a-1)a+\sigma_X^2\right)\left(a^3+3a\sigma_X^2-1\right)}\right. \\ & \left.+\frac{(a-1)k_{25}\left(a^2+\sigma_X^2-1\right)}{\left(a^3+3a\sigma_X^2-1\right)\left(a^4+6a^2\sigma_X^2+3\sigma_X^4-1\right)}\right) \\ & -\left(a\left(b^2-1\right)+b^2+1\right)\left((a-1)^2+\sigma_X^2\right)\end{aligned}$$

,

$$\begin{aligned}k_{25} = {} & a^6+a^5b^2+6a^5\sigma_X^2-2a^5-7a^4b^2\sigma_X^2+2a^4b^2+9a^4\sigma_X^2+a^4+30a^3b^2\sigma_X^4 \\ & -9a^3b^2\sigma_X^2+3a^3b^2+18a^3\sigma_X^4-9a^3\sigma_X^2-3a^3+21a^2b^2\sigma_X^4-6a^2b^2\sigma_X^2 \\ & +3a^2b^2+18a^2\sigma_X^4-3a^2\sigma_X^2+2a^2+18ab^2\sigma_X^6+12ab^2\sigma_X^4-6ab^2\sigma_X^2+2ab^2 \\ & -9a\sigma_X^4-12a\sigma_X^2-a+18b^2\sigma_X^6+6b^2\sigma_X^4-2b^2\sigma_X^2+b^2+3\sigma_X^2+2.\end{aligned}$$

The autocorrelation at the first lag $\tau = 1$ takes a relatively simple form for $b = 0$,

$$C_1(r_t^2) = \frac{(k_6+k_7+k_8+k_9)}{6(2a-1)\sigma_X^2\left(a^2+\sigma_X^2-1\right)+4\left(a^2+1\right)(a-1)^2}, \quad \text{if } b = 0,$$

where

$$\begin{aligned}k_6 &= (a-1)^4\left(a^2+1\right), \\ k_7 &= 2(a-1)^2\left(a\left(3a^2+a-3\right)+1\right)\sigma_X^2, \\ k_8 &= \left((12a-11)a^2+1\right)\sigma_X^4, \\ k_9 &= 6a\sigma_X^6.\end{aligned}$$

However, when $b \neq 0$, the expression is less convenient. For the lag $\tau = 1$,

$$C_1(r_t^2) = \frac{18b^2k_{12}\sigma_X^{10}+k_{18}\sigma_X^8+k_{17}\sigma_X^6+k_{12}k_{16}\sigma_X^4+k_{13}k_{14}^4k_{15}\sigma_X^2+k_{10}k_{11}k_{13}k_{14}^6}{2\left(k_{22}\sigma_X^8+k_{21}\sigma_X^6+k_{20}\sigma_X^4+k_{19}\sigma_X^2+2k_{10}k_{11}k_{13}k_{14}^4\right)},$$

where

$$
\begin{aligned}
k_{10} &= a^2 + a + 1,\\
k_{11} &= a^2 + 1,\\
k_{12} &= a^2 - 1,\\
k_{13} &= a + 1,\\
k_{14} &= a - 1,\\
k_{15} &= 6a^5 + 11a^4 - a^3 + a^2 - 7a + b^2 k_{10} k_{11} + 2,\\
k_{16} &= 30a^6 b^2 + 30a^6 - 27a^5 b^2 - 41a^5\\
&\quad - 22a^4 b^2 - 12a^4 - 22a^3 b^2 + 37a^3 + 16a^2 b^2 - 19a^2 + 5ab^2 + 6a - 4b^2 - 1,\\
k_{17} &= 78a^6 b^2 - 49a^5 b^2 - 104a^4 b^2 + 42a^4 k_{13} k_{14}\\
&\quad - 9a^3 b^2 - 33a^3 k_{13} k_{14} + 45a^2 b^2 - 8ab^2 - 3ak_{13} k_{14} + 11b^2,\\
k_{18} &= 66a^4 b^2 - 21a^3 b^2 - 72a^2 b^2 + 18a^2 k_{12} + 3ab^2 + 6b^2,\\
k_{19} &= 6a^5 k_{13} k_{14}^2 + 9a^4 k_{13} k_{14}^2 - 9a^3 k_{13} k_{14}^2 - 9ak_{13} k_{14}^2 + 2b^2 k_{10} k_{11} k_{13} k_{14}^2 + 3k_{13} k_{14}^2,\\
k_{20} &= 30a^4 b^2 k_{12} - 9a^3 b^2 k_{12} + 24a^3 k_{12} k_{14} - 6a^2 b^2 k_{12}\\
&\quad + 12a^2 k_{12} k_{14} - 12ab^2 k_{12} - 6ak_{12} k_{14} - 9b^2 k_{12} - 3k_{12} k_{14},\\
k_{21} &= 48a^3 b^2 k_{14} + 24a^2 b^2 k_{14} - 18a^2 b^2 + 18a^2 k_{13} k_{14} - 9ab^2 - 9ak_{13} k_{14} + 9b^2,\\
k_{22} &= 18a^2 b^2 - 9ab^2 - 18b^2.
\end{aligned}
$$