The Econometric Analysis of MS Models

Youwei Li, Bas Donkers, Bertrand Melenberg

June 25, 2004

Abstract

This paper studies how to compare different microscopic simulation (MS) models and how to compare a MS model with the real world. The parameters of interest are classified and characterized, and various econometric methods are applied to make the comparison. We illustrate the methodology by comparing various specifications of the MS model developed by Levy, Levy, and Solomon (2000) and by comparing this MS model with the real world.

1 Introduction

In financial markets, the observable behavior and phenomena are the consequences of aggregated individual movements at the macro level, but the determinants lie at the micro level. In general, it is very difficult to describe the individual behavior (decision making under risk and uncertainty), and the implied aggregated phenomena explicitly: economics, including financial markets, is a complex system. It is very difficult, if not impossible, to find analytical solutions for such systems. In order to get some insight into it, a possible approach is to do Microscopic Simulation (MS). The idea is to study complex systems by representing each of the microscopic elements individually (on a computer) and by simulating the behavior of the entire system, keeping track of all of the individual elements and their interactions over time. Throughout the simulation, the macroscopic variables that are of interest can be recorded, and their dynamics can be investigated.

The growing literature of MS in finance has resulted in various competing microscopic simulation models to explain observed phenomena in real-life financial markets. The works of Arthur et al. (1997), Chiarella and He (2002),LeBaron (2000), Levy et al. (2000), Lux (1998), among others, provide good examples of various MS models. So far, research has mainly focused on investigating whether a model shares some important characteristics of the actual financial markets, the stylized facts, such as short-term momentum, excess volatility, heavy trading volume, a positive correlation between volume and contemporaneous absolute returns, endogenous market crashes, etc. The typical way to do this is by running a 'representative realization' of the MS model, the recorded variables, such as stock returns, which can be analyzed by the standard financial econometric techniques as described in Cochrane (2001). Although much work has already been done along these lines (see, for instance, Levy et al., 2000, and Lux, 2000), to our knowledge, the systematic procedures to investigate the difference between two MS models, and to judge whether a MS model is realistic or not have not yet been developed.

The aim of this paper is to develop and apply econometric techniques to compare different microscopic simulation models, and, more importantly, to compare the MS model generated data with real life data. When comparing different models, important factors that drive MS economies can be detected and investigated. The statistical tools to compare different MS models can also easily be used to check the robustness of the outcomes of a MS model with respect to its initial conditions and parameter settings. In this way we might gain a better understanding of the underlying mechanism of MS models. When comparing an MS model with real life data, factors might be identified, which have to be adapted or integrated to create more realistic models, whose implications are comparable to the empirical findings in the financial markets. Confronting a MS model with real life data is not only a way to check the "realism" of the model, which will enhance our knowledge of financial markets, but it is also an essential step from a practical point of view. For example, when MS models are used to evaluate the impacts of government policies, or to forecast, we need to link the MS models with real life data.

Given a MS model and a set of model parameters, various outcomes, such as prices, returns, volatility, etc. of the MS economy can be generated. When comparing two different models, we first need to identify which outcomes are of interest and are to be used to describe the behavior of the MS economy. Next, we should specify the characteristics of the variables that we want to compare. For example, in the empirical finance literature, stock returns have been studied intensively. Besides simple descriptive statistics, many studies focus on, for example, the predictability of stock returns, excess volatility, and volatility clustering, etc.. So, the outcomes one might have in mind would be the time series behavior of the stock returns in the MS economy. To compare two MS economies, simple descriptive statistics, such as the mean and median, can be compared, but also the time series properties, such as autocorrelations. These statistics can be compared jointly, resulting in a test for a particular set of characteristics.

In principle, the characteristics of the MS economy can be retrieved with an arbitrary level of precision. Because we can run the MS model independently many times, the distribution functions of these characteristics can be obtained. The only uncertainty here is which of the states of nature has been realized in the simulations. Another type of uncertainty, the sampling uncertainty, arises when we are concerned with the question of how well a particular MS model fits real life data. The test procedures should take into account the sampling error in real life data which, for example, can be resolved by econometric methods. To compare on the basis of single parameters, such as the mean return or the level of first order autocorrelation, estimation uncertainty from the real life data can usually be quantified using analytical results, for example, the autocorrelation coefficient is obtained from a linear regression model with well known properties. Ideally, we should first estimate or calibrate the model on the basis of available data and then test whether the resulting model describes the actual data sufficiently well. Here, we will not estimate the model, but instead, we investigate the properties of simulated data and develop a comparison methodology.

The remainder of the paper is organized as follows. We give an econometric characterization of simulated data in Section 2. Section 3 is devoted to how to compare two different MS economies and Section 4 focuses on comparing MS economy with the real world. We illustrate how to apply the methods that are developed in the previous sections in Section 5 and we conclude in Section 6.

2 Econometric background

In this section, we start with a discussion of some properties of simulated data, then we characterize the simulated data from an econometric point of view.

2.1 Properties of simulated data

A MS model consists of inputs, a designed mechanism of the system, and outputs. Inputs include parameters, initial conditions, and also noise; the designed mechanism describes the functioning of the system, and how the dynamics evolve over time; the outputs are the observations of variables of interest. For instance, in a MS model of the stock market, Levy, Levy and Solomon (see Levy et al., 2000) set up a MS model (the LLS model from now on) in which the microscopic elements are individual investors. These investors make their decision according to standard utility maximization and they interact via buying and selling stocks and bonds within a temporary equilibrium mechanism. The investors decide upon the proportion of their wealth that they will invest in assets as a function of price. The price in each period is generated by equalizing that period's aggregated demand and supply. With this new temporary equilibrium price, the investors' expectations for asset returns will be updated, and price dynamics arise. In such a way, macrovariables, such as stock prices, stock returns, the distribution of total wealth among subgroups of investors, etc., can be recorded and, subsequently, studied.

We assume that the observable outcomes of interest, the state of a MS economy at time t, can be represented by a vector $x_t \in \mathbb{R}^K$ for some K. This state changes over time according to a (possibly) noisy law of motion, which depends on the designed mechanism, represented by a stochastic process $\{v_t\}$, so

$$x_t = G(x_{t-1}, v_t).$$

The function G is often assumed to be smooth. We might know the form G explicitly or implicitly, depending on the design of the MS model.

We use N to denote the number of simulation runs of a MS model, and use T to denote the number of periods of the outcomes that are observed for each run. Let $\{x_{n,t}\}$, $t = 1, ..., T$, be the observed series of a MS model for the nth simulation run, which is one realization of the stochastic process $\{x_t\}$. In order to make life easier, we can implement the simulation of the MS model such that the realizations of υ_t are independent over simulations. Throughout this paper we will make therefore the following assumption:

Assumption 1 *The outcomes of different simulations of a MS model are stochastically independent*[1].

Very often we also impose the assumptions that

Assumption 2 *The process $\{x_t\}$ is strictly stationary.*

Assumption 3 *The process $\{x_t\}$ is ergodic.*

These latter assumptions can be made less restrictive by making appropriate transformations. For instance, in real life, asset returns are more likely to satisfy these assumptions than asset prices. In MS models reflecting reality to some extent, the same might be the case, so that then an investigation in terms of asset returns is to be preferred over asset prices.

Given the outcomes of interest of MS models, we need to specify characteristics that can describe the outcome series properly. The interesting characteristics of the outcome series of a MS model often include simple descriptive statistics, such as means, medians, and variances at given time points, or mean, median, and variance of the whole observed periods; or they include the basic time series characteristics, such as autocorrelation coefficients. For application in finance, they can also include the statistics that describe the stylized facts of financial markets, such as fat tails, volatility clustering, excess volatility, and GARCH effects, etc. Statistically, these characteristics can be described as functionals of the distribution of the underlying process $\{x_t\}$.

In principle, all the parameters that we are interested in can be retrieved with an arbitrary level of precision under Assumption 1. Because by running the MS model independently many times, the distribution functions that are related to these parameters can be approximated arbitrarily closely. We notice the feature that in MS models, the outcome series $\{x_{n,t}\}$ is observable along both the dimension N and T. So, the parameters may be retrieved in different ways, for some of them, the asymptotic distributions can be obtained as N or T goes to infinity, and for some others, the asymptotic distributions can be obtained when both N and T go to infinity. Furthermore, if a parameter can be retrieved in different ways, a natural question that arises is which way is the most efficient one, either from a statistical point of view, or from a computational point of view, or both. For instance, when we consider the mean at a given time point

[1] For the resulting time series $\{x_t\}$, $t = 1, ...,$ of a single simulation, there are some attempts to distinguish between (noisy) deterministic chaos and randomness of underlying data generating process. Brock (1986) and Barnett and Serletis (2000) provide examples for such tests, but they found that the related tests are sensitive to the noise. In general, the theory of distinguishing noisy deterministic chaos and randomness based upon time series observations has not been established yet if the distinction is possible at all, see Dechert and Hommes (2000). Thus, in this paper we will focus more on the stochastic properties of observed series.

t_0, $t_0 \leq T$, without any other assumptions, besides its existence, it can only be retrieved when N goes to infinity. When we consider the mean over time, it can be retrieved as T goes to infinity. Under the assumptions of stationarity and ergodicity, these two parameters are equal, and can be retrieved when either N, T, or both, go to infinity. In order to distinguish these different situations and to get either statistically or computationally efficient estimators, it is important to classify the estimators of the parameters of interest first.

2.2 The econometric characterization

First, we introduce some notation. For observations $\{x_{n,t}\}$, $n = 1, ..., N$, $t = 1, ..., T$, we set

$$\mathbf{x}_{n,t}^{\tau}=(x_{n,t}, ..., x_{n,t+\tau-1})$$

For each $t \in \{1, ..., T-\tau+1\}$, we denote the empirical distribution function based upon $\{\mathbf{x}_{n,t}^{\tau}\}_{n=1}^{N}$ by $\widehat{F}_{t,N,\tau}$ and for each $n \in \{1, ..., N\}$ we denote by $\widehat{F}_{n,T,\tau}$ the empirical distribution function based upon $\{\mathbf{x}_{n,t}^{\tau}\}_{t=1}^{T+\tau-1}$, i.e.,

$$\begin{aligned} \widehat{F}_{t,N,\tau}(\mathbf{z}) &= \frac{1}{N}\sum_{n=1}^{N} 1_{(-\infty,\mathbf{z}]}(\mathbf{x}_{n,t}^{\tau}), \\ \widehat{F}_{n,T,\tau}(\mathbf{z}) &= \frac{1}{T+\tau-1}\sum_{t=1}^{T+\tau-1} 1_{(-\infty,\mathbf{z}]}(\mathbf{x}_{n,t}^{\tau}), \end{aligned}$$

where $1_{(-\infty,\mathbf{z}]}(\mathbf{x}_{n,t})$ is the usual indicator function such that $1_{(-\infty,\mathbf{z}]}(\mathbf{x}_{n,t}) = 1$ for $\mathbf{x}_{n,t} \leq \mathbf{z}$, and $1_{(-\infty,\mathbf{z}]}(\mathbf{x}_{n,t}) = 0$ otherwise, where $\mathbf{x}_{n,t}$ and $\mathbf{z}$ are vectors in $\mathbb{R}^{\tau}$, and $\leq$ is defined componentwise. The empirical distribution function based on all observations will be denoted by $\widehat{F}_{N,T,\tau}$, i.e.,

$$\widehat{F}_{N,T,\tau}(\mathbf{z}) = \frac{1}{N(T+\tau-1)}\sum_{n=1}^{N}\sum_{t=1}^{T+\tau-1} 1_{(-\infty,\mathbf{z}]}(\mathbf{x}_{n,t}^{\tau})$$

When τ is understood, we simplify

$$\mathbf{x}_{n,t} = \mathbf{x}_{n,t}^{\tau},\ \widehat{F}_{t,N} = \widehat{F}_{t,N,\tau},\ \widehat{F}_{n,T} = \widehat{F}_{n,T,\tau},\ \widehat{F}_{N,T} = \widehat{F}_{N,T,\tau}.$$

Under assumption 1 we have that the distribution of $\mathbf{x}_{n,t}$ does not depend on n. So, when is τ understood, we write

$$\mathbf{x}_t = \mathbf{x}_{n,t} = \mathbf{x}_{n,t}^{\tau},$$

and we denote by $F_{\mathbf{x}_t}$ the distribution function of $\mathbf{x}_t$. Notice that under the assumption of stationarity the distribution function $F_{\mathbf{x}_t}$ does not depend on t, and then we shall simply write $F_{\mathbf{x}}$.

As parameters of interest we investigate $\theta_{\mathbf{x}_t} \in \mathbb{R}^k$ as a function of the distribution function $F_{\mathbf{x}_t}$, i.e.,

$$\theta_{\mathbf{x}_t} := \varphi_t(F_{\mathbf{x}_t})$$

given some function $\varphi_t : D_{\varphi_t} \subseteq D(R^\tau) \to R^k$, where $D(R^\tau)$ is the set of all non-decreasing right continuous functions $z : R^\tau \to R$ such that $z(-\infty) = 0$, $z(\infty) = 1$, equipped with the uniform norm, and D_{φ_t} is a subset of $D(R^\tau)$, the relevant domains of the function φ_t. In case of stationarity, $\theta_{\mathbf{x}_t}$ will be independent of t if we take φ_t independent of t, and then we write

$$\theta_{\mathbf{x}} := \varphi(F_{\mathbf{x}}).$$

Typical examples of $\theta_{\mathbf{x}_t}$ are means, variances, or covariances. For instance, when $\tau = 1$, so that $\mathbf{x}_t$=$x_{n,t}$, and we consider $\theta_{\mathbf{x}_t} = E(x_{n,t})$, then the function $\varphi_t : D_\varphi \subseteq D(R) \to R$ is defined as $\varphi_t(F_{\mathbf{x}_t}) = \int x_{n,t} dF_{\mathbf{x}_t}(x_{n,t})$, where the subset D_{φ_t} is such that the mean is well defined. When we consider $\theta_{\mathbf{x}_t} = Var(x_{n,t})$, then the function $\varphi_t : D_{\varphi_t} \subseteq D(R) \to R$ is defined as $\varphi_t(F_{\mathbf{x}_t}) = \int (x_{n,t} - \int x_{n,t} dF_{\mathbf{x}_t}(x_{n,t}))^2 dF_{\mathbf{x}_t}(x_{n,1})$, and D_{φ_t} is such that the second moment exists.

Let us now first consider estimation of $\theta_{\mathbf{x}_t}$ for some given t. If we can estimate $F_{\mathbf{x}_t}$ consistently and obtain its limit distribution, then, under certain conditions on φ_t, $\theta_{\mathbf{x}_t}$ can also be estimated consistently, and its limit distribution can be derived by using an appropriate version of the delta method.

In principle, $F_{\mathbf{x}_t}$ can be estimated consistently by $\widehat{F}_{t,N} = \widehat{F}_{t,N,\tau}$ as $N \to \infty$. By Donsker's theorem (for instance, Theorem 19.3, Van der Vaart, 1998), it follows that

$$\sqrt{N}(\widehat{F}_{t,N} - F_{\mathbf{x}}) \overset{dist.}{\to} G_{F_{\mathbf{x}_t}},$$

where $G_{F_{\mathbf{x}_t}}$ is a Gaussian process in $D(R^\tau)$ specified by

$$\begin{aligned} E(G_{F_{\mathbf{x}_t}}(\mathbf{t})) &= 0, \ \mathbf{t} \in R^\tau \\ E(G_{F_{\mathbf{x}_t}}(\mathbf{t}_i) G_{F_{\mathbf{x}_t}}(\mathbf{t}_j)) &= F_{\mathbf{x}_t}(\mathbf{t}_i \wedge \mathbf{t}_j) - F_{\mathbf{x}_t}(\mathbf{t}_i) F_{\mathbf{x}_t}(\mathbf{t}_j), \end{aligned} \tag{1}$$

where $\mathbf{t}_i$, $\mathbf{t}_j \in R^\tau$, and $\mathbf{t}_i \wedge \mathbf{t}_j$ denotes the componentwise minimum of $\mathbf{t}_i$ and $\mathbf{t}_j$.

Now, we assume (see Section 20.2, Van der Vaart, 1998)

Assumption 4 *The function $\varphi_t : D_{\varphi_t} \subseteq D(\mathbb{R}^\tau) \to \mathbb{R}^k$ is defined on a subset D_{φ_t} of $D(\mathbb{R}^\tau)$ that contains $F_{\mathbf{x}_t}$ and is Hadamard differentiable at $F_{\mathbf{x}_t}$, where its derivative at $F_{\mathbf{x}_t}$, which is denoted by $\varphi'_{F_{\mathbf{x}_t}}$, $\varphi'_{F_{\mathbf{x}_t}} : D(\mathbb{R}^\tau) \to \mathbb{R}^k$, is a continuous linear map.*

Given this assumption, consider as an estimator of $\theta_{\mathbf{x}_t} = \varphi(F_{\mathbf{x}_t})$ its sample analogue

$$\widehat{\theta}_{\mathbf{x}_t} := \varphi_t(\widehat{F}_{t,N}).$$

Then it follows from the functional Delta method (see, for example, Theorem 20.8, Van der Vaart, 1998) that under Assumption 4

$$\sqrt{N}(\varphi_t(\widehat{F}_{t,N}) - \varphi(F_{\mathbf{x}_t})) = \varphi'_{F_{\mathbf{x}_t}}(\sqrt{N}(\widehat{F}_{t,N} - F_{\mathbf{x}_t})) + o_P(1).$$

Because $\varphi'_{F_{\mathbf{x}_t}}$ is continuous, this means, as a consequence of Riesz' representation Theorem (see, for example, result IV.6.3, Dunford and Schwartz, 1957), that there exists a function $\zeta_t : \mathbb{R}^\tau \to \mathbb{R}^k$, such that

$$\begin{aligned}
\varphi'_{F_{\mathbf{x}_t}}(\sqrt{N}(\widehat{F}_{t,N} - F_{\mathbf{x}_t})) &= \int \zeta_t(z)\, d\left(\sqrt{N}(\widehat{F}_{t,N} - F_{\mathbf{x}_t})\right)(z) \\
&= \sqrt{N}\frac{1}{N}\sum_n (\zeta_t(\mathbf{x}_{n,t}) - E(\zeta_t(\mathbf{x}_{n,t}))) \qquad (2) \\
&\overset{dist.}{\to} N(0, Var\{\zeta_t(\mathbf{x}_{n,t})\}). \qquad (3)
\end{aligned}$$

Next, we study the stationarity case (assumption 2) where we have $F_{\mathbf{x}} = F_{\mathbf{x}_t}$, and where we choose φ_t time-independent, so that $\theta_{\mathbf{x}} = \theta_{\mathbf{x}_t}$. As possible estimators, we then have

$$\widehat{\theta}^t_{\mathbf{x}} := \widehat{\theta}_{\mathbf{x}_t} = \varphi(\widehat{F}_{t,N})$$

for $t = 1, ..., T - \tau + 1$. In addition, it then makes sense to consider as estimator, for instance, the time average (indicated by the superscript Ti)

$$\widehat{\theta}^{Ti}_{\mathbf{x}} := \frac{1}{T - \tau + 1}\sum_{t=1}^{T-\tau+1} \widehat{\theta}^t_{\mathbf{x}}.$$

(Of course, other choices are possible as well; we consider this one for illustrative purposes.) Using

$$\sqrt{N}(\widehat{\theta}^t_{\mathbf{x}} - \theta_{\mathbf{x}}) = \sqrt{N}\frac{1}{N}\sum_n (\zeta(\mathbf{x}_{n,t}) - E\zeta(\mathbf{x}_{n,t})) + o_P(1)$$

with ζ now also time-independent, and denoting Γ_i as the ith order autocovariance of the series $\{\zeta(\mathbf{x}_{n,t}) - E\zeta(\mathbf{x}_{n,t})\}_{n=1}^N$, we get

$$\sqrt{N}(\widehat{\theta}^{Ti}_{\mathbf{x}} - \theta_{\mathbf{x}}) \overset{dist.}{\to} N\left(0, \frac{\Gamma_0}{T_\tau} + 2\sum_{i=1}^{T_\tau - 1}\frac{T_\tau - i}{T_\tau^2}\Gamma_i\right),$$

with $T_\tau = T - \tau + 1$, as $N \to \infty$.

Furthermore, in case of both Assumptions 2 and 3, the distribution function $F_{\mathbf{x}} = F_{\mathbf{x}_t}$ can also be estimated by $\widehat{F}_{n,T}$, and the parameter $\theta_{\mathbf{x}} = \varphi(F_{\mathbf{x}})$ can then be estimated by $\widehat{\theta}^n_{\mathbf{x}} := \varphi(\widehat{F}_{n,T})$, $n = 1, ..., N$. By the Riesz Representation Theorem, we have

$$\begin{aligned}
\sqrt{T}(\widehat{\theta}^n_{\mathbf{x}} - \theta_{\mathbf{x}}) &= \varphi'_{F_{\mathbf{x}}}(\sqrt{T}(\widehat{F}_{n,T} - F_{\mathbf{x}})) + o_P(1) \\
&= \int \zeta d\sqrt{T}(\widehat{F}_{n,T} - F_{\mathbf{x}}) + o_P(1) \\
&= \sqrt{T}\frac{1}{T}\sum_t (\zeta(\mathbf{x}_{n,t}) - E\zeta(\mathbf{x}_{n,t})) + o_P(1). \qquad (4)
\end{aligned}$$

Because ζ is a measurable function, under Assumption 2, 3, the process $\zeta(\mathbf{x}_{n,t})$ is also strictly stationary and ergodic (see, for instance,Pagan and Ullah, 1999, p371), we can use the CLT to derive the asymptotic normality. Next, we define as estimator of $\theta_{\mathbf{x}}$ the average of $\widehat{\theta}_{\mathbf{x}}^{n}$ over n, i.e.,[2]

$$\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si}} := \frac{1}{N}\sum_{n=1}^{N}\widehat{\theta}_{\mathbf{x}}^{n},$$

where the superscript Si indicates that this estimator is an average over independent simulations. Notice the difference between $\widehat{\theta}_{\mathbf{x}}^{Ti}$ and $\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si}}$: the estimator $\widehat{\theta}_{\mathbf{x}}^{Ti}$ is a time average of T_τ estimators, each of these estimators is estimated from independent observations of independent simulations. The estimator $\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si}}$ is an average over N simulations, where each of the estimators is estimated from observations of one realization of the MS model. We have under assumptions 1, 2, 3, and 4, that the parameter $\theta_{\mathbf{x}}$ can be estimated consistently by $\widehat{\theta}_{\mathbf{x}}^{Ti}$, with

$$\sqrt{N}(\widehat{\theta}_{\mathbf{x}}^{Ti} - \theta_{\mathbf{x}}) \overset{dist.}{\to} N\left(0, \frac{\Gamma_0}{T_\tau} + 2\sum_{i=1}^{T_\tau - 1}\frac{T_\tau - i}{T_\tau^2}\Gamma_i\right),$$

as $N \to \infty$, and for the estimator $\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si}}$, we have

$$\sqrt{T}(\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si}} - \theta_{\mathbf{x}}) \overset{dist.}{\to} N\left(0, \frac{1}{N}\left(\Gamma_0 + 2\sum_{i=1}^{\infty}\Gamma_i\right)\right),$$

as $T \to \infty$. When both N and T tend to infinity, we get

$$\sqrt{NT}(\widehat{\theta}_{\mathbf{x}}^{Ti} - \theta_{\mathbf{x}}) \overset{dist.}{\to} N\left(0, \Gamma_0 + 2\sum_{i=1}^{\infty}\Gamma_i\right)$$
$$\sqrt{NT}(\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si}} - \theta_{\mathbf{x}}) \overset{dist.}{\to} N\left(0, \Gamma_0 + 2\sum_{i=1}^{\infty}\Gamma_i\right)$$

Moreover, when both N and T tend to infinity, we can construct as estimator

$$\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si},Ti} = \varphi\left(\widehat{F}_{N,T}\right).$$

In this case we find

$$\begin{aligned}\sqrt{NT}(\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si},Ti} - \theta_{\mathbf{x}}) &= \varphi'_{F_x}\left(\sqrt{NT}\left(\widehat{F}_{N,T} - F_{n,T}\right)\right) + 0_p(1) \\ &= \sqrt{NT}\frac{1}{NT}\sum\sum\left(\zeta\left(\mathbf{x}_{n,t}\right) - E\left(\zeta\left(\mathbf{x}_{n,t}\right)\right)\right) + 0_p(1)\end{aligned}$$

so that

$$\sqrt{NT}(\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si},Ti} - \theta_{\mathbf{x}}) \overset{dist.}{\to} N\left(0, \Gamma_0 + 2\sum_{i=1}^{\infty}\Gamma_i\right).$$

[2] Again, other choices are possible as well; we consider this one for illustrative purposes.

We see that the three estimators, $\widehat{\theta}_{\mathbf{x}}^{\text{Si}}$, $\widehat{\theta}_{\mathbf{x}}^{Ti}$, $\widehat{\theta}_{\mathbf{x}}^{\text{Si},Ti}$, although generally different as long as φ is nonlinear as function of F, asymptoticaly are (first order) equivalent to each other.

For a specific parameter, the function ζ need to be determined specifically. A special case arises when we consider parameters that take the form $\theta_{\mathbf{x}} = \psi(\int gdF_{\mathbf{x}})$, where ψ and g satisfy

Assumption 5 *The function $g : \mathbb{R}^{\tau} \to \mathbb{R}^{l}$ is squared integrable, and the transformation $\psi : \mathbb{R}^{l} \to \mathbb{R}^{k}$ is continuously differentiable.*

Under Assumption 5, it can be readily checked that

$$\zeta = \psi' \circ g.$$

Now we turn to discuss some examples.

Example 1 *The mean $E(x_t)$*

The mean $\mu = E(x_t)$ can be rewritten as $\mu = \int x_t dF_{x_t} := \varphi(F_{x_t})$, the function φ is defined as $\varphi : D_{\varphi} \subseteq D(\mathbb{R}) \to \mathbb{R}$, where $D_{\varphi} = \{F \in D(\mathbb{R}),\ \int zdF(z) < \infty\}$. We know that when T is fixed, and N tends to infinity, the mean μ can be estimated by the estimators

$$\widehat{\mu}^{t} := \frac{1}{N}\sum_{n=1}^{N} x_{n,t},\ t \in \{1, 2, ..., T\},$$

and

$$\widehat{\mu}^{Ti} := \frac{1}{\tau}\sum_{t=1}^{\tau} \widehat{\mu}^{t}, \quad \tau \in \{1, 2, ..., T\},$$

we know that

$$\sqrt{N}(\widehat{\mu}^{Ti} - \mu) \to N\left(0, \frac{1}{\tau}\gamma_0 + 2\sum_{i=1}^{\tau-1}\frac{\tau - i}{\tau^2}\gamma_i\right),$$

for $\tau \in \{1, 2, ..., T\}$, as $N \to \infty$. When N is fixed and T tends to infinity, we define the estimators that based upon observations of the nth simulation as

$$\widehat{\mu}^{n} := \frac{1}{T}\sum_{t} x_{n,t},\ n \in \{1, 2, ..., N\},$$

and

$$\widehat{\mu}^{\text{Si}} := \frac{1}{N}\sum_{n=1}^{N}\widehat{\mu}^{n},$$

We then have

$$\sqrt{T}(\widehat{\mu}^{\text{Si}} - \mu) \to N\left(0, \frac{1}{N}\left(\gamma_0 + 2\sum_{i=1}^{\infty}\gamma_i\right)\right), \tag{5}$$

as $N \to \infty$.

It is easy to see that when both T and N tend to infinity, we have

$$\sqrt{TN}(\widehat{\mu}^{\mathrm{Si}} - \mu) \to N\left(0, \gamma_0 + 2\sum_{i=1}^{\infty}\gamma_i\right).$$

We note that $\widehat{\mu}^{Ti}$ and $\widehat{\mu}^{\mathrm{Si}}$ are the same estimators of mean.

We summarize the limit distribution of above estimators in the following table

Table 1: The limit distribution of the estimators of mean

$N \to \infty, T < \infty$	$\sqrt{N}(\widehat{\mu}^{Ti} - \mu) \to N(0, \frac{1}{\tau}\gamma_0 + 2\sum_{i=1}^{\tau-1}\frac{\tau-i}{\tau^2}\gamma_i)$
$N < \infty, T \to \infty$	$\sqrt{T}(\widehat{\mu}^{\mathrm{Si}} - \mu) \to N(0, \frac{1}{N}(\gamma_0 + 2\sum_{i=1}^{\infty}\gamma_i))$
$N \to \infty, T \to \infty$	$\sqrt{TN}(\widehat{\mu}^{\mathrm{Si}} - \mu) \to N(0, \gamma_0 + 2\sum_{i=1}^{\infty}\gamma_i)$

These results can be used to investigate how to do the simulations. This will be discussed in Section 2.3.

Example 2 *The AR coefficients*

Define γ_j as jth order autocovariance of $\{x_t\}$, then the jth order autocorrelation coefficient is $\beta_j = \gamma_j/\gamma_0$. For $\mathbf{x} = (x_0, x_j)$, the parameter autocorrelation coefficient $\beta_j = \gamma_j/\gamma_0 = cov(x_0, x_j)/var(x_0)$ takes the form $\theta_{\mathbf{x}} = \psi(\int gdF_{\mathbf{x}})$, and the functions ψ and g satisfy the Assumption 5. In fact, we can define $g : \mathbb{R}^2 \to \mathbb{R}^2$,

$$g(x_0, x_j) = ((x_0 - \int\int x_0 dF_{(x_0,x_j)})(x_j - \int\int x_j dF_{(x_0,x_j)}), (x_0 - \int\int x_0 dF_{(x_0,x_j)})^2)$$

and $\psi : \mathbb{R}^2 \to \mathbb{R}$, $\psi(z_1, z_2) = z_1/z_2$.

We notice that when T is fixed and N tends to infinity. If $T \geq j$, we define $\widehat{\gamma}_0^t$, $\widehat{\gamma}_j^t$ as the sample analogue of $\gamma_0^t = var\{x_t\}$, $\gamma_j^t = cov(x_t, x_{t+j}\}$. Then for each t, we can define estimators of β_j as

$$\widehat{\beta}_j^t := \frac{\widehat{\gamma}_j^t}{\widehat{\gamma}_0^t},$$

and

$$\widehat{\beta}_j^{Ti} := \frac{1}{T-j}\sum_{t=1}^{T-j}\widehat{\beta}_j^t.$$

We know from the previous analysis that when N goes to infinity,

$$\sqrt{N}(\widehat{\beta}_j^{Ti} - \beta_j) \to N(0, V_1),$$

where

$$V_1 = \lim_{N\to\infty} Var\left\{\frac{1}{T-j}\sum_{t=1}^{T-j}\widehat{\beta}_j^t\right\},$$

for the detailed expression of V_1, see the Appendix.

Similarly, we can also study the parameter β_j in case that N is fixed and T tends to infinity. Under assumption 2 and $E\{\varepsilon_t x_{t-j}\} = 0$, β_j can be estimated by the OLS estimator $\widehat{\beta}_j^n$ from

$$x_t = \alpha_j + \beta_j x_{t-j} + \varepsilon_t.$$

Asymptotically we have,

$$\sqrt{T}(\widehat{\beta}_j^n - \beta_j) \to N(0, Q^{-1}SQ^{-1}),$$

where

$$Q = p\lim_{T\to\infty}\frac{1}{T}\sum_{t=j+1}^{T} x_{t-j}^2 = Ex_t^2,$$

and S is the asymptotic covariance of the scaled sample mean $\frac{1}{\sqrt{T}}\sum_{t=j+1}^{T} x_{t-j}\varepsilon_t$,

$$\begin{aligned} S &= \lim_{T\to\infty} Var\frac{1}{\sqrt{T}}\sum_{i=0}^{T-1} x_{t-j-i}\varepsilon_{t-i} \\ &= E\{x_{t-j}^2\varepsilon_t^2\} + 2\sum_{i=1}^{\infty} E\{x_{t-j}x_{t-j-i}\varepsilon_t\varepsilon_{t-i}\}. \end{aligned}$$

We define

$$\widehat{\beta}_j^{\mathrm{Si}} := \frac{1}{N}\sum_{n=1}^{N}\widehat{\beta}_j^n,$$

then we get

$$\sqrt{T}(\widehat{\beta}_j^{\mathrm{Si}} - \beta_j) \to N(0, \frac{1}{N}Q^{-1}SQ^{-1}) \qquad (6)$$

as $T \to \infty$.

When both N and T go to infinity, we know from (2) that

$$\sqrt{TN}(\widehat{\beta}_j^{\mathrm{Si}} - \beta_j) \to N(0, Q^{-1}SQ^{-1}).$$

It can be prove that when both N and T go to infinity, the estimator $\widehat{\beta}_j^{Ti}$ and $\widehat{\beta}_j^{\mathrm{Si}}$ have the same asymptotic variance, see the appendix for a proof.

We summarize the limit distribution of the above estimators in the following table

Up to now, we assumed that τ is fixed, $1 \le \tau \le T < \infty$, and $\mathbf{x} = (x_1, ..., x_\tau)$. When $\tau = T = \infty$, we have $\mathbf{x} = (x_1, x_2, ...)$, and the parameter $\theta_{\mathbf{x}}$ depends on

Table 2: The limit distribution of the estimators of AR coefficients

$N \to \infty, T < \infty$	$\sqrt{N}(\widehat{\beta}_j^{Ti} - \beta_j) \to N(0, V_1)$
$N < \infty, T \to \infty$	$\sqrt{T}(\widehat{\beta}_j^{\mathrm{Si}} - \beta_j) \to N(0, \frac{1}{N} Q^{-1} S Q^{-1})$
$N \to \infty, T \to \infty$	$\sqrt{TN}(\widehat{\beta}_j^{\mathrm{Si}} - \beta_j) \to N(0, Q^{-1} S Q^{-1})$

the distribution function of the process $\{x_t\}$ over the whole time axis. For instance, the parameter $\theta_{\mathbf{x}}$ might be the integrated order d in the ARFIMA(p, d, q) process,. In this case, it does not make sense if we only require $N \to \infty$, because the parameter always depends on $F_{\mathbf{x}}$. Very often, under distributional assumptions, the probability density function can be written as a function of the parameters, such as d, and the observations. Accordingly, under smoothness condition for the probability density function, the parameter of interest can still be expressed as $\theta_{\mathbf{x}} = \varphi(F_{\mathbf{x}})$, although we may not be able to write the function φ explicitly. For instance, the parameter $\theta_{\mathbf{x}}$ may be the maximizer of the log-likelihood function, and estimation can be performed via maximum likelihood. Under conditions, such as those in Theorem 5.39 of Van der Vaart (1998), the asymptotic normality of the maximum likelihood estimator $\widehat{\theta}_{\mathbf{x}}^n$ of $\theta_{\mathbf{x}}$ for each simulation n can be derived, and the corresponding results for the estimator $\widetilde{\theta}_{\mathbf{x}}^{\mathrm{Si}}$, which is the average of $\widehat{\theta}_{\mathbf{x}}^n$ over independent simulations, comes out straightforwardly. So we have, under assumptions of Theorem 5.39 of Van der Vaart (1998),

$$\sqrt{T}(\widehat{\theta}_{\mathbf{x}}^n - \theta_{\mathbf{x}}) \overset{dist.}{\to} N\left(0, I_{\theta_{\mathbf{x}}}^{-1}\right),$$

as $T \to \infty$, where $I_{\theta_{\mathbf{x}}}$ is the Fisher information matrix, and moreover

$$\sqrt{TN}(\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si}} - \theta_{\mathbf{x}}) \to N\left(0, I_{\theta_{\mathbf{x}}}^{-1}\right)$$

as $T \to \infty$, and $N \to \infty$.

Example 3 *The ARFIMA processes*

Granger (1980) and Hosking (1981) introduced the ARFIMA(p, d, q) process

$$\Phi(L)(1 - L)^d x_t = \Theta(L)\varepsilon_t,$$

where $d \in (-\frac{1}{2}, \frac{1}{2}]$; L is the lag operator, and the fractional difference operator $(1 - L)^d$ is defined by

$$(1 - L)^d := \sum_{j=0}^{\infty} \begin{pmatrix} d \\ j \end{pmatrix} (-1)^j L^j;$$

L^j is the composition of j lag operators, $\Phi(L)$ and $\Theta(L)$ are lag polynomials with order p and q respectively,

$$\Phi(L) = 1 + A(1)L + A(2)L^2 + ... + A(p)L^p$$

and

$$\Theta(L) = 1 + M(1)L + M(2)L^2 + ... + M(q)L^q.$$

A process x_t is said to be fractionally integrated, if, after applying the operator $(1-L)^d$, it follows an ARMA(p, q) process. Generally, it is assumed that the roots of $\Phi(x)$ are simple, and the roots of $\Phi(x)$ and $\Theta(x)$ are outside the unit circle, and $\varepsilon_t \sim IIDN(0, \sigma^2)$. It is proved in Granger (1980) and Hosking (1981) that when $d \in (-\frac{1}{2}, \frac{1}{2}]$, x_t is stationary and ergodic. For $0 < d < \frac{1}{2}$, the process has long memory in the sense that its autocovariances are eventually positive and decay slowly (at a hyperbolic rate). For $-\frac{1}{2} < d < 0$, the autocovariances are eventually negative and decline slowly.

Sowell (1992) derives the MLE estimator for the parameter

$$\theta' = (d\ A(1)...A(p)\ M(1)...M(q)\ \sigma^2)$$

of the above ARFIMA(p, d, q) model. The log-likelihood is

$$\mathcal{L} = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\log(2\pi)\log|\Sigma| - \frac{1}{2}(\mathbf{x})'\Sigma^{-1}\mathbf{x},$$

where $\{\Sigma\}_{ij} = \gamma_{|i-j|}$ and $\mathbf{x}$ is the T-dimensional vector of the observations of x_t for simulation n. The parameter d can still be expressed as a function of $F_{\mathbf{x}}$, i.e., $d = \varphi(F_{\mathbf{x}})$, and we know from the implicit function theorem that the function φ satisfies the regularity conditions, the MLE estimator $\widehat{d^n}$ is $\sqrt{T}$ consistent and converges to a limiting normal distribution. For the fractional white noise process, $(1-L)^d x_t = \varepsilon_t$, $\varepsilon_t \sim N(0, \sigma^2)$, it turns out that

$$\sqrt{T}(\widehat{d^n} - d) \to N(0, \frac{6}{\pi^2}),$$

the asymptotic variance of the parameter estimates is independent of the value of d.

So, for the parameter $\widehat{d^{\mathrm{Si}}}$, which is the average of $\widehat{d^n}$ over N simulations, we can derive its asymptotic distribution easily.

2.3 Discussion

Based on the econometric characterization of simulated data, we now briefly discuss the efficiency of the estimators. When N is fixed and T goes to infinity, the estimator $\widehat{\theta}_{\mathbf{x}}^{\mathrm{Si}}$ is always the most efficient one within the class $\{\widehat{\theta}_{\mathbf{x}}^{n},\ n = 1, ..., N\}$. When T is fixed and N goes to infinity, it seems there's no general conclusion about the most efficient estimator among the set $\{\widehat{\theta}_{\mathbf{x}}^{t},\ t = 1, ..., T_\tau\}$. In some cases, it is very easy to find the most efficient one, but in some cases, such as mean and AR coefficients, it need to be analyzed case by case. Now we consider the example of mean, other kinds of parameters can be considered in the same way.

Example 1: The mean $E(x_t)$ (*continued*).

Under Assumption 2, we can find the estimator $\widehat{\mu}^{Ti}$ that has the smallest asymptotic variance within the set $\left\{\frac{1}{\tau}\sum_{t=1}^{\tau}\widehat{\mu}^{t},\ \tau=1,2,...,T\right\}$, recalling that

$$\widehat{\mu}^{Ti}=\frac{1}{\tau_0}\sum_{t=1}^{\tau_0}\widehat{\mu}^{t},$$

where τ_0 satisfies

$$\tau_0=\underset{\tau=1,\ldots,T}{Arg\min}\left\{\frac{1}{\tau}\gamma_0+2\sum_{i=1}^{\tau-1}\frac{\tau-i}{\tau^2}\gamma_i\right\}.$$

We note that

$$0\leq\frac{1}{\tau}\left(\gamma_0+2\sum_{i=1}^{\tau-1}\frac{\tau-i}{\tau}\gamma_i\right)\underset{\tau\to\infty}{\to}0,$$

so

$$0=\inf_{\tau=1,\ldots,\infty}\left\{\frac{1}{\tau}\gamma_0+2\sum_{i=1}^{\tau-1}\frac{\tau-i}{\tau^2}\gamma_i\right\}\leq\min_{\tau=1,\ldots,T-1}\left\{\frac{1}{\tau}\gamma_0+2\sum_{i=1}^{\tau-1}\frac{\tau-i}{\tau^2}\gamma_i\right\},$$

we also note that the elements in $\underset{\tau=1,\ldots,\infty}{Arg\inf}\left\{\frac{1}{\tau}\gamma_0+2\sum_{i=1}^{\tau-1}\frac{\tau-i}{\tau^2}\gamma_i\right\}$ can be finite or infinite. (For instance, when $\tau=2$, and $\gamma_1=-\gamma_0$, $\gamma_i\geq 0$ when $i\geq 2$, then $2\in\underset{\tau=1,\ldots,\infty}{Arg\inf}\left\{\frac{1}{\tau}\gamma_0+2\sum_{i=1}^{\tau-1}\frac{\tau-i}{\tau^2}\gamma_i\right\}$)

We define $M=NT$ as the total number of periods among N simulations, when M goes to infinity, this implies that either N goes to infinity, T goes to infinity, or both go to infinity. In these three situations, we want to know which approximation is the best one. Notice that, for the estimator $\widehat{\mu}^{\mathrm{Si}}$, when M is given, the approximated variance of $\widehat{\mu}^{\mathrm{Si}}$ is the same for T is large or for both N and T are large, so we only need to consider the estimators $\widehat{\mu}^{\mathrm{Si}}$ and $\widehat{\mu}^{Ti}$. Let M be given, the variances of these estimators can be approximated by

$$\begin{aligned}Var\left\{\widehat{\mu}^{Ti}\right\}&\approx\frac{1}{M}\left\{\gamma_0+2\sum_{i=1}^{\tau_0-1}\frac{\tau_0-i}{\tau_0}\gamma_i\right\},\\ Var\left\{\widehat{\mu}^{\mathrm{Si}}\right\}&\approx\frac{1}{M}\left\{\gamma_0+2\sum_{i=1}^{\infty}\gamma_i\right\},\end{aligned}$$

when N, and T go to infinity, respectively, at the same rate.

Then we see that if

$$\sum_{i=1}^{\tau_0-1}\frac{i}{\tau_0}\gamma_i+\sum_{i=\tau_0}^{\infty}\gamma_i\geq(\leq)0,$$

then $\widehat{\mu}^{Ti}$ ($\widehat{\mu}^{\mathrm{Si}}$) has the smallest variance. As we can see, if $\gamma_i\geq(\leq)0$, for $i=1,2,...,$ then this condition is easily satisfied. As another example, let's consider the MA(1) process

$$x_t=\mu+\varepsilon_t+\alpha\varepsilon_{t-1},\quad\varepsilon_t\sim iid(0,1)$$

then

$$\begin{aligned}\gamma_0 &= 1+\alpha^2,\\ \gamma_1 &= \alpha,\\ \gamma_2 &= \gamma_3 = ... = 0\end{aligned}$$

and, therefore,

$$\gamma_0 + 2\sum_{i=1}^{\infty}\gamma_i = (1+\alpha)^2,$$

$$\frac{1}{\tau}\gamma_0 + 2\sum_{i=1}^{\tau-1}\frac{\tau-i}{\tau^2}\gamma_i = \frac{(1+\alpha)^2}{\tau} - \frac{2\alpha}{\tau^2},$$

It is easy to see that $\gamma_0/\tau + 2\sum_{i=1}^{\tau-1}(\tau-i)\gamma_i/\tau^2$ is a decreasing function with respect to τ. So when $M = NT$ is fixed, we have an approximation

$$\begin{aligned}Var\left\{\widehat{\mu}^{Ti}\right\} &\approx \frac{(1+\alpha)^2}{M} - \frac{2\alpha}{MT},\\ Var\left\{\widehat{\mu}^{\mathrm{Si}}\right\} &\approx \frac{(1+\alpha)^2}{M}.\end{aligned}$$

It is clear that $\widehat{\mu}^{Ti}$ has a smaller (larger) variance than $\widehat{\mu}^{\mathrm{Si}}$ when $\alpha \geq 0$ ($\alpha \leq 0$).

While the general conclusion from the theoretical consideration of whether we should choose T or N be large in the simulations is ambiguous, we know that when both T and N tend to infinity, different estimators established above resulted in the same limit distribution. From practical point of view, what we always do is running a MS model for many time periods for each realization and then run many independent realizations, this means that we let T go to infinity at first, and then let N go to infinity. One realization of a model for many periods is necessary, this is because this represents a scenario of the economy that the model described, it contains the information about the dynamic evolution of the MS model. The realization of different scenarios provide more information on the understanding of the underlying economy. Technically, the parameter $\theta_{\mathbf{x}}$ can be estimated easily from the traditional time series context for nth realization, then we simply average these independent estimators, this provides a consistent estimator for the parameter θ, and it is easy to do statistical inference because we only use the central limit theorem for the i. i. d. situation.

3 Comparing different MS economies

As we mentioned before, in order to compare two different MS models, we need to specify the outcomes of interest of MS models at first, and then decide which characteristics of these outcomes will be compared. Statistically, the characteristics of outcomes of interest can be described by some parameters, so the problem that we are interested in becomes testing the equality of the parameters of two MS models. For the purpose of comparing two different MS economies, besides Assumption 1, we need to specify which T periods of the outcomes should be compared. Normally we need to filter out some initial transition periods, the dynamics that are generated after these periods are representative to a MS model.

Let $\{x_t\}$ and $\{y_t\}$ be the outcome series of two MS models, the parameters that we are interested in are $\theta_{\mathbf{x}}$ and $\theta_{\mathbf{y}}$ of two MS economies, where $\mathbf{x} = (x_1, ..., x_\tau)$, $\mathbf{y} = (y_1, ..., y_\tau)$ for $1 \leq \tau \leq T$, or $\mathbf{x} = (x_1, x_2...)$, and $\mathbf{y} = (y_1, y_2...)$. The null hypothesis and alternative hypothesis are $H_0 : \theta_{\mathbf{x}} = \theta_{\mathbf{y}}$, $H_1 : \theta_{\mathbf{x}} \neq \theta_{\mathbf{y}}$.

Various parameters can be compared by applying the results of the previous section to compare two different MS economies, with the purpose of demonstration, we start with a comparison of simple descriptive statistics, such as the mean, median and variance etc.. Then we compare the autocorrelation patterns of the outcomes series.

We start with analyzing the comparison on the basis of a finite dimensional parameter. After we decide which outcomes of MS models, such as stock returns, will be studied, we can compare the equality of k characteristics of two MS economies summarized in $\theta_{\mathbf{x}}$, $\theta_{\mathbf{y}} \in R^k$.

We assume that $\theta_{\mathbf{x}} = \varphi(F_{\mathbf{x}})$. For illustrate purpose, let us consider the estimators $\widehat{\theta}_{\mathbf{x}}^{Ti}$ and $\widehat{\theta}_{\mathbf{y}}^{Ti}$, We know from (2) that

$$\begin{aligned}\sqrt{NT}(\widehat{\theta}_{\mathbf{x}}^{Ti} - \theta_{\mathbf{x}}) &= \sqrt{NT}\frac{1}{NT}\sum_{n,t}(\zeta(\mathbf{x}_{n,t}) - E\zeta(\mathbf{x}_{n,t})) + o_p(1), \\ \sqrt{NT}(\widehat{\theta}_{\mathbf{y}}^{Ti} - \theta_{\mathbf{y}}) &= \sqrt{NT}\frac{1}{NT}\sum_{n,t}(\zeta(\mathbf{y}_{n,t}) - E\zeta(\mathbf{x}_{n,t})) + o_p(1).\end{aligned}$$

The null hypothesis is that the $k-$dimensional vector of characteristics of the MS economy is the same for both MS models: $H_0 : \theta_{\mathbf{x}} = \theta_{\mathbf{y}}$. This equality can be tested by using the well known Wald test

$$W = NT(\widehat{\theta}_{\mathbf{x}}^{Ti} - \widehat{\theta}_{\mathbf{y}}^{Ti})'\widehat{\Sigma}^{-1}(\widehat{\theta}_{\mathbf{x}}^{Ti} - \widehat{\theta}_{\mathbf{y}}^{Ti}),$$

where $\widehat{\Sigma}$ is the sample analogue of the covariance matrix, $\Sigma = Var(\zeta(\mathbf{x}_{n,t}) - \zeta(\mathbf{y}_{n,t}))$, which is the covariance of $\zeta(\mathbf{x}_{n,t}) - \zeta(\mathbf{y}_{n,t})$.

For example, we are interested in studying the dynamics of the stock returns in an MS economy. For each simulation run of a MS model, the outcome is a time series of returns. We run each MS model independently N times and we want to test whether the average return of the two different economies are the same or not at certain time points: $0 \leq T_1 \leq ... \leq T_l \leq T$. More precisely, let $\{x_{n,t}\}_{t=1}^T$ and $\{y_{n,t}\}_{t=1}^T$ be the time series of stock returns of two different MS models, let $\widehat{\mu}_x^t$, $\widehat{\mu}_y^t$ denote the mean of the returns at time t over N independent simulation runs for each of the models respectively. So

$$\widehat{\theta}_x = \begin{pmatrix} \widehat{\mu}_x^{T_1} \\ ... \\ \widehat{\mu}_x^{T_l} \end{pmatrix}, \ \widehat{\theta}_y = \begin{pmatrix} \widehat{\mu}_y^{T_1} \\ ... \\ \widehat{\mu}_y^{T_l} \end{pmatrix},$$

under the null hypothesis that the averaged return are equal at time points, $t = T_1, ..., T_l$, we have

$$\sqrt{N}(\widehat{\theta}_x - \widehat{\theta}_y) \xrightarrow{d} N(0, \Sigma),$$

where Σ is the covariance matrix of the differences of the averaged returns which can be estimated by its sample analogue $\widehat{\Sigma}$.

We notice that in some cases we might need to design and implement two simulation models independently, such that the numbers of simulations N_1 and N_2 of the two MS models be different. This might be the case when one MS model need much more computational time than the other one. In this case, without loss of generality, we assume that $N_1 \geq N_2$, and

$$\sqrt{N_1}(\widehat{\theta}_x - \theta) \xrightarrow{d} N(0, \Sigma_1),$$
$$\sqrt{N_2}(\widehat{\theta}_y - \theta) \xrightarrow{d} N(0, \Sigma_2),$$

moreover, we assume that $\lim_{N_2 \to \infty} N_1/N_2 = c \geq 1$, then we get

$$\sqrt{N_2}(\widehat{\theta}_x - \widehat{\theta}_y) \xrightarrow{d} N(0, \Sigma_1/c + \Sigma_2).$$

For one MS model, we can also use Wald test statistics to detect the variation of the averaged returns at different time points, for instance, we can test the null hypothesis: $\theta_x^{T_1} = \theta_x^{T_2} = ... = \theta_x^{T_l}$.

Another example of descriptive statistic that is of interest is the median, which describe the central tendency of a distribution. After the estimation of the variance matrix Σ, the Wald type test can also be implemented. Also for one particular MS model, after we estimate the mean and median, as a character of the MS model itself, the skewness of the distribution of the character can be tested by testing the equality of them, this can be done based on Hausman test.

4 Comparing MS model with real data

After comparison of two different economies, we focus on the comparison between model generated data and real life data. For models of financial markets, one might, for example, compare the return process of a MS economy with the returns on the S&P 500. We are interested in the problem that whether a MS can generate dynamics that mimic the real market dynamics. Here we illustrate how to test whether a MS model can provide a good description on particular aspects of the actual data.

We denote the parameter of the real world by θ_R, and we denote its counterpart that comes from MS models as θ_{MS}. We want to test $H_0 : \theta_R = \theta_{MS}$. Because the estimator of θ_{MS} can have a convergence rate of $\sqrt{NT}$, we can treat the estimator of θ_{MS} as if it is the true value, the Wald statistics

$$W = T(\widehat{\theta}_R - \widehat{\theta}_{MS})'\widehat{\Sigma}^{-1}(\widehat{\theta}_R - \widehat{\theta}_{MS}), \tag{7}$$

can be constructed under the null hypothesis, where $\widehat{\Sigma}$ is the sample analogue of the covariance matrix related to the influence function of $\widehat{\theta}_R$, which can be obtained in a similar way as in the previous sections.

Alternatively, for the parameters, such as the autocorrelation coefficients to describe the dynamics in the MS model. First, similar to the situation of comparing two different MS economies, for the actual data, we estimate the mean of autocorrelation coefficients and then construct confidence intervals, then we can see whether the autocorrelation coefficients coming from simulated data lie in these intervals. Of course, for the MS model generated data, according to the limit distribution of the estimators of $\theta_{\mathbf{x}}$, the confidence interval for $\theta_{\mathbf{x}}$ can also be constructed, and then we can compare the confidence intervals that come from the actual data and simulated data.

Also, due to the asymptotic normality of the estimators of actual data and simulated data, a Wald test statistics can be constructed.

5 An application

In this section, we illustrate how the proposed econometric tools can be used to analyze the model by Levy et al. (2000) (LLS model from now on). First, we introduce the initial conditions parameters, then we turn to compare the LLS model with an adapted LLS model where a new type of investors is introduced, and finally, we confront the LLS model with real life data.

5.1 The MS models

We use the LLS model as an illustration. In Appendix B we describe this model in details. Now we turn to introduce the benchmark economy that we will simulate and analyze.

In the simulations time periods represent quarters of a year. The other parameter settings and initial conditions are as follows.

- Number of investors = 1000, with 96% RII investors and 4% EMB investors. There are two types of EMB investors, with memory span 5 and 15, respectively. Both groups are equally large.
- Number of shares $N = 10000$.
- Quarterly riskless interest rate $r = 1\%$.
- At time $t = 1$ each investor is endowed with a total wealth of \$1000, which is composed of 10 shares worth an initial price of \$20.94 per share, and the remainder in cash.
- Required quarterly rate of return on the stock $k = 4\%$.
- The initial dividend is set at \$0.5.
- Maximal one-period dividend decrease $z_1 = -7\%$.

- Maximal one-period dividend growth $z_2 = 10\%$.
- $\tilde{z}$ is uniformly distributed between z_1 and z_2; thus, the average dividend growth rate is $g = (z_1 + z_2)/2 = 1.5\%$.
- The standard deviation of the random noise affecting the EMB's decision making is $\sigma = 0.2$.

Our initial conditions and parameter setting are comparable to those made by LLS. The quarterly interest rate is taken to be 1%, yielding a 4.06% annual interest rate. The initial price is set as the first period price that the RII investors expect; the initial quarterly dividend is set at \$0.5 which corresponds to an annual dividend yield of about 4%. The average quarterly dividend growth rate of 1.5% represents the firm's growth and yields an annual growth rate of 6.1%, which is close to the long run average dividend growth rate of the S&P. The risk aversion parameter is chosen as $\alpha = 1.5$ because this value conforms with the estimate of the risk aversion parameter found empirically and experimentally, as described in Levy et al.(2000).

To understand the dynamics of the LLS economy, we first consider its price dynamics in Figure **??**. For one simulation of the benchmark model, we see that during the first 100 periods, the RII investors dominate the market, and the price has a clear upward trend, due to increasing dividends; then, around period 100, a relatively high dividend realizes and, as a consequence, a relatively high return is generated. This high return leads the EMB investors to increase their investment proportion in the stock at the next trading period. This increased demand of the EMB investors is large enough to affect the next period's price, and, thus, a second high return is generated. At this point in time, the EMB investors look at a set of ex post returns with two high returns, and they increase their investment proportion even further. Because the EMB investors keep buying aggressively, this positive feedback loop cannot be broken by the RII investors, even though they realize that the stock price is overvalued relative to its fundamental price, so they start selling stocks. However, when the price keeps going up, the EMB investors invest all their wealth in the stock. The price will stay at a high level, but the returns will become lower. Notice that when the EMB investors who look back for 5 periods have already "forgotten" the high return, the EMB investors with memory span of 15 periods are still investing aggressively in the stock. When they "forget" the high return, they cut their investments in the stock sharply and this causes a crash. Once the stock price goes back to its fundamental value, the RII investors start to buy again and the crash ceases. After a few periods, the cycles transit to shorter cycles induced by the EMB investors with the short memory span of 5 periods. The reason is that when a population becomes dominant and dictates the price dynamics, this population typically starts underperforming. This can be seen as follows. For the EMB investors, because the investors affect the price with their actions, they push the price up when buying, and, therefore, buy high. Similarly, they push the price down when selling, so they sell low.

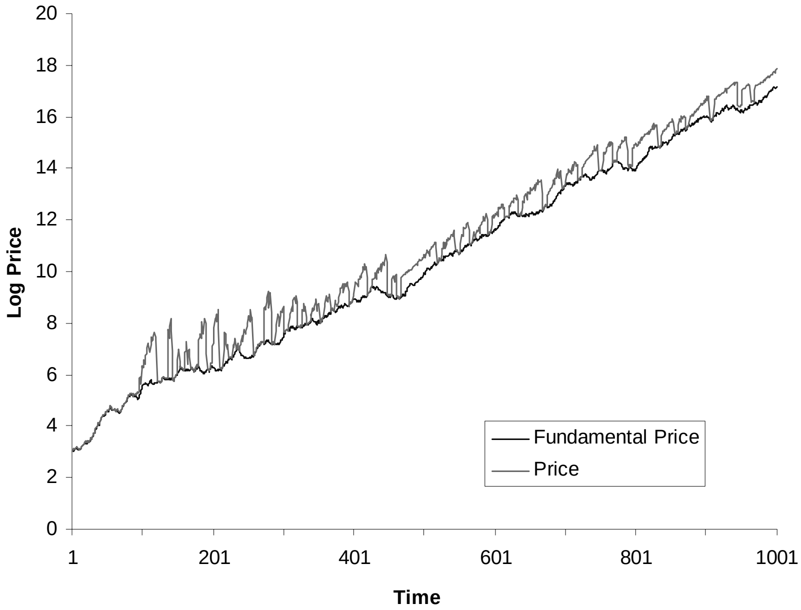

Figure 1: Log of the market price and the fundamental price

The above analysis is based on a single simulation. It makes sense to repeat the simulations many times, in order to investigate whether the findings based on a single simulation are robust to different drawings of random numbers, keeping the parameters and initial conditions the same. In this way we also may get an impression of the average behavior of the LLS-economy.

5.2 Comparing MS models

In the benchmark model, there are two types of investors, the RII and the EMB investors. The simulated stock market is a rising market, due to the assumed dividend growth, with price cycles caused by the trading strategies of the EMB investors with different memory spans. Because the two subgroups of EMB investors buy stocks at relatively high prices, and sell at low prices, in the end, they achieve a poor performance. So, it might be interesting to investigate what will happen when we introduce a new type of investors, who are, so to speak, at the opposite side of the market. Similar to Zschischang and Lux (2001), we consider as deviation from the benchmark model an economy with a new type of investors, constant portfolio investors, who always invest a constant proportion of their wealth in the stocks. Zschischang and Lux (2001) investigate the LLS model where initially all the investors are EMB investors (consisting of three or more subgroups). The authors found, when the market is invaded by only a small amount of constant portfolio investors (1%), that, even when these new investors are endowed with a small initial wealth and hold 1.5% of their portfolio in the stock, they eventually achieve dominance and asymptotically gain 100% of the available wealth. As an alternative economy, we consider an economy where 0.5% of the investors are constant portfolio investors instead of RII-investors (having the same initial wealth as the other investors). These constant portfolio investors invest 1.5% of their wealth in the stock. We keep the other characteristics of the economy the same as the benchmark model.

We performed a Wald test to investigate whether the introduction of the constant portfolio investors has a significant impact on the economy. The comparison results with the benchmark model in terms log return, log price and proportion of total wealth held by two groups EMB investors are summarized in table 3. It is clear that none of the comparison statistics is significant, thus the constant portfolio investors do not cause a significant impact on the economy.

Figure 2 shows the average proportion of total wealth of the constant portfolio investors across 5000 simulations. As the figure shows, we find that the wealth of the constant portfolio investors decreases gradually.

Notice that in the Zschischang and Lux-analysis the constant portfolio investors are the only investors who are at the opposite side of the market in case of the cycles, so that eventually they are able to gain all wealth. But in the economy considered here, the RII investors for a large part take over this role by buying or selling, depending on the price being lower or higher than its fundamental value, resulting in a gradually decreasing wealth held by the constant portfolio investors.

In this table, we also report the results of sensitivity analysis, the resulting

Table 3: The comparison results with the benchmark model in terms of the mean

	Log Return	Log Price	Wealth (ms=15)	Wealth (ms=15)
	101.88 (99)	42085.5(99)	122.20 (99)	125.10 (99)
t=901,902,...,1000	101.68 (99)	43413.2(99)	124.84 (99)	113.73 (99)
	77.95 (100)	76.95 (100)	97.29 (100)	79.13 (100)
	9.53 (5)	192436.7(5)	383.93 (5)	772.40 (5)
t=500,600,...,1000	2.65 (5)	194884.1(5)	388.87 (5)	757.85 (5)
	8.33 (6)	7.44 (6)	3.37 (6)	2.45 (6)

Note: t = 901, 902, ..., 1000 means that the periods under consideration are the last 100 periods (of the 1000 periods), t = 700, 800, 900, 1000 means that only these four time points are considered. Within the row named 'Log Return', the first subrow reports the Wald statistics of the benchmark economy, for instance, 101.88 is the Wald statistic corresponding to the null hypothesis of equality of the average log return in periods t = 901, 902, ..., 1000 (with degrees of freedom between brackets), and so on, the second subrow reports the Wald statistics of the new economy, and the third subrow reports the results of comparing the new economy with the benchmark model.

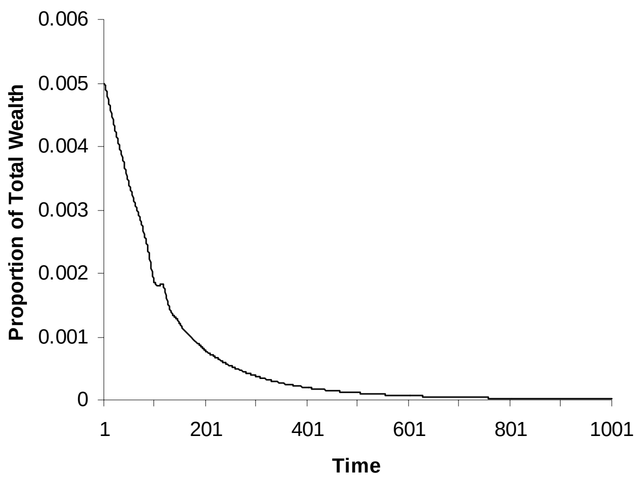

Figure 2: Proportion of total wealth held by constant portfolio investors, averaged over 5000 simulations

Wald test statistics show that both the benchmark model and the new model are robust in terms of log-returns, and the wealth hold by EMB-investors, and for loh-prices, the results are expected because of the obvious increase trend.

5.3 Comparing MS model with real life data

We use quarterly data of S & P 500 from Datastream as representation of the real life situation, which starts from 1965 to the first quarter of 2003. See table 4 for the sample statistics. We illustrate by comparing AR coefficients and the coefficients of the ARFIMA(p, d, q) process.

First, for the actual data, we calculate the autocorrelations, construct the confidence interval by using the Newey-West corrected standard deviation for each of the autocorrelations, and we report the results in table 5.

Table 4: Sample statistics of reurns of S & P 500

Mean	Median	Std.Dev.	Max	Min	Skew.	Kurt.
0.0194	0.0162	0.0848	0.2923	-0.2548	-0.0575	3.800

Then we estimate the averaged autocorrelations that comes from LLS models. We can then check if the averaged autocorrelations lie in the intervals of that of actual data. We calculate the autocorrelation for different lags, $j = 1, 2, ..., 60$. Table 6 summarizes the average autocorrelation, the average Newey-West corrected standard deviation, and average t-value for 5000 independent simulations. Also the number of significant positive and negative t-values, out of the total 5000 t-values calculated, are presented.

It is obvious that at periods 6, 7, 8 and 16, the autocorrelations are significantly negative. We also report the results in Figure 3,

We find that the means of the autocorrelation lie entirely in the 95% confidence intervals of that of real life data. It seems that the LLS model fits the real world very well in terms AR coefficients.

Next, we illustrate comparing methods in terms of the coefficients of the ARFIMA(p, d, q) process. We estimate the ARFIMA $(0, d, 0)$ model and the ARFIMA $(1, d, 1)$ model for stock returns. We summarize the results in the table 7

We see from the table that in both of the ARFIMA $(0, d, 0)$ and ARFIMA $(1, d, 1)$ model, the parameter d is not significant, there is no evidence of long memory for the quarterly stock return process. We estimate the ARFIMA $(0, d, 0)$ model and the ARFIMA $(1, d, 1)$ model for the returns of the benchmark LLS model, we run 5000 independent simulations, and for each realization we estimate the two models. The table 8 summarized the averaged results

We know from the table that for the ARFIMA $(0, d, 0)$ model, the estimated d from the LLS model lies within the 95% confidence interval of estimates of d from actual data, which is $(-0.1506, 0.114)$. However, for the ARFIMA $(1, d, 1)$ model, there's significant difference between the actual data and the data of

Table 5: Autocorrelation pattern of stock returns of S & P 500

Lag	AR coe.	Newey-West Std.	t-statistics	Lower bound	Upper bound*
1	0.0117	0.0702	0.1667	-0,1259	0,1493
2	-0.0455	0.0585	-0.7779	-0.1602	0.0692
3	0.0359	0.1017	0.3533	-0.1634	0.2352
4	-0.0578	0.0801	-0.7210	-0.2148	0.0992
5	-0.0021	0.0713	-0.0299	-0.1418	0.1376
6	-0.0056	0.1028	-0.0544	-0.2071	0.1959
7	-0.0954	0.0613	-1.5571	-0.2155	0.0247
8	-0.0132	0.0657	-0.2004	-0.1420	0.1156
9	0.0816	0.0863	0.9456	-0.0875	0.2507
10	0.0562	0.0754	0.7457	-0.0916	0.2040
11	-0.0787	0.0774	-1.0176	-0.2304	0.0730
12	0.0425	0.0730	0.5829	-0.1006	0.1856
13	-0.0671	0.0935	-0.7174	-0.2504	0.1162
14	-0.0176	0.0863	-0.2057	-0.1867	0.1515
15	0.0215	0.0883	0.2441	-0.1516	0.1946
16	0.0549	0.0981	0.5598	-0.1374	0.2472
17	0.2033	0.1345	1.5118	-0.0603	0.4669
18	-0.0081	0.0755	-0.1072	-0.1561	0.1399
19	-0.0738	0.0767	-0.9619	-0.2241	0.0765
20	-0.1346	0.0885	-1.5207	-0.3081	0.0389
21	-0.0827	0.0973	-0.8498	-0.2734	0.1080
22	0.0212	0.0948	0.2234	-0.1646	0.2070
23	-0.0836	0.0706	-1.1833	-0.2220	0.0548
24	-0.1015	0.0647	-1.5687	-0.2283	0.0253
25	-0.0681	0.0919	-0.7404	-0.2482	0.1120
26	0.1436	0.0990	1.4500	-0.0504	0.3376
27	0.0736	0.0901	0.8166	-0.1030	0.2502
28	-0.0545	0.0918	-0.5938	-0.2344	0.1254
29	-0.1372	0.0873	-1.5772	-0.3083	0.0339
30	-0.0506	0.1094	-0.4631	-0.2650	0.1638

* Significant at 95%

Table 6: Autocorrelation pattern of stock returns of LLS model

Lag	AR coe.	Average NW Std.	Average t-value	Significant Posit. t*	Significant Negat. t*	Confidence intervals	
1	0.0026	0.0358	0.1015	768	653	0.0011	0.0041
2	0.0170	0.0300	0.5595	826	138	0.0161	0.0179
3	0.0037	0.0281	0.0806	321	273	0.0029	0.0046
4	-0.0029	0.0295	-0.1629	267	493	-0.0038	-0.0020
5	0.0472	0.0297	1.5924	2468	76	0.0460	0.0484
6	-0.2041	0.0454	-4.7671	0	4964	-0.2053	-0.2029
7	-0.1253	0.0382	-3.4071	1	4552	-0.1266	−0.1240
8	-0.0719	0.0332	-2.2163	2	3545	-0.0729	-0.0709
9	-0.0372	0.0317	-1.1988	22	1724	-0.0383	-0.0362
10	0.0028	0.0324	0.0385	395	361	0.0017	0.0039
11	-0.0407	0.0313	-1.4209	116	2361	-0.0420	-0.0394
12	0.0209	0.0372	0.4383	817	249	0.0195	0.0224
13	0.0341	0.0382	0.7702	1107	109	0.0328	0.0355
14	0.0347	0.0374	0.8231	1151	82	0.0334	0.0360
15	0.0417	0.0384	1.0386	1451	549	0.0404	0.0429
16	-0.1180	0.0486	-2.5671	0	4057	-0.1191	-0.1168
17	-0.0364	0.0380	-0.9819	140	1560	-0.0379	-0.0350
18	-0.0099	0.0353	-0.3352	401	916	-0.0113	-0.0085
19	-0.0002	0.0350	-0.0811	442	607	-0.0015	0.0011
20	0.0038	0.0352	0.0429	439	412	0.0026	0.0050
21	-0.0183	0.0352	-0.5811	120	907	-0.0194	-0.0172
22	0.0351	0.0376	0.8825	1206	61	0.0339	0.0362
23	0.0394	0.0367	1.0243	1428	43	0.0382	0.0406
24	0.0401	0.0380	1.0048	1403	43	0.0389	0.0414
25	0.0373	0.0383	0.9039	1278	75	0.0360	0.0386
26	0.0235	0.0377	0.5437	778	142	0.0223	0.0248
27	0.0279	0.0379	0.6686	905	110	0.0268	0.0292
28	0.0131	0.0373	0.2647	500	246	0.0119	0.0144
29	0.0023	0.0371	-0.0269	310	420	0.0011	0.0035
30	-0.0053	0.0369	-0.2289	212	556	-0.0065	-0.0041

Table 7: Maximum likelihood estimation of ARFIMA(p,d,q) model for S & P 500

ARFIMA(p, d, q)	Coefficient	Std. Error	t-value	P-value
$(0, d, 0)$	-0.0183	0.0675	-0.272	0.786
	-0.0527	0.0813	-0.648	0.518
$(1, d, 1)$	-0.5668	0.4687	-1.21	0.229
	0.6365	0.4240	1.5	0.135

Note: The estimated coefficients of ARFIMA$(1, d, 1)$ model are listed in the order: d, AR coefficient, MA coefficient. This is also true for other tables in this section.

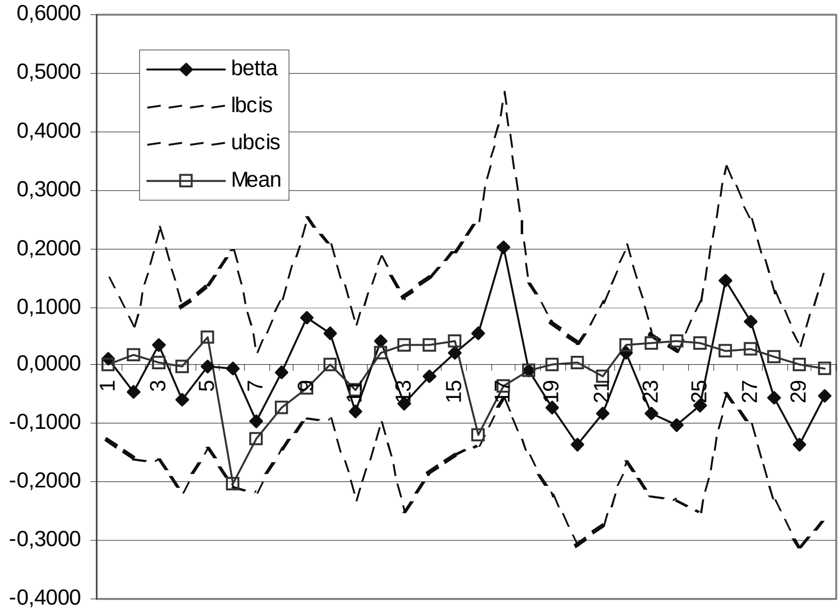

Figure 3: The confidence intervals for the autocorrelations of S&P 500 and the averaged autocorrelation of LLS model. (Note: In this table, betta means the estimated autocorrelation of S&P 500, lbcis and ubcis means its lower and upper confidence intervals, and "Mean" is the averaged autocorrelation of LLS model over 5000 simulations.)

Table 8: Maximum likelihood estimation of ARFIMA(p,d,q) model for LLS model

ARFIMA(p,d,q)	Coefficient	t-value	P-value	No. Sig.
$(0,d,0)$	-0.0272	-0.3699	0.5210	379
	-0.7444	-3.3721	0.0339	4409
$(1,d,1)$	0.6983	5.7899	0.0151	4672
	0.0208	-0.3973	0.5342	250

LLS model, for the simulated data, the estimated parameter d is significantly negative, which means that the return process of LLS model has short memory.

Besides the benchmark LLS model, we also run the LLS model for different situations, for instance, for different initial price, initial dividend, initial wealth, different risk aversion parameter, etc. We report the estimation results of the ARFIMA $(0, d, 0)$ model in table 9 and in table 10 we also report the t-test for the difference of estimated d between the benchmark model and the models with different initial parameters.

Table 9: Maximum likelihood estimation of ARFIMA(0,d,0) model for LLS models

		d	t-value	P-value	No. Sig.
p(0)	16	-0.0286	-0.3868	0.5201	386
	26	-0.0282	-0.3844	0.5247	381
D(0)	0.4	-0.0272	-0.3666	0.5228	370
	0.6	-0.0247	-0.3376	0.5193	333
RV	1.45	-0.0281	-0.3840	0.5172	393
	1.55	-0.0273	-0.3688	0.5233	372
z_1	-0.08	0.0077	0.0847	0.6621	33
	-0.06	-0.1806	-2.6182	0.0848	3509
IW	$unif.$	-0.0239	-0.3267	0.5241	346
	50%	-0.0279	-0.3787	0.5230	386

Table 10: The t-test of ARFIMA(0,d,0) model for LLS models

	p(0)		D(0)		RV		Z_1		IW	
	16	26	0.4	0.6	1.45	1.55	-0.08	-0.06	Unif.	505
t	1.167	0.840	0.047	1.935	0.760	0.125	26.696	105.53	2.546	0.554

We see from the table 10 that the LLS model is robust with respect to the initial prices, initial dividend, risk aversion parameter in terms of d. However, the changes of maximal one-period dividend decrease Z_1 has a big impact, this is because the change of Z_1 also changes the whole distribution of the dividend process. We notice that in the benchmark model each investor is endowed equally with a total wealth of \$1000. If half of the investors is endowed \$500 and the other half endowed with \$1500, then, compared to the benchmark model, the difference in d is not significant. However, the difference in d is significant when all of the investors initial wealth is drawn from a uniform distribution on $[500, 1500]$.

We report the estimation results of the ARFIMA $(1, d, 1)$ model in table 11, and we also report the results of Wald test for the difference in the estimated parameters between the benchmark model and the models with different initial parameters in the table 12.

Table 11: Maximum likelihood estimation of ARFIMA(1,d,1) model for LLS models

	Coefficient	t-value	P-value	No. Sig.
	−0.7417	−3.3707	0.0341	4368
p(0)=16	0.6849	5.6697	0.0171	4630
	0.0298	−0.2950	0.5350	248
	−0.7400	−3.3690	0.0361	4377
p(0)=26	0.6848	5.6989	0.0182	4619
	0.0282	−0.2504	0.5383	250
	−0.7488	−3.3982	0.0314	4416
D(0)=0.4	0.6941	5.7334	0.0155	4654
	0.0292	−0.1310	0.5363	232
	−0.7414	−3.3582	0.0361	4386
D(0)=0.6	0.6887	5.7031	0.0163	4653
	0.0299	−0.2926	0.5343	255
	−0.7452	−3.3935	0.0325	4406
RV=1.45	0.6846	5.6860	0.0196	4620
	0.0336	−0.1651	0.5387	231
	−0.7520	−3.4161	0.0309	4419
RV=1.55	0.6936	5.7428	0.0173	4645
	0.0324	−0.1240	0.5405	218
	−0.7963	−3.8913	0.0092	4683
z_1=-0.08	0.7225	6.2001	0.0029	4806
	0.0785	0.0568	0.4804	223
	−0.4247	−1.7905	0.2511	2158
z_1=-0.06	0.2754	2.4540	0.2010	2774
	−0.0537	−0.8731	0.4030	764
	−0.7483	−3.4077	0.0326	4429
IW=unif.	0.6940	5.8031	0.0158	4674
	0.0321	−0.2320	0.5328	247
	−0.7377	−3.3562	0.0357	4372
IW=50%	0.6837	5.7035	0.0197	4615
	0.0277	−0.3603	0.5420	242

Table 12: The Wald test of ARFIMA(1,d,1) model for LLS models

	p(0)		D(0)		RV		Z_1		IW	
	16	26	0.4	0.6	1.45	1.55	-0.08	-0.06	Unif.	505
w	8.695	7.757	4.899	7.538	13.68	11.66	573.9	10681.5	9.419	7.513

We see from table 12 that in terms of the ARFIMA$(1, d, 1)$ model, the robustness of the LLS model with respect to the initial conditions and parameters is ambiguous. Compared with the 95% quantile of the χ^2_3 distribution, 7.815, the most significant impact is still caused by Z_1.

6 Conclusion

Microscopic Simulation (MS-)models are a promising way to study financial markets, since they allow for the possibility to include all kinds of realistic and complex behavior of interacting economic agents, without having to worry about analytical tractability. However, in many cases judgements of the outcomes of MS models seems to be based solely on visual inference.

In this paper we propose to investigate Microscopic Simulation (MS) models using statistical and econometric techniques. Such techniques can be used to study the impact of changes in the initial parameter settings and initial conditions on the simulated time series behavior of the relevant quantities. But also different MS economies can be compared using these techniques, in order to find out whether particular adaptations are crucial or not. We also present the methodology to compare real life data with the MS economies. Here it is important to take into account measurement uncertainty in both the simulation data and the real life observation. Comparison of "simple" statistics, such as the mean or a single autocorrelation coefficient, is rather straightforward. However, we also show how one can compare "global" characteristics of an economy by testing for differences in the spectral density estimate.

A The asymptotic variance of the estimator of AR coefficients

Here we discuss the case that T is fixed and N goes to infinity.

If $T \geq j$, then β_j can be estimated consistently by

$$\widehat{\beta}_j^t = \frac{\widehat{\gamma}_{j,t}}{\widehat{\gamma}_{0,t}}$$

as N goes to infinity for each t, where $\widehat{\gamma}_{0,t}$, $\widehat{\gamma}_{j,t}$ are the sample analogue of $\gamma_{0,t} = var\{x_t\}$, $\gamma_{j,t} = cov(x_t, x_{t+j}\}$. For the estimator

$$\widehat{\beta}_j^{Ti} = \frac{1}{T-j} \sum_{t=1}^{T-j} \widehat{\beta}_{j,t},$$

we can prove that

$$\sqrt{N}(\widehat{\beta}_j^{Ti} - \beta_j) \to N\left(0, V_1\right),$$

where

$$V_1 = \lim_{N\to\infty} Var\left\{\frac{1}{T-j}\sum_{t=1}^{T-j}\widehat{\beta}_j^t\right\}.$$

We know that

$$\begin{pmatrix} \beta_j^1 \\ \beta_j^2 \\ \\ \beta_j^{T-j} \end{pmatrix} = \begin{pmatrix} \frac{cov(x_1,x_{j+!})}{Vx_1} \\ \frac{cov(x_2,x_{j+2})}{Vx_2} \\ \\ \frac{cov(x_{T-j},x_T)}{Vx_{T-j}} \end{pmatrix} = f\begin{pmatrix} Ex_1 \\ Ex_1^2 \\ Ex_{j+1} \\ Ex_1x_{j+1} \\ \dots \\ \dots \\ Ex_{T-j} \\ Ex_{T-j}^2 \\ Ex_T \\ Ex_{T-j}x_T \end{pmatrix},$$

where

$$f\begin{pmatrix} z_1 \\ \dots \\ z_{4(T-j)} \end{pmatrix} = H_{T-j,1},$$

$$h_{i,1} = \frac{z_{4i} - z_{1+4(i-1)}z_{3+4(i-1)}}{z_{2+4(i-1)} - z_{1+4(i-1)}^2},\ i = 1, ..., T-j$$

$$\left(\frac{\partial f}{\partial z'}\right)_{i,k} = \begin{cases} \frac{2z_{1+4(i-1)}(z_{4i}-z_{1+4(i-1)}z_{3+4(i-1)})}{(z_{2+4(i-1)}-z_{1+4(i-1)}^2)^2} - \frac{z_{3+4(i-1)}}{z_{2+4(i-1)}-z_{1+4(i-1)}^2},\ k = 1+4(i-1) \\ -\frac{z_{4i}-z_{1+4(i-1)}z_{3+4(i-1)}}{(z_{2+4(i-1)}-z_{1+4(i-1)}^2)^2},\ k = 2+4(i-1) \\ -\frac{z_{1+4(i-1)}}{z_{2+4(i-1)}-z_{1+4(i-1)}^2},\ k = 3+4(i-1) \\ \frac{1}{z_{2+4(i-1)}-z_{1+4(i-1)}^2},\ k = 4i \\ 0,\ others. \end{cases}$$

Thus

$$\sqrt{N}(\begin{pmatrix} \widehat{\beta}_j^1 \\ \widehat{\beta}_j^2 \\ \\ \widehat{\beta}_j^{T-j} \end{pmatrix} - \begin{pmatrix} \beta_j^1 \\ \beta_j^2 \\ \\ \beta_j^{T-j} \end{pmatrix}) \to N(0, V),$$

and

$$V = KVar\{X\}K',$$

where $X = (x_1, x_1^2, x_{j+1}, x_1x_{j+1}, ..., x_{T-j}, x_{T-j}^2, x_T, x_{T-j}x_T)'$, and $Var\{X\} = BX_{(T-j)\times(T-j)}$, BX is a block matrix, each block is 4×4, and

$$(K)_{i,k} = \begin{cases} \frac{2Ex_{1+4(i-1)}cov(x_{1+4(i-1)},x_{j+1+4(i-1)})}{(Vx_{1+4(i-1)})^2} - \frac{Ex_{j+1+4(i-1)}}{Vx_{1+4(i-1)}},\ k = 1+4(i-1) \\ -\frac{cov(x_{1+4(i-1)},x_{j+1+4(i-1)})}{(Vx_{1+4(i-1)})^2},\ k = 2+4(i-1) \\ -\frac{Ex_{1+4(i-1)}}{Vx_{1+4(i-1)}},\ k = 3+4(i-1) \\ \frac{1}{Vx_{1+4(i-1)}},\ k = 4i \\ 0,\ others \end{cases},$$

$i = 1, ..., T - j,\ k = 1, ..., 4(T - j)$. We define

$$K_i = \begin{pmatrix} K_{i,1+4(i-1)} & K_{i,2+4(i-1)} & K_{i,3+4(i-1)} & K_{i,4i} \end{pmatrix}\ i = 1, ..., T - j,$$

then

$$V_{ik} = K_i \left(BX\right)_{ik} K_k',\ i, k = 1, ..., T - j$$

therefore

$$V_1 = \frac{1}{(T-j)^2} \iota_{1\times(T-j)} V \iota_{(T-j)\times 1}.$$

Under the relationship that

$$x_t = \alpha_j + \beta_j x_{t-j} + \varepsilon_t,$$

we can prove that the asymptotic variance of the estimator $\widehat{\beta}_j^{Ti}$ is the same as that of $\widehat{\beta}_j^{\text{Si}}$ when both T and N go to infinity.

We notice that the limits of the average of the diagonal terms of matrix V corresponds to the first term in the expression of S, $E\{x_{t-j}^2 \varepsilon_t^2\}$, and the limits of the average of the other terms corresponds to $E\{x_{t-j} x_{t-j-i} \varepsilon_t \varepsilon_{t-i}\}$ in S. For simplicity, we assume that $E\{x_t\} = 0$, then

$$V_{ik} = \begin{pmatrix} 0 & -\frac{\gamma_j}{\gamma_0^2} & 0 & \frac{1}{\gamma_0} \end{pmatrix} \left(BX\right)_{ik} \begin{pmatrix} 0 \\ -\frac{\gamma_j}{\gamma_0^2} \\ 0 \\ \frac{1}{\gamma_0} \end{pmatrix},$$

so

$$\begin{aligned} V_{ii} &= -\frac{1}{\gamma_0^4} \left(\gamma_j^2 Cov(x_i, x_i^2) - 2\gamma_0 \gamma_j Cov(x_i^2, x_i x_{i+j}) + \gamma_0^2 Var(x_i x_{i+j})\right) \\ &= -\frac{1}{\gamma_0^2} Var\left(x_i x_{i+j} - \beta_j x_i^2\right) = -\frac{1}{\left(E\{x_t^2\}\right)^2} E\{x_{t-j}^2 \varepsilon_t^2\}, \quad i = 1, ..., T - j, \end{aligned}$$

and

$$\begin{aligned} V_{ik} &= -\frac{1}{\gamma_0^4} \left[\gamma_j^2 Cov(x_i^2, x_k^2) - \gamma_0 \gamma_j Cov(x_i x_{i+j}, x_k^2)\right. \\ &\quad \left. -\gamma_0 \gamma_j Cov(x_i^2, x_k x_{k+j}) + \gamma_0^2 Cov(x_i x_{i+j}, x_k x_{k+j})\right] \\ &= -\frac{1}{2}\frac{1}{\gamma_0^4} \left[-Var\left(x_i x_{i+j} - \beta_j x_i^2 - x_k x_{k+j} + \beta_j x_k^2\right)\right. \\ &\quad \left. +Var\left(x_i x_{i+j} - \beta_j x_i^2\right) + Var\left(x_k x_{k+j} - \beta_j x_k^2\right)\right] \\ &= -\frac{1}{2}\frac{1}{\gamma_0^4} \left[-Var\left(x_i \varepsilon_i - x_k \varepsilon_k\right) + Var\left(x_i \varepsilon_i\right) + Var\left(x_k \varepsilon_k\right)\right] \\ &= -\frac{1}{\left(E\{x_t^2\}\right)^2} E\{x_i \varepsilon_i x_k \varepsilon_k\}, \quad i, k = 1, ..., T - j, i \neq k. \end{aligned}$$

thus

$$(T-j)V_1 = \frac{1}{(T-j)}\iota_{1\times(T-j)} V \iota_{(T-j)\times 1} \to Q^{-1}SQ^{-1}$$

as $T \to \infty$, $N \to \infty$. This proves that

$$\sqrt{NT}(\widehat{\beta}_j^{Ti} - \beta_j) \to N\left(0, Q^{-1}SQ^{-1}\right).$$

B The Levy-Levy-Solomon Model

In the model by Levy et al. (2000), LLS economy from now on, there are two assets: a stock and a bond. The bond is assumed to be a riskless asset, while the stock is a risky asset. The stock serves as a proxy for the market portfolio, for example, the Standard & Poor's index. The bond is exogenous with infinite supply, so the investors can buy from it as much as they wish at a given rate of return, r. The stock is in finite supply. There are N outstanding shares of the stock. The return on the stock is composed of two parts:

(i) Capital gain. If an investor holds a stock, any rise (fall) in the price of the stock contributes to an increase (decrease) in the investor's wealth.

(ii) Dividend Payments. The company earns income and distributes dividends. It is assumed that the firm pays a dividend of D_t per share at time t. The dividend is a stochastic variable that follows a multiplicative random walk, that is, $\widetilde{D}_t = D_{t-1}(1+\widetilde{z})$, where $\widetilde{z}$ is a random variable[3] with some probability density function $f(z)$ with support $[z_1, z_2]$. For simplicity, $\widetilde{z}$ is distributed uniformly in the range $[z_1, z_2]$. The overall rate of gross return on the stock in period t, denoted by $\widetilde{R}_t$, is now given by

$$\widetilde{R}_t = \frac{\widetilde{P}_t + \widetilde{D}_t}{P_{t-1}} \tag{8}$$

where $\widetilde{P}_t$ is the stock price at time t.

The investors are expected utility maximizers, characterized by the utility index $U(W) = W^{1-\alpha}/1-\alpha$, which reflects their personal preference. The investors will be divided into two groups, the first group will be referred to as the rational informed investors (RII), and the second group will be referred to as the efficient market believers (EMB). RII investors evaluate the "fundamental value" of the stock as the discounted stream of all future dividends. They believe that the stock price may deviate from the fundamental value in the short run, but if it does, it will eventually converge to the fundamental value. The EMB investors believe in market efficiency. They believe that the stock price accurately reflects the stock's fundamental value at every point in time. Therefore, their investment decision is reduced to optimal diversification between the stock and the bond. This diversification decision requires the ex ante return distribution for the stocks, but as the ex ante distribution is not available, the EMB investors assume that the process generating the returns is fairly stable,

[3] We will use $\widetilde{}$ to denote a random variable to distinguish it from its realization.

and they employ the ex post distribution of stock returns in order to estimate the ex ante return distribution.

The RII investors

In the LLS model, it is assumed that the RII investor believes that the convergence of the price to the fundamental value will occur in the next period. Furthermore, RII investors estimate the next period fundamental value of stock price P^f_{t+1} by

$$P^f_{t+1} = \frac{E_{t+1}[\widetilde{D}_{t+2}]}{k-g} \tag{9}$$

according to Gordon's dividend stream model. Here k is the discount factor, and g is the expected growth rate of the dividend, i.e., $g = E(\widetilde{z}) = \int_{z_1}^{z_2} f(z)zdz$, which is known to the investors. The expectation at time $t+1$ of $\widetilde{D}_{t+2}$ depends on the realized dividend observed at $t+1$, D_{t+1}, but at time t, D_t is known, not D_{t+1}. However, the RII investors know the distribution of $\widetilde{D}_{t+1}$: $\widetilde{D}_{t+1} = D_t(1+\widetilde{z})$. Consequently, RII investors believe that P_{t+1} is a random variable given by

$$\widetilde{P}_{t+1} = \widetilde{P}^f_{t+1} = \frac{D_t(1+\widetilde{z})(1+g)}{k-g} \tag{10}$$

The RII investors' investment decision is based on the rate of return of the stocks, $\widetilde{R}_{t+1}$, that is implied by the price process above. For every hypothetical price, P_h, RII investor i believes that if she invests a proportion x^i_h of her wealth in the stock at time t, then at time $t+1$ her wealth will be

$$\widetilde{W}^i_{t+1} = W^i_h[(1-x^i_h)(1+r) + x^i_h\widetilde{R}_{t+1}] \tag{11}$$

where W^i_h is the wealth of investor i at time t. This wealth level depends on the hypothetical price of the stock, P_h, since

$$W^i_h = W^i_{t-1} + N^i_{t-1}D_{t-1} + (W^i_{t-1} - N^i_{t-1}P_{t-1})r + N^i_{t-1}(P_h - P_{t-1}). \tag{12}$$

Here N^i_{t-1} is the number of shares held by investor i at time $t-1$.

For every hypothetical price, P_h, the investor's decision is now to find the proportion of her wealth to invest in stocks, denoted by x^i_h, which maximizes her expected utility $E\left\{U(\widetilde{W}^i_{t+1})\right\}$:

$$\begin{aligned} E\left\{U(\widetilde{W}^i_{t+1})\right\} &= E\left\{U\left(W^i_h[(1-x^i_h)(1+r) + x^i_h\widetilde{R}_{t+1}]\right)\right\} \\ &= \int_{z_1}^{z_2} \frac{1}{1-\alpha}\left(W^i_h[(1-x^i_h)(1+r) + x^i_h\widetilde{R}_{t+1}]\right)^{1-\alpha} f(z)dz. \end{aligned}$$

A solution for this optimization problem can be found by solving the first order conditions, see Appendix 1.

With investor wealth and the optimal share of this wealth that is invested in stocks, the number of shares demanded by RII investor i is

$$N^i_h(P_h) = \frac{x^i_h(P_h)W^i_h(P_h)}{P_h} \tag{13}$$

The EMB investors

EMB investor i uses the most recent m^i returns on the stock to estimate the ex ante distribution. Although the investor might realize that all past returns matter, he has only a limited memory, so only the last m^i returns are taken into account. At time t, each of these past returns on the stock $R_{j,}$ $j = t, t-1, ..., t-m^i+1$ is given an equal probability $1/m^i$ to reoccur in the next period $(t+1)$. Therefore, the expected utility of EMB investor i is given by

$$E\left\{U(W_{t+1}^i)\right\} = \frac{1}{m^i}\sum_{j=1}^{m^i}\frac{1}{1-\alpha}[(1-x_h^i)W_h^i(1+r)+x_h^iW_h^i(1+R_{t-j})]^{1-\alpha}. \quad (14)$$

The EMB investor maximizes this expected utility yielding the optimal proportion of wealth, x_h^{*i}, that will be invested in the stock. This determines his demand for the stock.

However, many empirical studies suggest that the behavior of investors is driven not only by rational expected utility maximization but by a multitude of other factors as well. To model the effects of all these factors causing the investor to deviate from the optimal portfolio, a normally distributed random variable is added to the optimal investment proportion. To be more specific, LLS assume that

$$x^i = x^{*i} + \widetilde{\varepsilon}^i$$

where $\widetilde{\varepsilon}^i$ is a random variable drawn from a normal distribution with mean zero and standard deviation σ. For simplicity, noise is only added to the portfolio share of stocks for the EMB investors.

With the total supply of shares N fixed, the equilibrium stock price at time $t+1$, P_{t+1}, can be determined. It is the hypothetical price, P_h, that equates the aggregate demand for stocks of the RII and EMB investors with total supply. This price can be recorded, so as the other market characteristics of interest, such as the investor's wealth levels. The new price leads to updated expectations and a new equilibrium arises in the next period, and so on.

References

[1] Arthur, W. B., Durlauf, S. N., and Lane, D. A. (1997) eds, The Economy as an Evolving Complex System II, Addison-Wesley, Reading, MA.

[2] Baillie, R. T. (1996), Long Memory processes and Fractional Integration in Econometrics, Journal of Econometrics, 73, 5-59.

[3] Barberis, N., M. Huang and T. Santos (2001), Prospect Theory and Asset Prices, Quarterly Journal of Economics, 116, 1-54.

[4] Barnett, W. A. and A. Serletis (2000), Martingales, Nonlinearity, and Chaos, Journal of Economic Dynamics & Control, 24, 703-724.

[5] Billingsley, P. (1968), Convergence of Probability Measures, John Wiley & Sons, New York.

[6] Black, F. (1986), Noise, The Journal of Finance, 41, No. 3, 529-543.

[7] Brock, W. A. (1986), Distinguishing Random and Deterministic Systems: Abridged Version, Journal of Economic Theory, 40, 168-195.

[8] Campbell, J. Y., A. W. Lo, and A. C. Mackinlay (1997), The Econometrics of Financial Markets, Princeton University Press, Princeton, New Jersey.

[9] Chatfield C. (1975), The Analysis of Time Series: Theory and Practice, Chapman and Hall, London.

[10] Chen, S., T. Lux and M. Marchesi (2001), Testing for Non-linesr Structure in an Artificial financial Market, Journal of Economic Behavior & Organization, 46, 327-342.

[11] Chiarella, C. and X. He (2002), An Adaptive Model on Asset Pricing and Wealth Dynamics with Heterogeneous Trading Strategies, Working Paper, University of Technology, Sydney.

[12] Cochrane, J.H. (2001), Asset Pricing, Princeton University Press, Princeton.

[13] Diebold, F. X., L. E. Ohanian, and J. Berkwitz (1998), Dynamic Equilibrium Economies: A Framework for Comparing Models and Data, Review of Economic Studies, 65, 433-452.

[14] Dunford N. and J. T. Schwartz (1957), Linear Operators, Part 1, Wiley, New York.

[15] Franke, J. and W. Härdle (1992), On Bootstrapping kernel spectral estimates, The Annals of Statistics, 20, No. 1, 121-145.

[16] Granger, C. W. J. (1980), Long Memory Relationships and the Aggregation of Dynamic Models, Journal of Econometrics, 14, 227-238.

[17] Haliassos, and Hassapis (2001), Non-expected Utility, Saving and Portfolios, The Economic Journal, 111, 69-102.

[18] Härdle W., J. Horowitz, and J. -P. Kreiss (2002), Bootstrap Methods for Time Series, Working Paper.

[19] Hosking, J. R. M. (1981), Fractional Differencing, Biometrika, 68, 165-176.

[20] Ilinski, K. (2001), Physics of Finance, John Wiley & Sons, LTD.

[21] Kreiss J. -P., and E. Paparoditis (2000), Autoregressive Aided Periodogram Bootstrap for Time Series, Working Paper, Institute for Mathematical Stochastics, Technical University of Braunschweig, Germany.

[22] LeBaron, B. (2000), Agent-based Computational Finance: Suggested Readings and Early Research, Journal of Economic Dynamics & Control, 24, 679-702.

[23] Levy, M., H. Levy, and S. Solomon (2000), Microscopic simulation of financial markets, Academic Press, New York.

[24] Li, Q. (1996), Nonparametric Testing of Closeness between Two Unknown Distribution Functions, Econometric Reviews, 15, 261-274.

[25] Lux, T. (1998), The Social-economic Dynamics of Speculative Markets: Interacting Agents, Chaos, and the Fat Tails of Return Distributions, Journal of Economic Behavior & Organization, Vol. 33, 143-165.

[26] Lux, T. (2000), Multi-Agent Simulation of Financial Markets and the Stylized Facts, Conference "Beyond Equilibrium and Efficiency", Santa Fe Institute, May 18-20.

[27] Lux, T. and M. Marchesi (1998), Volatility Clustering in Financial Markets: A Micro-Simulation of Interacting Agents, Working paper.

[28] Mantegna, R. N. and H. E. Stanley (2001), An Introduction to Econophysics, Cambridge University Press.

[29] Priestly, M. B. (1981), Spectral Analysis and Time Series, Volume 1, Academic Press.

[30] Scheinkman, J. A. and B. LeBaron (1989), Nonlinear Dynamics and Stock Returns, Journal of Business, 40, no. 3, 311-337.

[31] Solomon, S. and M. Levy (2000), Market Ecology, Pareto Wealth Distribution and Leptokurtic Returns in Microscopic Simulation of the LLS Stock Market Model, http://xxx.lanl.gov/cond-mat/0005416.

[32] Sowell, F. (1992), Maximum Likelihood Estimation of Stationary Univariate Fractionally Integrated Time Series Models, Journal of Econometrics, 53, 165-188.

[33] Swanepoel J. W. H., and J. W. J. van Wyk (1986), The Bootstrap Applied to Power Spectral Density Function Estimation, Biometrika, 73, 1, 135-141.

[34] van der Vaart, A. W. (1998)Asymptotic Statistics, Cambridge University Press.

[35] Voit, J. (2001), The Statistical Mechanics of Financial Markets, Springer-Verlag, Berlin, Heidelberg, New York.

[36] Zschischang, E. and T. Lux (2001), Some New Results on the Levy, Levy and Solomon Microscopic Stock Market Model, Physica A, 291, 563-573.