SVERIGES RIKSBANK
WORKING PAPER SERIES

177

SVERIGES RIKSBANK

# Estimation of an Adaptive Stock Market Model with Heterogeneous Agents

*Henrik Amilon*

JANUARY 2005

WORKING PAPERS ARE OBTAINABLE FROM

Sveriges Riksbank • Information Riksbank • SE-103 37 Stockholm
Fax international: +46 8 787 05 26
Telephone international: +46 8 787 01 00
E-mail: info@riksbank.se

# Estimation of an Adaptive Stock Market Model with Heterogeneous Agents

Henrik Amilon*

Sveriges Riksbank Working Paper Series
No. 177
January 2005

**Abstract**

Standard economic models based on rational expectations and homogeneity have problems explaining the complex and volatile nature of financial markets. Recently, boundedly rational and heterogeneous agent models have been developed and simulated returns are found to exhibit various stylized facts, such as volatility clustering and fat tails. Here, we are interested in how well the proposed models can explain all the properties seen in real data, not just one or a few at a time. Hence, we do a proper estimation of some simple versions of such a model by the use of efficient method of moments and maximum likelihood and compare the results to real data and more traditional econometric models. We discover two main findings. First, the similarities with observed data found in earlier simulations rely crucially on a somewhat unrealistic modeling of the noise term. Second, when the stochastic is more properly introduced we find that the models are able to generate some stylized facts, but that the fit generally is quite poor.

Key words: Efficient method of moments; heterogeneous expectations; bounded rationality; evolutionary dynamics; adaptive beliefs

JEL Classification Numbers: C13, C15, C32, C51, G12

---

*European Central Bank, Postfach 16 03 19, D-60066 Frankfurt am Main, Germany. Phone: +49 69 1344 5824. E-mail: henrik.amilon@ecb.int. I would like to thank Carl Chiarella, Tony He and Marianna Grimaldi for many constructive and helpful discussions. Also Malin Adolfson and Mattias Villani provided most useful comments. A major part of this paper was performed while I was visiting the Quantitative Finance Research Centre at the School of Finance and Economics, University of Technology, Sydney, and the Research Department at Sveriges Riksbank. Their kind hospitality is gratefully acknowledged, and so is financial support from *Bankforskningsinstitutet*, *Föreningssparbankernas Forskningsstiftelse*, the Crafoord Foundation, and the Royal Academy of Science. The views expressed in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank.

# 1 Introduction

Many of the well established models in economics and finance rely on two important cornerstones: homogeneous investors and market efficiency. There is little doubt that people differ in preferences, knowledge, and beliefs, but the homogeneity assumption may still serve as a useful approximation. That depends on if the heterogeneity averages out and can be captured by a single representative agent, a fact that simplifies the analysis tremendously. The other major assumption, the efficient market hypothesis (EMH), is closely related to the rational expectation hypothesis (REH), first introduced by Muth (1961). There are strong and weak versions of the EMH, but generally it states that since all agents are rational and capable of processing information immediately and accurately, prices should reflect all available information. In such a market everyone agrees upon the fundamental price and fluctuations around this equilibrium arise only because of unexpected and random changes in fundamentals.

Empirical investigations of financial series show signs of volatility clustering, excess kurtosis, high persistence and asymmetry in volatility, small dependences in returns, but substantial correlation in absolute or squared returns, see Pagan (1996) for a comprehensive study of characteristic features in financial market data. These stylized facts are difficult to explain with just efficient market fluctuations and call for alternative explanations.

It is often argued, see e.g. Fama (1970), that an efficient market has no predictable patterns (conditioned on public information) since they would disappear as rational traders exploit them. Interestingly, this reveals three important mechanisms. The first is that the EMH is a very conservative theory. The philosophy is that it is of no use to try and find profitable methods, because if they exist they would already be in use. This is in contrast to how the real world works, not just the financial world, where a never-ending search for improvements and superiority is driving the evolution forward. Second, the EMH is sustained because some of the traders do not believe in it since they are constantly looking for predictability, and thirdly, that their actions are fed back and influence future prices. There is a continuous interaction between the market and its participants. Suppose a subset of traders test a strategy where if the price hits a threshold, it will continue up. By chance, they are right and their success is observed by the other agents. Some more agents want to try this strategy which raises the price further, reinforcing and motivating the strategy in a self-fulfilling manner. It is not perfectly rational to rely on such a strategy, especially not if the fundamental price is known, but in an uncertain world it can be hard to tell if a price rise is due to changes in fundamentals, or a speculative overvaluation. Sooner or later the price drops, by chance or because a view that the asset is overvalued is spread among the traders, inducing people to sell. The price drops further and may trigger another trading strategy, and so on. On average, forecastable structures may disappear as agents exploit them, as indicated by the low autocorrelations in asset returns in empirical studies, but may be present during certain periods of time.[1] This indicates that heterogeneity, uncertainty, adaptation, and expectation feedback are plausible components of real markets, and arise naturally if just the extreme mechanisms behind the efficient market hypothesis are slightly lightened.

Bounded rationality, see e.g. Sargent (1993), as opposed to perfect rationality is often used to describe how agents, with limited information about fundamentals, develop expectation models of what moves price and dividends. They are not irrational, but given their limited information they adapt to what they believe is optimal, and given their information set they act rationally. This endogenous uncertainty of the state of the world prevents the agents from forming and solving life-time optimization problems in favour of more simple reasoning and rules of thumb, see e.g. Shefrin (2000).

Recently, the concept of bounded rationality and evolutionary adaptive agents have been modeled in e.g. Brock and Hommes (1997, 1998), Chiarella and He (2002a, 2002b), Gaunersdorfer and Hommes (2000), Hommes (2001), and in a more computationally oriented multi-agent framework in Arthur et al. (1997), LeBaron et al. (1999), LeBaron (2001a), Lux (1995), and Lux and Marchesi (1999).[2] Common to all these heterogenous agent models is the existence of different trader types: Fundamentalists, who believe the asset price is determined solely by economic fundamentals, and technical traders who try to predict future prices by searching for patterns in historical prices. A general and most interesting finding is that these models qualitatively explain a number of the stylized facts mentioned above and that, in contrast to classic financial theory, the technical traders are not driven out of the market. Both types of agents continue to coexist, as they do in real markets.[3]

[1] Many empirical investigations support the fact that technical trading strategies yield significant profits during certain time intervals. The classic paper is Brock et al. (1992). See also Jeegadesh and Titman (2001), and Chiarella and He (2002) for recent references on this topic.

[2] The agents in Arthur et al. (1997), LeBaron et al. (1999), and LeBaron (2001) not only choose the best forecasting rule, but also have the ability to further develop and update it, i.e. the agents can learn. See Evans and Honkapohja (2001) for a recent treatment of algorithmic learning.

[3] There is a close connection to earlier work in behavioral finance on noise-trader models, see DeLong et al. (1990). The distinction of two trader types is made there, but the traders are not adaptive. The less rational noise-traders do systematic mistakes, and continue to do so without adjusting to the outcome

The Adaptive Belief System (ABS) of Brock and Hommes (1997, 1998) models the financial markets as an evolutionary interaction of competing agents, each with a specific trading strategy. The agents are boundedly rational since they choose the strategy that has worked best in the past according to some fitness function such as realized profits, accumulated wealth, or the utility of these quantities. The model includes many desirable key features such as adaptation and expectation feedback. It is complex, yet analytically tractable, and has inspired further extensions for example in Gaunersdorfer and Hommes (2000), Chiarella and He (2002b), and De Grauwe and Grimaldi (2004).

Do the above models also quantitatively explain financial market movements? Rough calibrations in Brock and Hommes (1997), Chen et al. (2001), Gaunersdorfer and Hommes (2000), Gaunersdorfer (2001), and LeBaron (2001b) indicate that some of the statistical properties of the simulated returns resemble those of the real data, but some do not. A close fit to some moments of the data is at the expense of a worse fit to others. It is important to stress that the heterogeneous expectation models are not without dynamic noise. The nonlinear models are fed with an exogenous stochastic process, but the noise process is "nice", that is normally distributed. Instead it is the internal dynamics of the models that should amplify and distort the randomness into the complicated and realistic price fluctuations we observe. This is in sharp contrast to the statistical models used in empirical finance of which the ARCH-class models, and the Markov switching models by far are the most popular, see Bollerslev et al. (1994) and Hamilton (1994) for numerous applications. They have proved to be quite successful in modeling financial data but they do not offer any explanation of why volatility tends to cluster, or why there are switches between different magnitudes in volatility. It would be most satisfactory to be able to explain these phenomena with a structural model, where deviations from the fundamental value are trigged by randomness, and amplified by realistically modeled agents. Besides, if the existing statistical models are approximations of such underlying dynamics, the estimation of the structural model directly would most likely lead to econometric improvements.

In this paper we perform a proper estimation of an adaptive heterogeneous agent model to see if the preliminary simulation results stand to face reality. Unfortunately, when we correct the earlier simulations in Gaunersdorfer and Hommes (2000) and Gaunersdorfer (2001) with a more realistic description of the model stochastics many of the similarities to real data are lost. We focus on the extended version of the model in Brock and Hommes (1998) described in De Grauwe and Grimaldi (2004). The same methodology can, however,

of their strategy. They are truly irrational and not boundedly rational.

be used to estimate many other heterogeneous agent models, at least if they contain a reasonable number of parameters. The usual technique based on maximum likelihood estimation is not always feasible. In these cases, we instead rely on the efficient method of moment (EMM), or indirect inference, approach of Gallant and Tauchen (1996) and Gourieroux et al. (1993). EMM is usually used in estimation of stochastic volatility models, auction data models, and nonlinear rational expectations models, as pointed out in Gallant and Tauchen (1996), and the estimation of a heterogeneous adaptive agent model is therefore a new area to apply this technique. It is our hope that a proper objective estimation also is a helpful tool for the further theoretical development in this field, by pointing out what the models maybe fail to capture, the overall impact of certain parameters, or what features are of importance and should be focused upon.

The paper is organized as follows. Section 2 presents the model under scrutiny, the Adaptive Belief System, while Section 3 describes how stochastic is introduced in the model and how to estimate it. The empirical results of the estimation, and a comparison with real data as well as other statistical models, are given in Section 4. Finally, Section 5 presents some concluding remarks and suggestions for further research. A more detailed description of the EMM is found in the Appendix.

# 2 The adaptive heterogeneous agent model

In this section we present a somewhat generalized version of the Adaptive Belief System of Brock and Hommes, which De Grauwe and Grimaldi (2004) use in a related exchange rate framework. The model consists of three parts: (i) Utility maximizing agents select optimal portfolios based on (ii) different forecasts rules or beliefs about the next period price, and (iii) evaluate the different rules and adopt in the coming period the one with best performance or highest fitness.

## 2.1 The utility maximization

Suppose there are two securities in the economy: one risky asset that pays an uncertain dividend, and one infinitely supplied risk-free asset that pays the constant rate $r$. Let $p_t$ be the ex-dividend stock price, and $y_t$ its stochastic dividend. Following the framework of Brock and Hommes (1998), the wealth of investor $h$ evolves according to

$$W_{h,t+1} = (1+r)W_{h,t} + (p_{t+1} + y_{t+1} - (1+r)p_t)\, z_{h,t} \tag{1}$$

where $z_{h,t}$ is the number of shares at time $t$. Let $E_{h,t}$ and $V_{h,t}$ denote the $h$ investor's expectation and variance operators, conditioned on the information set $F_t = \{p_{t-1}, y_{t-1}, p_{t-2}, ...\}$ of past prices and dividends. Assuming myopic investors with a mean-variance utility function and a common risk-aversion parameter $a > 0$, each investor solves

$$\max_{z_{h,t}} E_{h,t}\left[W_{h,t+1}\right] - \frac{a}{2} V_{h,t}\left[W_{h,t+1}\right] \tag{2}$$

for his optimal amount of shares, which yields

$$z_{h,t} = \frac{E_{h,t}\left[R_{t+1}\right]}{aV_{h,t}\left[R_{t+1}\right]}, \tag{3}$$

where $R_{t+1} = p_{t+1} + y_{t+1} - (1+r)p_t$ is the excess profit. Thus, the investors are characterized by the same utility function and differ only in how they form their beliefs about the conditional expectation and variance.

Suppose $n_{h,t}$ is the fraction of investors with the same beliefs at time $t$. Summing over the demands from all groups of investors gives us the aggregated demand. With a fixed total number of shares in the market, $Z$, we have

$$\sum_{h=1}^{H} n_{h,t} z_{h,t} = Z. \tag{4}$$

Using (3) in (4), the market clearing equilibrium price $p_t$ is determined by

$$(1+r)p_t = \left( \sum_{h=1}^{H} \frac{n_{h,t}}{\sigma_{h,t}^2} E_{h,t}\left[p_{t+1} + y_{t+1}\right] - aZ \right) \frac{1}{\sum_{h=1}^{H} \frac{n_{h,t}}{\sigma_{h,t}^2}}, \tag{5}$$

where the summation is over $H$ groups, and $\sigma_{h,t}^2$ is short-hand notation for $V_{h,t}\left[R_{t+1}\right]$. Assuming net zero supply of the risky asset ($Z = 0$), and an IID dividend process with constant mean $E_t\left[y_{t+1}\right] = \bar{y}$, the market clearing price equation becomes:[4]

$$(1+r)p_t = \left( \sum_{h} \frac{n_{h,t}}{\sigma_{h,t}^2} \left( E_{h,t}\left[p_{t+1}\right] + \bar{y} \right) \right) \frac{1}{\sum_{h=1}^{H} \frac{n_{h,t}}{\sigma_{h,t}^2}}. \tag{6}$$

It is important to note that the fundamental rational expectations (RE) price is nested within the above model. With homogeneous expectations, the arbitrage market equilibrium (6) reduces to

$$(1+r)p_t = E\left[p_{t+1}\right] + \bar{y}. \tag{7}$$

[4] In order to render tractability, Brock and Hommes (1997, 1998) made the additional assumption that beliefs about conditional variance are equal and constant for all types and times, that is, $\sigma_{h,t}^2 = \sigma^2 \forall h, t$.

Using the law of iterated expectations and assuming transversality, the RE price is given by

$$p^* = \sum_{k=1}^{\infty} \frac{\bar{y}}{(1+r)^k} = \frac{\bar{y}}{r}. \tag{8}$$

Here we also allow the dividends to follow a random walk. In that case the market clearing price is determined by

$$(1+r)p_t = \left( \sum_h \frac{n_{h,t}}{\sigma^2_{h,t}} \left( E_{h,t}\left[p_{t+1}\right] + y_t \right) \right) \frac{1}{\sum_{h=1}^{H} \frac{n_{h,t}}{\sigma^2_{h,t}}}, \tag{9}$$

which implies that the RE price also follows a random walk

$$p_t^* = \sum_{k=1}^{\infty} \frac{y_t}{(1+r)^k} = \frac{y_t}{r}. \tag{10}$$

In a perfectly rational world, all investors agree upon the fundamental price of the risky asset. Asset prices change because of unexpected changes in dividends only. In a heterogenous world, on the other hand, where prices are determined by (6), (9), or more generally by (5), asset dynamics can show a much more complex behavior.

## 2.2 The forecast rules

So far we have said nothing about how the agents form their beliefs, that is, their conditional expectations about the future price. Let us assume the existence of two types of traders with the following prediction rules:

$$E_{f,t}\left[p_{t+1}\right] = p^*_{t-1} + \upsilon \left(p_{t-1} - p^*_{t-1}\right), \qquad 0 \leq \upsilon \leq 1, \tag{11}$$

$$E_{h,t}\left[p_{t+1}\right] = p_{t-1} + \lambda \sum_{i=0}^{T} \alpha^{m,h}(1-\alpha^{m,h})^i \left(p_{t-1-i} - p_{t-2-i}\right), \qquad |\lambda| < 1, 0 < \alpha^{m,h} < 1. \tag{12}$$

The first investor is called a *fundamentalist* who believes that tomorrow's price will mean-revert towards the fundamental price by a factor $\upsilon$. When $\upsilon = 1$, these traders believe in the Efficient Market Hypothesis (EMH) and that prices follow a random walk. The second category is *chartists* or *technical traders.* They extrapolate into the future a geometrically declining moving average of past price changes, where $\alpha^{m,h}$ determines the effective time window and $\lambda$ gives the degree of extrapolation. Usually, only deviations from the latest observed price are investigated, that is $\alpha^{m,h} = 0$ and $i = 0$, but other extrapolation rules and lags are analyzed in Chiarella and He (2003). The chartists are further categorized as momentum traders ($\lambda > 0$) or contrarians ($\lambda < 0$), and we see that $\lambda = 0$ also corresponds to EMH-believers. Note that the timing of the information set is of importance. In the Walrasian market equilibrium used here, the market clearing price depends on expectations of $p_{t+1}$. When forming these expectations the agents have not yet observed $p_t$, and therefore use the most recent information from $t-1$. Furthermore, in the special case of a simple dividend process with constant mean, $p^*$ replaces $p^*_{t-1}$. From now on we assume a three-agent model consisting of one group of fundamentalists ($h = f$), trend chasers ($h = tc$), and contrarians ($h = co$), respectively.[5]

The linear updating rule of the chartists is stable if $|\lambda| < 1$ but may lead to too large and too frequent deviations from the fundamental price. We therefore follow De Grauwe and Grimaldi (2004) and introduce a stabilizing force that becomes active if the price deviates too much from its fundamental value, by assuming that the risk aversion of the fundamentalists declines as the misalignment increases:

$$a_{f,t} = \frac{a}{1 + \phi \left| p_{t-1} - p^*_{t-1} \right|}, \qquad \phi \geq 0, \tag{13}$$

$$a_{h,t} = a, \qquad h = tc, co, \tag{14}$$

where $\phi$ measures the sensitivity to the misalignment $|p_{t-1} - p^*_{t-1}|$. If $\phi = 0$ we are back to the case where all agents share the same risk aversion.

In this more general framework, the agents also care about the time-varying risk of their portfolio since $\sigma^2_{h,t}$ enters the market price equilibrium. Here, we follow De Grauwe and Grimaldi (2004) and define the variance term as the geometrically declining weighted average of the squared (one period ahead) forecasting error made by the chartists and fundamentalists, respectively:

$$\sigma^2_{h,t} = \sum_{i=0}^{T} \alpha^{v,h} (1 - \alpha^{v,h})^i \left( E_{h,t-1-i} \left[ p_{t-i} \right] - p_{t-i} \right)^2 \qquad , 0 < \alpha^{v,h} < 1, \forall h. \tag{15}$$

## 2.3 The evolutionary fitness measure

One important thing remains in order to complete the model and that is to specify how the fractions $n_{h,t}$ evolve over time. Let us assume the existence of an evolutionary fitness function or performance measure, $U_{h,t}$. Based on the performance measure, agents make a

[5] The choice of traders is partly motivated by empirical studies who discover profitability for momentum strategies over short time intervals, while contrarian strategies generates profits over longer time intervals. See Chiarella and He (2002b) for further references.

decision of which group to join and whose belief they should rely on. The probability that an agent chooses strategy $h$ is formed on the basis of discrete choice or 'Gibbs' probabilities (see Manski and McFadden, 1981, and Anderson et al., 1993, for a discussion and economic applications of discrete choice models):

$$n_{h,t} = \frac{\exp\left(\beta\left(U_{h,t-1} - C^h\right)\right)}{\sum_h \exp\left(\beta(U_{h,t-1} - C^h)\right)}, \tag{16}$$

where $C^h \geq 0$ measure the cost of strategy $h$, and $\beta \geq 0$ is the intensity of choice measuring how fast agents switch between different prediction strategies. Usually $C^f > 0$ for the fundamentalists to represent an information cost associated with revealing the fundamental price ($p^*_{t-1}$ or $p^*$), in the spirit of Routledge (1999). If $\beta = 0$, the traders are indifferent to differences in fitness and all fractions will be constant and equal to $1/H$. The other extreme case, $\beta = \infty$, corresponds to the case where all traders immediately switch to the most successful trading strategy last period. In the intermediate case with a finite and positive $\beta$, agents make their predictions according to their fitness, but choose less optimal strategies with a certain probability. The market display herd behavior, but with an inertia and a scepticism about the optimal strategy.

With mean-variance investors a natural performance measure, adopted in De Grauwe and Grimaldi (2004), is the utility from past profits of a one unit investment or, for short, past risk-adjusted profits.[6] The risk-adjusted profit for strategy $h$ at time $t$ is given by

$$\pi_{h,t} = R_t s(z_{h,t-1}) - \frac{a}{2}\sigma^2_{h,t}, \tag{17}$$

where $s(x)$ is the signum function, that is, $s(x) = 1$, if $x > 0, s(x) = 0$ if $x = 0$, and $s(x) = -1$ if $x < 0$. Thus from (3), when agents' forecasts of the sign of the excess profit $E_{h,t-1}\left[R_t\right]$ are correct their risk adjusted profits increase. A suitable performance measure can then be defined as

$$U_{h,t} = \pi_{h,t} + \eta U_{h,t-1}, \tag{18}$$

where $0 \leq \eta \leq 1$ is a memory parameter. If $\eta = 0$, only the performance in the last period is of interest, while with a positive $\eta$, the weights given to past utilities of profits decrease exponentially. The complete Adaptive Belief System is thus described by the price equation (6), the evolutionary dynamics (16), and the fitness function (18).

[6]It should be noted that the approach in Brock and Hommes is somewhat different as they use the utility from past realized profits, that is $\pi_{h,t} = R_t z_{h,t-1} - a\sigma^2 z^2_{h,t-1}/2$, where $z_{h,t-1} = E_{h,t-1}\left[R_t\right]/(a\sigma^2)$ is the demand from (3) with a constant $\sigma^2_{h,t}$. The profitability of a rule will then also depend on if the demand is temporarily high or low. Both approaches were tested here with small differences as a result.

# 3 The model stochastics

The dynamic properties of a very similar two-agent Adaptive Belief System with a fundamentalist and a trend-follower, both with a common and constant perception of risk and a constant dividend process, have been thoroughly investigated in Gaunersdorfer and Hommes (2000), and Gaunersdorfer et al. (2000). For different parameter configurations, the system is shown to undergo different bifurcations and to exhibit complicated non-linear, and even chaotic, price behavior. Still, the purely deterministic model is too simple to capture the dynamics of real stock markets. Until now we have been deliberately vague about how the stochastic is introduced in the model. Adding IID noise to a constant dividend enters the price equation (9) only from $R_t$ in (17) via (18) and (16), and is insufficient to render realistic price dynamics.[7] In fact, Gaunersdorfer (2001) and Gaunersdorfer and Hommes (2000) just use a constant dividend and instead add a dynamic Gaussian noise term $\varepsilon_t$, representing a model approximation error, to the market clearing equation (6):

$$p_t = \frac{1}{1+r}\left(\sum_h \frac{n_{h,t}}{\sigma^2_{h,t}}\left(E_{h,t}\left[p_{t+1}\right]+\bar{y}\right)\right)\frac{1}{\sum_{h=1}^H \frac{n_{h,t}}{\sigma^2_{h,t}}} + \varepsilon_t, \qquad \varepsilon_t \sim N(0,\rho). \tag{19}$$

This results in a noisy chaotic system from which returns, $r_t = p_t/p_{t-1} - 1$, are calculated. In a number of simulations in the above references, this model generates return series that exhibits low autocorrelation, volatility clustering, and excess kurtosis, the trademarks of real financial markets. The problem is that in (19) prices are stationary around a constant $p^*$. By adding noise of the same variance to all prices, those prices that are relatively low will vary more than prices that are above $p^*$. This is not a realistic way to model noise and return series calculated from such prices will show signs of the above stylized facts by construction only. In fact, if the non-linearities of the system are turned off and prices are generated from a random walk, $p_t = p_{t-1} + \varepsilon_t$, the resulting return series will have all the desired properties mentioned above. A more proper way to model the stochastic part is instead as usual:

$$p_t = \frac{1}{1+r}\left(\sum_h \frac{n_{h,t}}{\sigma^2_{h,t}}\left(E_{h,t}\left[p_{t+1}\right]+\bar{y}\right)\right)\frac{1}{\sum_{h=1}^H \frac{n_{h,t}}{\sigma^2_{h,t}}} + p_{t-1}\varepsilon_t \equiv \hat{p}_t + p_{t-1}\varepsilon_t, \quad \varepsilon_t \sim N(0,\rho). \tag{20}$$

[7] It should be noted that De Grauwe and Grimaldi (2004) compute their returns as price *differences.* This means that their noise is correctly modeled if their price (exchange rate) is interpreted as the log of the price. On the other hand, it is not obvious in their utility based exchange rate model why agents should agree upon the log of the price and not the price.

In this case, pure IID returns are nested within the model as should be a minimum requirement for any return specification. As we will see, the ability of the model to generate real dynamics is greatly affected.

Our main concern is to estimate the model. Assuming a suitable error distribution, in our case we choose the Gaussian since any non-normal behavior should stem from the intrinsic nonlinear dynamics, the likelihood function for estimating the model from the observed prices ($p_t$) is given by:

$$\ln L = -\frac{1}{2}\sum_{t=1}^{T}\left(\ln(2\pi) + \ln\left(\rho^2\right) + \left(\frac{p_t - \hat{p}_t}{p_{t-1}}\right)^2 \frac{1}{\rho^2}\right).$$

If the dividend yield is assumed to follow a random walk, the stochastic enters the pricing mechanism not only through $n_{h,t}$ as before, but also directly in (9), and more significantly through $p_t^*$ from (10). It should be clear that by a random walk we mean

$$y_t = y_{t-1} + y_{t-1}\varepsilon_t^{rw}, \qquad \varepsilon_t^{rw} \sim N(0, \rho^{rw}) \tag{21}$$

so that the fundamental price process and the resulting return process do not suffer from the shortcomings previously mentioned. This way of introducing stochastics gives much richer dynamics but nothing prevents from adding a model approximation error as well, yielding

$$p_t = \frac{1}{1+r}\left(\sum_h \frac{n_{h,t}}{\sigma_{h,t}^2}\left(E_{h,t}\left[p_{t+1}\right] + y_t\right)\right)\frac{1}{\sum_{h=1}^{H}\frac{n_{h,t}}{\sigma_{h,t}^2}} + p_{t-1}\varepsilon_t, \qquad \varepsilon_t \sim N(0,\rho), \tag{22}$$

where $\varepsilon_t^{rw}$ and $\varepsilon_t$ are independent and normally distributed.

Unfortunately, the error terms are now more embedded in the nonlinear structure, which means that a computable likelihood function is no longer available. We therefore have to rely on simulation based econometric methods, where some moments of the simulated data are matched to those of the real data. This idea follows Hansen's (1982) generalized method of moments and is described in the simulated method of moment procedure of Duffie and Singleton (1993). In order to avoid selecting the moments on an ad hoc basis, Gallant and Tauchen (1996), and Gourieroux et al. (1993) systemize which moments to match in what has become known as the efficient method of moments (EMM) or indirect inference. This estimation procedure involves simulating data from the structural Adaptive Belief System with noise added as in (21) and (22), compute simulated returns, $r_t = p_t/p_{t-1} - 1$, evaluate a score vector of an estimated auxiliary model with these simulated returns, and then use this vector as moment conditions to be minimized in order to obtain the unknown parameters of the Adaptive Belief System. Further details are given in the appendix.

# 4 Empirical results

The complete Adaptive Belief System with noise is thus described by the price equations (20) or (22), the dividend process, the evolutionary dynamics (16), and the fitness function (18). No proper estimations of an Adaptive Belief System has to our knowledge been performed, and that is exactly what we do here, in order to more objectively investigate if any of the versions of the model is capable of describing real market data. First, we estimate an auxiliary GARCH model to be used in the EMM estimation of a two-agent Adaptive Belief System with the more complex stochastic structure in (22). Here, we are forced to economize on the number of parameters to estimate. Second, we do a maximum likelihood estimation of a similar Adaptive Belief System but with the additive noise distribution in (20) only, in order to see how much, if anything, is lost with this simpler approach. At this stage, we estimate a more flexible three-agent Adaptive Belief System to investigate the gain of this added complexity. With a pure additive noise term, the number of parameters is no longer a constraint.

It should be noted that the choice of dividend yield processes used here imply that prices do not grow exponentially over the long run in sharp contrast to real prices, but by using a growing non-stationary dividend process, rapidly increasing prices can be obtained.[8] In order to have a close connection to the theoretical results and earlier simulations, we hold on to the original formulation of the dividend yield processes. In our empirical work we therefore use detrended daily data of the S&P 500 index from January 1980 to December 2000, that is we calculate return series from the index, normalize these to zero mean, and re-compute index observations from the demeaned return series. Figure 1 shows the detrended index series and the return series. It is evident from Figure 1a that the variability of the price series depends on the index level, which would not have been the case for a price series generated from (19).

In order to capture as much of the (non-normal) characteristics of the data as possible it is important to have as long series as possible. Still, a few events like the 1987 crash have a large impact on the statistical properties of the data set and the resulting estimation, see e.g. Engle and Lee (1994). Since, to our knowledge, no structural or statistical model has successfully captured these extreme observations jointly with the rest of the data, we choose to eliminate the observations belonging to the crash and the two consecutive days.[9] Another reason is to facilitate the comparisons with the simulations in Brock et al.

[8]See also the related model in Chiarella and He (2002a), where the underlying CRRA utility function leads to growing price (and wealth) processes.

[9]The excluded returns are -20.4%, 5.3%, and 9.1%, resulting in a sample length of 4549 returns.

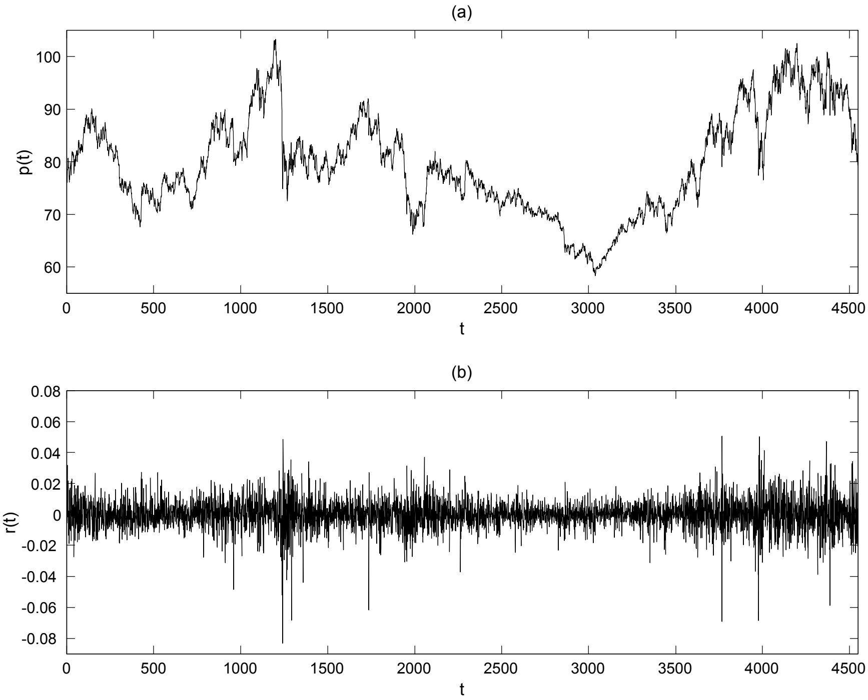


Figure 1: (a) Daily S&P 500 detrended prices December 1982 – December 2000. The observations belonging to the October 1987 crash and the two concecutive days are excluded. (b) Corresponding S&P 500 returns.

(2001), Gaunersdorfer and Hommes (2000) and Hommes (2001), who also exclude these data points from their sample. The volatility clustering and the extreme events are seen in Figure 1a. Even after the censoring, it is quite a challenge to find a model that fits this data set.[10]

## 4.1 ML estimation of the auxiliary model

Natural choices of auxiliary models are the ARCH-class models, initially proposed by Engle (1982). They can potentially capture many stylized facts of financial data, such as excess kurtosis, and volatility clustering and asymmetry. There have been many extensions of the original model, and several of them have proved to be quite successful in modeling many financial variables, not just equity indices, see e.g. the survey in Bollerslev et al. (1994). Important for our concern, they also have a well-defined likelihood function.

[10]For reasons that will become clear later, the likelihood of the GARCH estimates is based on $t > 500$. Hence, the first two years of observations are excluded in Figure 1.

Table 1: GARCH estimation with normally distributed errors.

| | $\beta_1$ | $\alpha_0$ | $\alpha_2$ | $\alpha_3$ | $\alpha_L$ | $\omega \times 10^5$ | $\gamma_1$ | $\gamma_2$ | $\gamma_3$ |
|---|---|---|---|---|---|---|---|---|---|
| Estimate | 0.049 | 0.011 | 0.591 | 0.169 | 0.064 | 0.164 | 0.799 | −0.779 | 0.885 |
| $t$-value | 3.20 | 2.48 | 2.73 | 0.85 | 4.77 | 4.39 | 52.6 | −34.0 | 37.2 |

The specification we use is the AR($o$)-GARCH($p,q$) process of Bollerslev (1986), extended with a leverage term to account for the fact that large downward moves tend to have larger effects on future volatility than upward moves of comparable size. Pagan and Schwert (1990) discuss how this asymmetry should be modelled, and our choice is similar in spirit to the one used in Engle and Lee (1994) who model the S&P 500 index from 1971 to 1990:

$$
\begin{aligned}
r_t &= \beta_1 r_{t-1} + \varepsilon_t, \\
\varepsilon_t &= \sigma_t u_t, \\
g_t &= (\alpha_0 + \alpha_L D_t)\varepsilon_t^2 \\
\sigma_t^2 &= \omega + \sum_{j=1}^{p} \gamma_j \sigma_{t-j}^2 + \sum_{j=1}^{q} \alpha_j g_{t-j},
\end{aligned}
\tag{23}
$$

where $D_t = 1$ if $\varepsilon_t < 0$, zero otherwise, and $\alpha_1 \equiv 1$.[11] A positive $\alpha_L$ thus reflects the leverage effect of the lagged squared innovations.

We allow the disturbances $u_t$ to be generated by a standard Normal distribution, but a specification with the normalized Student-$t$ distribution of Bollerslev (1987) was also tested. The Student-$t$ distribution is capable of handling the fat tails of the distribution, but the estimation does not rely on the powerful QML results of Bollerslev and Wooldridge (1992): The QML theory applies if the assumed normality assumption is false, but not if an assumed Student-$t$ distribution is misspecified. The Gaussian results may therefore be more robust, and the worse fit of the EMM estimation of the structural model when using a GARCH-model with Student-$t$ distributed errors as an auxiliary model showed indeed some support for this observation. Hence, only estimates of the Gaussian case is presented.

The GARCH estimation results are presented in Table 1. We use three lags in the conditional variance process, i.e. $p = q = 3$, which is more than usually used in modeling asset returns. Almost all estimates are significant, but standard information criteria (Hannan-Quinn and Schwarz, but not Akaike) show signs of overparametrization. However, there

[11] For numerical reasons, the threshold $D_t$ is implemented with the differentiable function $d(x) = \exp(-K * x) / (1 + \exp(-K * x)), K = 10000$.

Table 2: Moments and diagnostics of the standardized residuals and the S&P 500 return data.

| Data | Std. Dev. | Skewness | Kurtosis | $Q(10)$ | $Q^2(10)$ |
|---|---|---|---|---|---|
| Std. res. | 1.00 | −0.43 | 6.36 | 12.8 | 2.95 |
| | 1.00 | 0.00 | 3.00 | 0.24 | 0.98 |
| S&P 500 | 0.0095 | −0.46 | 8.65 | 40.2 | 697.3 |
| | | | | 0.00 | 0.00 |

Note: $Q(10)$ and $Q^2(10)$ are the Ljung-Box statistics for the return and squared return data with 10 lags, respectively. Small numbers are the moments predicted by the statistical distributions, except for the Ljung-Box statistics, where they are $p$-values of a $\chi^2_{10}$ distribution.

are two reasons for choosing this specification: first and foremost, the order condition of identification is that there are at least as many parameters in the auxiliary model as there are in the structural model. Second, the information criteria favor additional parameters in the conditional variance process compared to the conditional mean, and this specification maximizes the Hannan-Quinn and Schwarz information criteria for this particular number of parameters. The relevant persistence measure is $\sum \gamma_j + (1 + \sum \alpha_j)(\alpha_0 + \alpha_L/2)$, meaning that volatility forecasts decay with the power of 0.981. The significance of $\alpha_L$ supports the leverage effect in the data.

Table 2 takes a closer look at the observed data and the standardized residuals of the estimations. The Ljung-Box test indicates that the model successfully removes the dependences in the second moment, but the issue of non-normality remains. We see that the standardized residuals retain almost all skewness and a large part of the excess kurtosis of the original data. By comparing the sample moments by the moments predicted by the assumed distributions, we see that the models overestimate the skewness and underestimate the kurtosis. The implication is that more flexible distributions may be needed, like the exponential generalized beta family in Wang et al. (2001).[12]

## 4.2 EMM and ML estimations of the structural model

To obtain our EMM estimates, the GARCH scores are evaluated with simulated data from the structural model as to minimize (25), see the appendix. We choose the sample size to five times the length of the original data, $N = 5 \times T$.[13] This is the effective sample

[12] A further expansion into semi-nonparametric (SNP) conditional densities introduced in Gallant and Nychka (1987) is also possible, and applied, amongst others, to interest rate modeling in Andersen and Lund (1997).

[13] It would be desirable to have a larger $N$ in order to reduce the Monte Carlo error. Unfortunately, the time consuming estimations prevent us from this at the present.

Table 3: ML and EMM estimates of the ABS.

| Model | $\beta$ | $\upsilon$ | $\lambda^{tc}$ | $\lambda^{co}$ | $\phi$ | $a$ | $\rho$ ($\times 10^3$) | $\rho^{rw}$ ($\times 10^3$) | $\alpha^{m,tc}$ | $\alpha^{m,co}$ | $\alpha^{v,f}$ | $\alpha^{v,tc}$ | $\alpha^{v,co}$ | $C^f$ | $\eta$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ML | 1.94 | 0.02 | 0.01 | -0.32 | 0.00 | 1.07 | 9.47 | | 0.32 | 0.13 | 0.05 | 0.20 | 0.51 | 0.20 | 0.08 |
| S.E | 0.01 | 0.03 | 0.02 | 0.00 | 0.02 | 0.02 | 0.00 | | 0.02 | 0.01 | 0.01 | 0.04 | 0.02 | 0.02 | 0.01 |
| EMM | 1.99 | 0.55 | 0.11 | | 0.07 | 1 | 5.25 | 6.36 | 0.16 | | 0.42 | 0.42 | | 0 | 0 |
| S.E | 20.44 | 0.75 | 0.98 | | 0.25 | | 2.58 | 0.83 | 2.13 | | 0.22 | 0.22 | | | |

Note: Parameters with no entries for standard errors are fixed in the estimations. In the EMM estimation $\alpha^{v,f}$ and $\alpha^{v,tc}$ are restricted to be identical.

size, that is after the first 2000 simulated values are discarded in order to let any initial effects to wear off.[14] It should be mentioned that we encountered some numerical difficulties. Different starting values resulted in different terminal values, suggesting that the minimizing function is very flat, or has several local minima. The statistical properties of the different solutions are quite similar though. A global optimization technique, such as simulated annealing (see Press et al., 1992, and Goffe et al., 1994 for implementations and further references) would be useful, but is for the moment not feasible due to computational limitations.

The value of $r$ is chosen to correspond to a yearly risk-free rate around 4% and $\bar{y}$ to give a fundamental price of 100. Furthermore, since identification requires that the parameters in the structural model to be less than in the auxiliary model and since EMM estimation is sensitive to overfitting of the auxiliary model, we restrict the number of parameters to estimate with EMM to eight and fix the rest to some plausible figures.

When we use the simpler noise structure in (20) we can allow ourselves to estimate a larger number of parameters. In this case, we estimate all 14 parameters of a three-agent model.

The estimates and standard errors of our two Adaptive Belief Systems are displayed in Table 3. Starting with the ML estimates, we see that most of the parameters are highly significant. The mean-reverting parameter of the fundamentalists, $\upsilon$, is very small and insignificant and so is the extrapolation parameter, $\lambda^{tc}$, of the trend chasers. The third agent is, on the other hand, significantly identified as a contrarian ($\lambda^{co} < 0$). Interestingly, since $\phi$ is insignificant all agents share the same risk aversion, and there seems not to be a need for a stabilizing force driving the chartists out of the market via equation (13). The success of contrarian strategies are in the literature documented for longer time horizons, while the momentum strategies seem to be more profitable on shorter horizons, see e.g. Chiarella and He (2002b). Here, the conclusions are somewhat mixed since the contrarians

[14] We further discard the first 500 draws of the score vector because of possible initial transients. The same is also done when estimating $W_T$ in (26).

use a longer time-window than the trend chasers in their updates of the variance, while the opposite is true for the use of past price observations when they form their beliefs about the future price. Furthermore, the information cost of the fundamentalists, $C^f$, is clearly identified and some weight is also attributed to past utilities, although the investors are restricted in this respect since $\eta$ is identical for all trader types. In spite of the significant estimates, the standard deviation of the noise term is not far from the standard deviation of the raw data in Table 2, which suggests a quite modest model fit.

The EMM estimates of a two-agent system with fewer free parameters but also a more complex disturbance structure in (22) show quite a different picture. Here, the risk aversion parameter is fixed but the fundamentalists are still allowed to differ by the influence of $\phi$. Besides, both fundamentalists and trend chasers are assumed to share the same time-window in determining the time-varying risk in (15). The $\beta$-parameter that determines the speed of the switches between investor groups is of the same magnitude as in the earlier case, but it is no longer significant. In fact, no parameter is estimated to be significant at the 5% level, except for the variances of the two exogenous error terms! Although insignificant, the mean-reversion behavior of the fundamentalists is now much stronger, and the only group of technical traders consists in this case of momentum traders since the estimate of the extrapolation parameter is positive, 0.11.

Despite the disappointing findings from the EMM estimations, the top part of Figure 2 shows a representative part of around four years of the difference between the simulated prices, Equation (22) and the fundamental prices, Equations (10) and (21). We see that periods of small price deviations are constantly interrupted by episodes where prices wander away from the fundamental price quite substantially and persistently. From the bottom part of the figure, we see that these movements coincide with the weight of the trend-followers: when the price deviations increase, the trend chasers tend to dominate the market and vice versa. In fact, these bubbles and crashes are typically triggered by exogenous noise and then further reinforced by the growing population of chartists until the price starts to move in the other direction, again triggered by a shock of such magnitude that the perception among the two investor groups of where the market is going is altered. This everlasting struggle among the investors determines the whole dynamics of the system and cannot be stressed enough. When prices start to move away, say up, from the fundamental price, the fundamentalists expect a back-drop but are overwhelmed by the steadily increasing chartists group who profit from the price increase which, in turn, attracts more investors to become chartists in a self-fulfilling manner. When almost everyone has become a chartist, the upward pressure on prices slows down and the more pessimistic view of the fundamentalists becomes relatively important. A negative shock

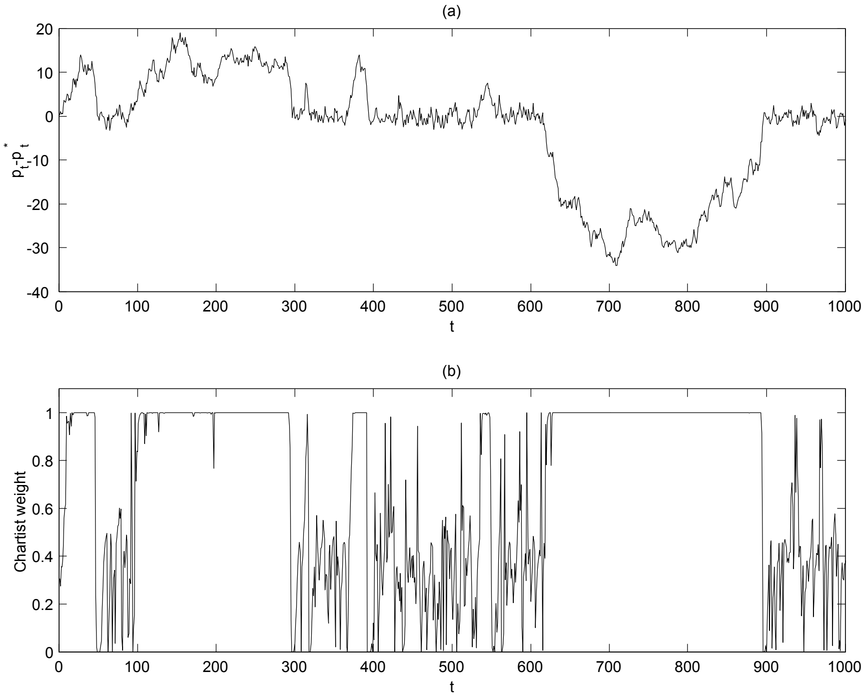


Figure 2: (a) 1000 differences between the simulated and fundamental prices from the EMM estimation, and (b) the corresponding weights of the momentum traders.

may then result in a price decline, which increases the profits of the fundamentalists, and so the journey back towards the fundamental price levels starts again.

Such a boundedly rational exuberance is interesting per se, but it does not reveal if the resulting return series show any resemblances with true data. Table 4 therefore shows a more thorough investigation of the model diagnostics. We see from the simulated EMM return series that the model generates some non-normal behavior. The kurtosis is 3.2, which is clearly non-normal, but still far from the 8.6 of the S&P 500. The Ljung-Box portmanteau test for dependences in returns resembles that of the real data, while the same test statistic for dependences in squared returns, which often is regarded as an indication of ARCH-effects, although highly significant is lower than for the real data. Unfortunately, the small but promising features of the model do not come without a cost. The standard deviation of the EMM series is considerably less compared to the S&P 500. It is not obvious why the model in some sense trades the second moment for the fourth one, but a simulation of 1000 series of the same length as the original data (4549 return observations) with the EMM parameter estimates confirms a slight negative correlation between the standard deviation and the kurtosis of the 100 series with largest kurtosis.

Table 4: Moments and diagnostics of the return series of the ML and EMM estimations.

| | Std.Dev. $(\times 10^3)$ | Skewness | Kurtosis | $Q(10)$ | $Q^2(10)$ | $\chi^2$ |
|---|---|---|---|---|---|---|
| EMM | 7.06 | −0.014 | 3.22 | 24.73<br>0.01 | 182.45<br>0.00 | 29.27<br>0.00 |
| ML | 9.53 | 0.00 | 3.04 | 18.17<br>0.05 | 15.94<br>0.10 | |
| S&P 500 | 9.50 | −0.46 | 8.65 | 40.16<br>0.00 | 697.27<br>0.00 | |

Note: The test of overidentifying restrictions is distributed as $\chi_1^2$. Small numbers indicate $p$-values.

Not surprisingly, the $\chi^2$ test for overidentifying restrictions in Table 3, Equation (28), provides striking evidence against the two-agent structural model with two exogenous noise processes in favor of the auxiliary GARCH model. The $p$-value is basically zero.

The poor fit of the ML model that was suggested already in Table 3 is again seen in Table 4, which shows the results from 1000 simulations from the three-agent model with the simpler noise structure. The kurtosis is now only slightly (but actually significantly) larger than three, and now also without a deterioration of the standard deviation. But unfortunately, the Ljung-Box tests show no longer any indications of any volatility clustering which together with excess kurtosis must be regarded as the most apparent stylized facts.

The Ljung-Box statistics are useful summary statistics but it could be a point to examine the cause of the rejections somewhat more closely. Figure 3 therefore shows the autocorrelation function of returns and squared returns for the S&P 500 data, a simulation of the auxiliary AR-GARCH model, and the EMM simulation. We see that the autocorrelations of squared returns are quite different for the structural model compared to the other return series. It appears as if the Ljung-Box rejection stems from the first lags only, which is in stark contrast to the behavior of the real data and the GARCH model which decays very slowly. The sample auto-correlation in returns is small for all lags for all series, and it seems also to be somewhat less erratic for the structural model than for the other two series. There is little doubt that the GARCH model more successfully captures the stylized facts of real market returns. It should be stressed that higher excess kurtosis and autocorrelation in the second moment still are possible to generate with the structural models, but only at the expense of a much worse fit to other moments, mainly an unrealistically large autocorrelation in the first moment. This is also the reason why these parameter configurations are rejected by the data.

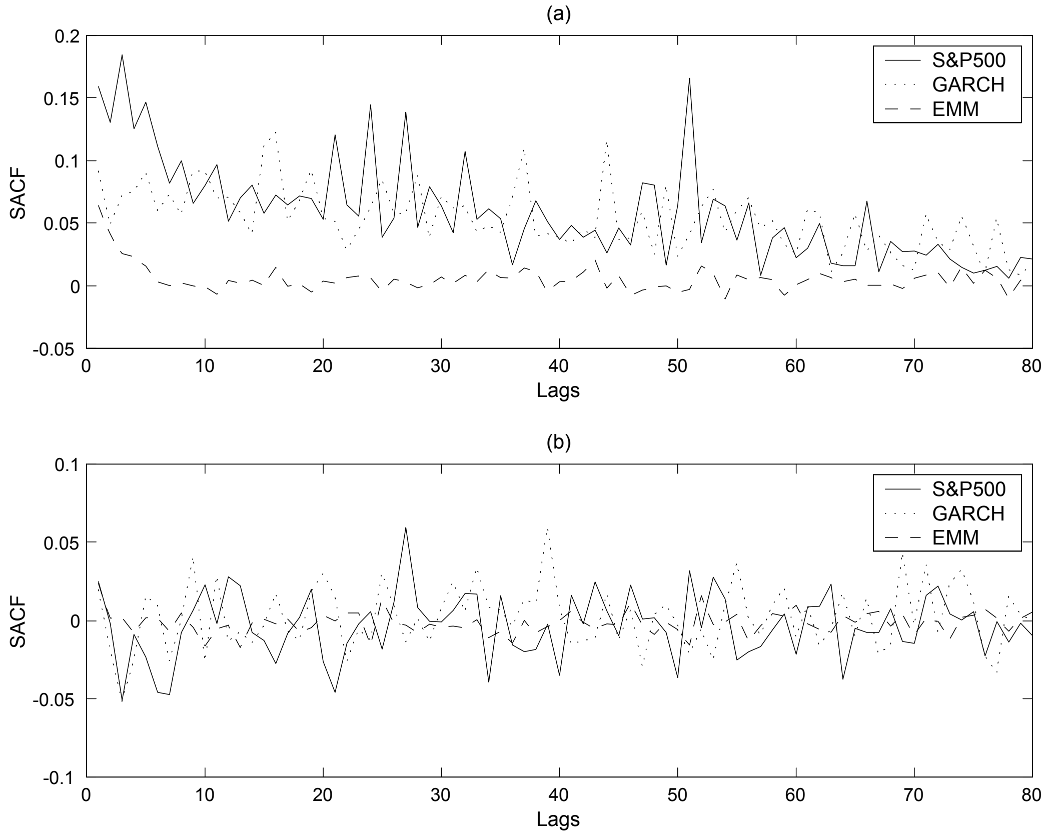


Figure 3: (a) Sample autocorrelation functions (SACF) of squared returns for the S&P 500, a simulted series from the auxiliary AR(1)-GARCH(3,3) model, and the EMM simulation. (b) The same for returns.

# 5 Conclusions

In recent years, different structural models that try to explain the complex behavior of financial markets have been proposed. A class of models that have shown promising theoretical results are the Adaptive Belief Systems, originating from Brock and Hommes (1997), where heterogeneous agents equipped with different expectations determine the market price. A key feature is adaptation: a successful forecast rule will attract other investors and vice versa. The quantitative aspects of the model are, however, not as carefully explored. This has mostly been limited to comparing the size of different moments and the autocorrelation structure to some stock market indices and by fitting GARCH models to simulated series, as in Chen et al. (2001) and Gaunersdorfer (2000). Here, we try to find out how well the proposed models can explain all the properties seen in real data, not just one or a few at a time.

Hence, in this paper we estimate two versions of an Adaptive Belief System by the use of maximum likelihood and the EMM technique of Gallant and Tauchen (1996) and Gourieroux et al. (1993). We discover two main findings. First, the similarities with observed data found in earlier simulations rely crucially on a somewhat unrealistic modeling of the noise term. Second, when the stochastic is more properly introduced we find that the models are able to generate some stylized facts, but that the fit generally is quite poor. The results are in some sense disappointing since we cannot find an adequate fit to the observed data. On the other hand should we be encouraged since the models under scrutiny are simple prototypes and still seem to explain some empirical facts. We did also discover local minima. It may therefore still be the case that there exists a global minima that generates the desired real market behaviour, but which we failed to find. In this respect, a global optimization algorithm would be most helpful, but unfortunately also very time consuming.

A more serious shortcoming is that the model only involves a few trader types, while in reality there are many. It is straightforward to extend the model to a true multi-agent framework, such as Arthur et al. (1997) and Lux (1995), but at the loss of tractability and an increasing number of parameters. An elegant theory is developed in Brock et al. (2001), who introduces the Large Type Limit (LTL) system. The basic idea is that the many agents' parameters are assumed to be drawn from some convenient distribution, typically a multivariate normal, thus reducing the degree of freedom tremendously while still keeping the essence of a multi-agent model.

The estimation technique used in this paper can be applied to the extended models as well. Hopefully, estimations similar to this will help future research to augment the models with features that match the observed data more successfully. It is difficult to see, though, how this can be accomplished without a specification of a time-varying second moment of an exogenous stochastic process.

# References

Andersen, T. and J. Lund, 1997, Estimating continuous-time stochastic volatility models of the short-term interest rate, Journal of Econometrics, 343-377.

Anderson, S., de Palma, A. and J. Thisse, 1993, Discrete choice theory of product differentiation (MIT Press, Cambridge).

Arthur, W., J. Holland, B. LeBaron, R. Palmer and P. Taylor, 1997, Asset pricing under endogenous expectations in an artificial stock market, in: W. Arthur, D. Lane and S. Durlauf, The economy as an evolving complex system II (Addison-Wesley, Reading).

Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, 307-327.

Bollerslev, T., 1987, A conditional heteroskedastic time series model for speculative prices and rates of return, Review of Economics and Statistics 69, 542-547.

Bollerslev, T., R. Engle and D. Nelson, 1994, Arch models, in: R. Engle and D. McFadden, Handbook of Econometrics, Vol. 4 (Elsevier, Amsterdam), 2959-3038.

Bollerslev, T. and J. Wooldridge, 1992, Quasi-maximum likelihood estimation of dynamic models with time varying covariances, Econometric Reviews 11, 143-172.

Brock, W. and C. Hommes, 1997, A rational route to randomness, Econometrica 65, 1059-1095.

Brock, W. and C. Hommes, 1998, Heterogeneous beliefs and bifurcation routes to chaos in a simple asset pricing model, Journal of Economic Dynamics and Control 22, 1235-1274.

Brock, W., C. Hommes and F. Wagener, 2001, Evolutionary dynamics in financial markets with many trader types, Working Paper, CeNDEF, University of Amsterdam.

Brock, W., J. Lakonishok and B. LeBaron, 1992, Simple technical trading rules and the stochastic properties of stock returns, Journal of Finance 47, 1731-1764.

Chen, S-H., T. Lux and M. Marchesi, 2001, Testing for non-linear structure in an artificial financial market, Journal of Economic Behavior and Organization 46, 327-342.

Chiarella, C. and X. He, 2002a, An adaptive model on asset pricing and wealth dynamics with heterogeneous trading strategies, Research Paper 84, Quantitative Finance Research Group, University of Technology, Sydney.

Chiarella, C. and X. He, 2002b, Heterogeneous beliefs, risk and learning in a simple asset pricing model, Computational Economics 19, 95-132.

Chiarella, C. and X. He, 2003, Dynamics of beliefs and learning under $a_L$-process – the heterogeneous case, Journal of Economic Dynamics and Control 27, 503-531.

De Grauwe, P. and M. Grimaldi, 2004, Exchange rate puzzles: a tale of switching attractors, Working Paper 2004:163, Sveriges Riksbank.

Duffie, D. and K. Singleton, 1993, Simulated moments estimations of Markov models of asset prices, Econometrica 61, 929-952.

Engle, R., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987-1007.

Engle R., and G. Lee, 1994, Estimating diffusion models of stochastic volatility, in: P. Rossi, Modelling stock market volatility: Bridging the gap to continuous time (Academic Press, San Diego).

Fama, E., 1970, Efficient capital markets: A review of theory and empirical work, Journal of Finance 25, 383-417.

Gallant, R. and D. Nychka, 1987, Semi-nonparametric maximum likelihood estimation, Econometrica 55, 363-390.

Gallant, R., and G. Tauchen, 1996, Which moments to match?, Econometric Theory 12, 657-681.

Gaunersdorfer, A., 2001, Adaptive beliefs and the volatility of asset prices, Central European Journal of Operations Research, 5-30.

Gaunersdorfer, A. and C. Hommes, 2000, A nonlinear structural model for volatility clustering, Working Paper 00-02, CeNDEF, University of Amsterdam.

Gaunersdorfer, A. , C. Hommes and F. Wagener, 2000, Bifurcation routes to volatility clustering, Working Paper 00-04, CeNDEF, University of Amsterdam.

Goffe, W., G. Ferrier and J. Rogers, 1994, Global optimization of statistical functions with simulated annealing, Journal of Econometrics 60, 65-99.

Gourieroux, C., A. Monfort and E. Renault, 1993, Indirect inference, Journal of Applied Econometrics 8, 85-119.

Hamilton, J., 1994, Time series models (Princeton University Press).

Hansen, L., 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029-1054.

Hommes, C., 2001, Financial markets as nonlinear adaptive evolutionary systems, Quantitative Finance 1, 149-167.

Jegadeesh, N. and S. Titman, 2001, Profitability of momentum strategies: An evaluation of alternative explanations, Journal of Finance 56, 699-720.

LeBaron, B., 2001a, Evolution and time horizons in an agent based stock market, Macroeconomic Dynamics 5, 225-254.

LeBaron, B., 2001b, Empirical regularities from interacting long and short horizon investors in an agent-based stock market, IEEE Transactions on Evolutionary Computation 5, 442-455.

LeBaron, B., W. Arthur and R. Palmer, 1999, Time series properties of an artificial stock market, Journal of Economic Dynamics and Control 23, 1487-1516.

Lux, T., 1995, Herd behavior, bubbles and crashes, Economic Journal 105, 881-896.

Lux, T. and M. Marchesi, 2000, Volatility clustering in financial markets: A microsimulation of interacting agents, International Journal of Theoretical and Applied Finance 3, 675-702.

Manski, C. and D. McFadden, 1981, Structural analysis of discrete data with econometric applications (MIT Press, Cambridge).

Muth, J., 1961, Rational expectations and the theory of price movements, Econometrica 29, 315-335.

Pagan, A., 1996, The econometrics of financial markets, Journal of Empirical Finance 3, 15-102.

Pagan, A. and G. Schwert, 1990, Alternative models for conditional stock volatility, Journal of Econometrics 45, 267-290.

Press, W., B. Flannery, S. Teukolsky and W. Vetterling, 1992, Numerical recipes in C: The art of scientific computing (Cambridge University Press, New York).

Routledge, B., 1999, Adaptive learning in financial markets, Review of Financial Studies 12, 1165-1202.

Sargent, T., 1993, Bounded rationality in macroeconomics, (Clarendon Press, Oxford).

Shefrin, H., 2000, Beyond greed and fear: Understanding behavioral finance and the psychology of investing (Harvard Business School Press, Boston).

Wang, K-L., C. Fawson, C. Barrett and J. McDonald, 2001, A flexible parametric GARCH model with an application to exchanges rates, Journal of Applied Econometrics 16, 521-536.

# Appendix

The key idea behind the efficient method of moment, which builds upon quasi maximum likelihood (QML) principles, is surprisingly simple: use the score of an auxiliary model (or score generator) evaluated under the structural model as the vector of moment conditions in order to calibrate the parameters of the structural model. The auxiliary model that generates the scores should approximate the actual distribution of the data closely, but it does not have to nest it. If it does, then one obtains ML efficiency. Furthermore, identification requires that the number of auxiliary model parameters is larger than those of the structural model.

To be more specific, suppose that the log likelihood function of the auxiliary model is $\frac{1}{N}\sum \ln f\left(r_t|X_t,\beta\right)$. This is not the true data generating process and the estimates, $\hat{\beta}$, may or may not be consistent. The data generating process is instead our structural model, parametrized by $\theta$, and we assume that there is a value $\theta^0$ such that the density of the observed data, $r_t$, is the same as of the simulated returns, $r_t\left(\theta^0\right)$. If we further assume the existence of a binding function, $\beta = b\left(\theta\right)$, we have that the unknown density of the structural model $p\left(r_t|X_t,\theta^0\right) = f\left(r_t\left(\theta^0\right)|X_t\left(\theta^0\right),\hat{\beta}\right)$. The binding function defines $\beta^0$, the quasi true vector, by $\beta^0 = b\left(\theta^0\right)$, from which it now follows that $p\lim\hat{\beta} = \beta^0$, the consistency result we need.

We can now simulate a time series of size $N$ from the structural model, denoted $\left\{r_t\left(\theta\right), X_t\left(\theta\right)\right\}$, in order to generate the moment conditions:

$$m_N\left(\theta,\hat{\beta}\right) = \frac{1}{N}\sum_{t=1}^{N}\frac{\partial \ln f\left(r_t\left(\theta\right)|X_t\left(\theta\right),\hat{\beta}\right)}{\beta}, \tag{24}$$

which converges to zero as $T \to \infty$ when $\theta = \theta^0$. This occurs because then $\hat{\beta}$ converges to $\beta^0$ which, in turn, is the QML estimate with first order condition given by (24). For other $\theta$, (24) will not converge to zero. It is essential that $N$ is large so that the Monte Carlo variance becomes negligible.

The EMM estimator of $\theta$ is defined by

$$\hat{\theta} = \arg\min_{\theta} m_N\left(\theta, \hat{\beta}\right)' W_T^{-1} m_N\left(\theta, \hat{\beta}\right), \tag{25}$$

where $W_T$ is a weighting matrix. Following GMM theory, the optimal choice of $W_T$ is a consistent estimator of the asymptotic covariance matrix of the scores. If the auxiliary model is a reasonable approximation of the data, $W_T$ is often estimated from the outer product gradient

$$W_T\left(\hat{\beta}\right) = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial \ln f\left(r_t|X_t, \hat{\beta}\right)}{\hat{\beta}} \frac{\partial \ln f\left(r_t|X_t, \hat{\beta}\right)}{\hat{\beta}'}. \tag{26}$$

Most conveniently, $W_T$ does not depend on the structural parameter vector $\theta$.

The estimated asymptotic covariance matrix of $\hat{\theta}$ is

$$Cov\left(\hat{\theta}\right) = \frac{1}{T}\left(\frac{\partial m_N\left(\hat{\theta}, \hat{\beta}\right)}{\theta} W_T^{-1}\left(\hat{\beta}\right) \frac{\partial m_N\left(\hat{\theta}, \hat{\beta}\right)}{\theta'}\right)^{-1} = \frac{1}{T}\left(M_\theta' W_T^{-1}\left(\hat{\beta}\right) M_\theta\right)^{-1}, \tag{27}$$

where the Jacobian $M_\theta$ in general must be computed numerically. A general specification test is also available. Under the null hypothesis that the structural model is true, $T$ times the minimized value of the EMM criterion function is asymptotically distributed as $\chi^2$ with degrees of freedom equal to the number of overidentifying restrictions:

$$T m_N\left(\hat{\theta}, \hat{\beta}\right)' W_T^{-1} m_N\left(\hat{\theta}, \hat{\beta}\right) \longrightarrow^d \chi^2_{\dim(\beta)-\dim(\theta)}. \tag{28}$$

# Earlier Working Papers:

For a complete list of Working Papers published by Sveriges Riksbank, see www.riksbank.se

Sveriges Riksbank
Visiting address: Brunkebergs torg 11
Mail address: se-103 37 Stockholm

Website: www.riksbank.se
Telephone: +46 8 787 00 00, Fax: +46 8 21 05 31
E-mail: registratorn@riksbank.se

ISSN 1402-9103     Riksbank Printing Office Stockholm 2005