Eur. Phys. J. B **55**, 183–187 (2007)
DOI: 10.1140/epjb/e2006-00385-4

THE EUROPEAN
PHYSICAL JOURNAL B

Empirical validation of stochastic models of interacting agents

A "maximally skewed" noise trader model

S. Alfarano[1,a], T. Lux[1], and F. Wagner[2]

[1] Department of Economics, University of Kiel, Olshausenstrasse 40, 24118 Kiel, Germany
[2] Department of Physics, University of Kiel, Leibnizstrasse 15, 24098 Kiel, Germany

Received 31 March 2006 / Received in final form 11 September 2006
Published online 20 October 2006 –

Abstract. The present paper expands on recent attempts at estimating the parameters of simple interacting-agent models of financial markets [S. Alfarano, T. Lux, F. Wagner, Computational Economics **26**, 19 (2005); S. Alfarano, T. Lux, F. Wagner, in *Funktionsfähigkeit und Stabilität von Finanzmärkten*, edited by W. Franz, H. Ramser, M. Stadler (Mohr Siebeck, Tübingen, 2005), pp. 241–254]. Here we provide additional evidence by (i) investigating a large sample of individual stocks from the Tokyo Stock Exchange, and (ii) comparing results from the baseline noise trader/fundamentalist model of [S. Alfarano, T. Lux, F. Wagner, Computational Economics **26**, 19 (2005)] with those obtained from an even simpler version with a preponderance of noise trader behaviour. As it turns out, this somewhat more parsimonious "maximally skewed" variant is often not rejected in favor of the more complex version. We also find that all stocks are dominated by noise trader behaviour irrespective of whether the data prefer the skewed or the baseline version of our model.

PACS. 89.65.Gh Economics; econophysics, financial markets, business and management

1 Introduction

After more than a decade of research on agent-based models in finance, estimation of the parameters of such models ranks high on the agenda of future research. However, while basic stylized facts of financial data can be captured to some extent by various microscopic models, estimation of their parameters has hardly been tried in the literature. A few exceptions include attempts at estimating models of chartist/fundamentalist interaction as regime-switching processes [1] and a heuristic estimation of the parameters of Kirman's seminal ant model [2] with financial data [3]. A very similar model has also been estimated by Alfarano et al. [4,5]. In contrast to Kirman [2] and Gilli and Winker [3], their model allows for asymmetric movements between the two groups of the population (labeled noise traders and fundamentalists). This generalization allows for a broad range of unimodal or bimodal outcomes of the distribution of the population together with almost any degree of skewness in the pdf for the stationary distribution of group occupation numbers. Depending on the underlying parameters, this model could give rise to dominance of either noise traders or fundamentalists on average within a particular market together but would also allow for intermittent deviations from the unconditional mean of the population configuration. Alfarano et al. [4] derive closed-form solutions for the distribution of returns of this model which enables them to estimate the parameters via a maximum likelihood fit for the unconditional distribution. The results allow insights into whether — within the framework of this particular model — the data suggest a dominance of noise traders or fundamentalists behaviour. As it turns out, noise traders seem to dominate in stock markets [4], while fundamentalists dominate in foreign exchange markets [5]. Here we expand on these studies by applying the same tools to a larger ensemble of stocks, namely 100 individual stocks for the Tokyo Stock Exchange with daily records extending from 1975 to 2001. As it turns out, the dominance of noise traders in stock markets is also confirmed for *each individual entry* in this data set. Inspired by this unequivocal finding, we also explore whether a more parsimonious version of our baseline model would be sufficient to capture the main features of the data. To this end, we develop an asymptotic benchmark of a "maximally skewed" process in which the tendency of agents to switch from the fundamentalist group to the noise trader group (parametrized by a Poissonian transition rate ε_1) by far dominates over movements in the opposite direction (parametrized by a Poissonian transition rate ε_2). We investigate how this extreme but simple model performs in comparison to our benchmark case explored in [4] and [5]. As it turns out the more parsimonious model with $\varepsilon_1 \gg \varepsilon_2$ cannot be rejected via likelihood ratio tests for 40% of the stocks in our sample.

[a] e-mail: alfarano@bwl.uni-kiel.de

2 The herding model and the artificial market

Our market is populated by a fixed number of traders N, belonging to two categories: (i) *fundamentalists*, who buy or sell according to the deviation between the actual price p and the fundamental value p_F and (ii) *noise traders* who are subject to "irrational" fads or moods as introduced in the seminal paper by De Long et al. [6]. The basic building block of our model is a herding mechanism among investors inspired by the recruitment-infection model introduced by Kirman [2]. We formalize this approach as a jump Markov process in continuous time. For sufficiently small time increments $\Delta\tau$, at most one agent will change its type. The resulting increase or decrease of group occupation numbers can be formalized by the following conditional transition probabilities $\rho(\cdot|\cdot)$:

$$\begin{aligned}\rho(n+1, t+\Delta\tau|n,t) &= (N-n)(a_1+bn)\Delta\tau,\\ \rho(n-1, t+\Delta\tau|n,t) &= n\big(a_2+b(N-n)\big)\Delta\tau,\end{aligned} \quad (1)$$

where n and $N-n$ are the respective numbers of noise traders and fundamentalists. The constant parameters a_1, a_2 and b specify the behaviour of the investors. The transition probabilities in equation (1) are composed of two terms: the first term, which is linear in the number of investors in the state, governs the idiosyncratic propensity to switch to the other strategy. The major difference to Kirman's model is that we allow for asymmetric transition rather than assuming a common constant a for the autonomous switching probabilities. Note that this enables us to generate dynamics with an arbitrary degree of asymmetry and dominance of one group over the other. The second term in the transition rates captures the herding tendency, since it is proportional to the product of the number of agents in the opposite state, i.e. $(N-n)$ or n. The constant b is a parameter for the strength of the herding behaviour. For finite but large N, the dynamics of the discrete variable n can be approximated by a Fokker-Planck equation expressed in terms of the intensive variable $z=\frac{n}{N}$, with drift and diffusion functions given by:

$$A(z) = a_1-(a_1+a_2)z \qquad \text{and} \qquad D(z) = 2b(1-z)z. \quad (2)$$

Introducing two simple behavioural rules for the noise traders' and fundamentalists' excess demand, we can relate the composition of the population of the investors in the market to the log-returns of the market price. Fundamentalists' excess demand is assumed to be given by:

$$ED_F = (N-n)\ln\frac{p_F}{p}, \quad (3)$$

which means that they react with a proportional increase of excess demand to deviations between the supposed fundamental value (p_F) and the current market price (p). Noise traders' aggregate excess demand takes the form:

$$ED_C = -r_0 n\xi, \quad (4)$$

where the stochastic variable ξ represents the current aggregate 'mood' of the noise traders, n is the number of noise traders and r_0 is a constant which is used for scaling their contribution to aggregate excess demand vis-à-vis the fundamental component. The negative sign on the right-hand side of equation (4) is simply chosen for notational convenience. Our noise traders are close in spirit to the original contribution of De Long et al. [6] in that they are characterized by random misperceptions rather than systematic trend following or other heuristic trading rules. Within a Walrasian scenario, we can compute the equilibrium price by setting total excess demand equal to zero, which yields to the following formula for the market price:

$$p = p_F \exp\left(r_0\frac{z}{1-z}\xi\right). \quad (5)$$

The variables z and $1-z$ in equation (5) are the fractions of the noise traders and fundamentalists among agents, respectively. While we concede that prevalence of a market clearing equilibrium might be too restrictive of an assumption for high-frequency data, we also conjecture that price adjustment speed is high enough to prevent large imbalances between demand and supply for daily data. Thus, we believe that the temporary equilibrium assumption of equation (5) is not too much at odds with empirical observations. Assuming for simplicity that the fundamental value does not change over time, the log-returns can be expressed as (see [4]):

$$r(t,\Delta t) = r_0\left(\frac{z(t+\Delta t)}{1-z(t+\Delta t)}\xi(t+\Delta t) - \frac{z(t)}{1-z(t)}\xi(t)\right). \quad (6)$$

Assuming that $\xi(t)$ changes much faster than the population configuration, this can be approximated by:

$$\begin{aligned} r(t,\Delta t) &= r_0\,\frac{z(t)}{1-z(t)}\left(\xi(t+\Delta t)-\xi(t)\right)\\ &= \sigma(t)\,\eta(t,\Delta t),\end{aligned} \quad (7)$$

where r_0 is a constant of proportionality, $\sigma(t)=\frac{z(t)}{1-z(t)}$ is the volatility (depending on the population composition) and $\eta(t,\Delta t)$ is an IID random variable with mean zero. Equation (7) turns out to be analytically tractable, providing us with closed-form solutions for a wide range of conditional and unconditional properties of returns (equilibrium distribution and autocorrelation function of absolute returns). The equilibrium distribution for the volatility σ is given by:

$$p(\sigma) = \frac{1}{r_0}\frac{1}{B(\varepsilon_1,\varepsilon_2)}\left(\frac{\sigma}{r_0}\right)^{\varepsilon_1-1}\left(\frac{r_0}{\sigma+r_0}\right)^{\varepsilon_1+\varepsilon_2}, \quad (8)$$

where $\varepsilon_{1,2}\equiv\frac{a_{1,2}}{b}$ and $B(\cdot,\cdot)$ is the Beta function. If we normalize the time series of returns according to the condition $E[|r_t|]=1$, the constant r_0 can be expressed as a function of the two behavioural parameters ε_1 and ε_2:

$$r_0 = \frac{\varepsilon_2-1}{\varepsilon_1}\frac{1}{E[|\eta|]}. \quad (9)$$

Under the previous normalization and in the limit $\varepsilon_1 \gg 1$, the probability density of equation (8) converges to a

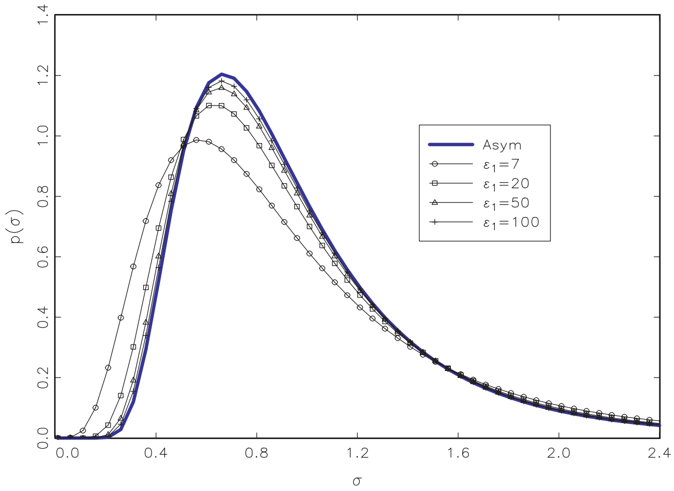

Fig. 1. Comparison of the equilibrium distribution of the volatility σ (Eq. (8)) for increasing values of the parameter ε_1 and its asymptotic distribution in the case $\varepsilon_1 \gg 1$ given by equation (10).

non-degenerate distribution:

$$p_\infty(\sigma) = \frac{1}{\Gamma(\varepsilon_2)} \frac{1}{k} \left(\frac{k}{\sigma}\right)^{\varepsilon_2+1} \cdot \exp\left(-\frac{k}{\sigma}\right), k \equiv \frac{\varepsilon_2 - 1}{E[|\eta|]}, \tag{10}$$

which is the inverse Gamma distribution. Interestingly, the inverse Gamma is the resulting stationary distribution of well-known stochastic volatility models introduced in the literature on financial mathematics [7,8]. The derivation can be found in Appendix A. Note that neglecting changes of the fundamental value in equation (6) can be justified by the dominance of the noise trading contribution, i.e. $p(\sigma)$, whose regular variation with index ε_2 would constitute the dominant component to the pdf of returns as long as changes of fundamental value would have tails decaying faster than with a power of ε_2 (which would be the case for Normally distributed changes of fundamentals and a broad range of other distribution functions). Figure 1 shows the convergence of the distribution (8) to its asymptotic form (10). Assuming that the noise η is uniformly distributed[1], it is also possible to compute the pdf of absolute returns, see [4], which takes the form:

$$p(|r|) = \frac{1}{r_0} \frac{\varepsilon_2}{\varepsilon_1 - 1} \left[1 - \beta\left(\frac{|r|}{|r| + r_0}\,; \varepsilon_1 - 1, \varepsilon_2 + 1\right)\right], \tag{12}$$

where $\beta(\cdot;\cdot,\cdot)$ is the incomplete Beta function. The limiting case $\varepsilon_1 \gg 1$ of equation (10) leads to the following parametric family for the distribution of absolute returns:

$$p_\infty(|r|) = \frac{1}{2} \frac{\varepsilon_2}{\varepsilon_2 - 1} \Gamma\left(\varepsilon_2 + 1, \frac{2(\varepsilon_2 - 1)}{|r|}\right), \tag{13}$$

where $\Gamma(\cdot,\cdot)$ is the incomplete Gamma function — see Appendix B.

The average percentage of noise traders in the market is $E[z] = \frac{\varepsilon_1}{\varepsilon_1+\varepsilon_2}$. Therefore, a very large value of the parameter ε_1 is an indication of a market that is mostly dominated by noise traders.

Table 1. Summary of the estimation results for the entire sample of 100 stocks. The entries for ε_∞ refer to the asymptotic version of the model based on equation (13). The last row reports the descriptive statistics of the estimates for which the likelihood ratio test does not reject the limiting case of the model at a p-value of 2.5%.

	Mean	Std	Median	Max	Min
ε_1	17.8	12.6	14.4	75	6.0
ε_2	4.3	0.7	4.2	6.4	2.9
ε_∞	3.5	0.3	3.5	4.2	2.7
$\varepsilon_{\infty,p}$	3.4	0.3	3.4	4.1	2.7

3 Estimation of the parameters

Given equations (12) and (13), we can estimate the parameters of the model, ε_1 and ε_2. We apply the Maximum Likelihood procedure as explained in [4]. The focus of our present estimation exercise is to test whether the model based on the two parameters in equation (12) performs better than its limiting asymptotic version in describing the unconditional distribution of financial returns. The aim is to see whether an even simpler model of interacting agents with an overwhelming tendency towards noise trading would work as well as our more elaborate baseline framework to describe the returns distribution. In order to do so, we estimate the parameters for a large ensemble of financial data, consisting of 100 stocks from the Japanese market at daily sampling intervals with a time horizon ranging from January 4th 1975 to December 28th 2001. Table 1 summarizes the results of our estimation exercise and Figure 2 illustrates the empirical distribution of the estimates across the whole pool of stocks. The ML estimation procedure for the case of two parameters, (ε_1 and ε_2) is applied with the constraint $\varepsilon_1 \le 100$. For 6 out of the entire sample of 100 stocks, the previous constraint is binding. The descriptive statistics in Table 1 refer to the sample without those 6 cases. Using likelihood ratio tests, we can compare the relative performance of the two-parameter model and its limiting version for $\varepsilon_1 \gg 1$. It turns out that in 40 out of 100 cases[2], the asymptotic version cannot be rejected. Interestingly, even in the case of non-rejection of the two-parameter model, the condition $\varepsilon_1 > \varepsilon_2$ holds for all the considered time series. These results confirm and extend our previous finding [4] of stock markets being dominated by noise traders. The condition $\varepsilon_1 > \varepsilon_2$ seems to be a salient feature of stock market

[1] Under this assumption the normalization factors, k and r_0, are:

$$r_0 = \frac{2(\varepsilon_2 - 1)}{\varepsilon_1}\,, \qquad k = 2(\varepsilon_2 - 1)\,. \tag{11}$$

[2] The p-value of the test is fixed at 2.5%.

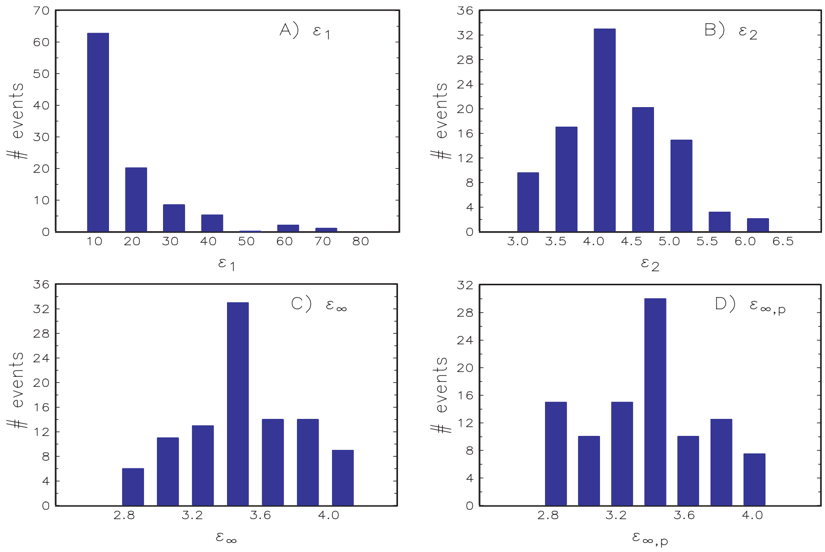

Fig. 2. The four panels show the histograms of the estimates. Panel (A) and (B) refer to the parameters ε_1 and ε_2 respectively, without the 6 cases of binding constraints in the ML procedure. Panel (C) shows the histogram for the estimates of the parameter ε_2 in the limiting case $\varepsilon_1 \gg 1$ as given by equation (13) and labeled as ε_∞. Panel (D) shows the distribution of the estimates of parameter ε_2 for those cases for which the asymptotic model cannot be rejected at a p-value of 2.5%, labeled as $\varepsilon_{\infty,p}$.

data since it emerges as a robust feature of all applications of our model. Interestingly, using exchange rates instead of stock market data, we observe the opposite case, namely a market dominated by fundamentalists as in [5]. We might, therefore, conclude that the model can capture some underlying difference in the two categories of financial time series. Within the framework of our model, the higher degree of fluctuations of stock markets vis-à-vis foreign exchange markets would have to be attributed to a dominance of noise traders' activity in the former. It might seem natural that the higher volatility (in terms of the relative range of fluctuations) of stock markets against foreign exchange markets is attributed to a larger number of noise traders in the former within a behavioural finance setting.

However, note that different measures of volatility exist and stock and foreign exchange markets appear equally volatile under many of them (e.g. tail index estimates). Nevertheless, the naked eye sees a lower range of fluctuations in the foreign exchange markets as illustrated in [5]. This translates into relatively subtle differences of distributions, for which no clear concept of measurement exists so far, but which are apparently detected by our model. Hence, while we have an intuitive explanation of the different findings for both types of markets, the details of how this explanation is borne out by our estimates are far from trivial.

4 Conclusion

In the present paper, we have estimated the parameters of a very simple interacting-agents model of financial market dynamics. As it turns out, the Japanese stock market data unanimously speak in favor of the dominance of noise traders in the pertinent time series. In a large number of cases, we were not even able to reject the extreme version of the model with a dominant tendency of adherence to the noise trader group among agents. One would certainly like to investigate how robust these results are by varying the details of the traders' behaviour and the design of the market process. Unfortunately, there is not much hope that one might arrive at similar 'nice' closed-form solutions with more complicated models. Furthermore, while we think that these results are interesting as a first approach towards estimation of agent-based models, estimation based on an unconditional distribution is certainly not fully satisfactory. As the next step, one should try to estimate the parameters via full maximum likelihood, taking into consideration the underlying Markov process resulting from agents' strategy choices. This would enable one to estimate probabilities for occupation numbers from which forecasts of volatility could be computed.

We thank the comments presented by two anonymous referees. Financial support by the European Commission under STREP Contract No. 516446 is gratefully acknowledge.

Appendix A: Asymptotic distribution of the volatility for $\varepsilon_1 \gg 1$

We start by rearranging the components of equation (8):

$$p(\sigma) = \frac{\Gamma(\varepsilon_1+\varepsilon_2)}{\Gamma(\varepsilon_1)\ \varepsilon_1^{\varepsilon_2}} \frac{k^{\varepsilon_2}}{\Gamma(\varepsilon_2)} \frac{1}{\left(\frac{k}{\varepsilon_1\sigma}+1\right)^{\varepsilon_1}} \cdot \frac{1}{\left(\sigma+\frac{k}{\varepsilon_1}\right)^{\varepsilon_2}} \cdot \frac{1}{\sigma}. \quad \text{(A.1)}$$

Under the condition $\varepsilon_1 \gg 1$, the first term converges to 1, the third term converges to $\exp(-\frac{k}{\sigma})$ and the forth term converges to $\sigma^{-\varepsilon_2}$. Therefore the overall distribution converges to the expression given by equation (10).

Appendix B: Asymptotic distribution of the absolute returns for $\varepsilon_1 \gg 1$

The unconditional distribution of the absolute returns, $|r| = v$, can be computed by the following formula:

$$p_\infty(v) = \int_0^\infty p\,(v \mid \sigma)\, p_\infty(\sigma)\; d\sigma, \tag{B.1}$$

where the conditional probability $p(v|\sigma)$ takes the form:

$$p(v|\sigma) = \begin{cases} 0 \text{ if } & v > \sigma \\ \frac{1}{\sigma} \text{ if } & v < \sigma, \end{cases} \tag{B.2}$$

due to the uniform distribution of the variable η. Plugging equation (10) into equation (B.1) yields:

$$p_\infty(v) = \frac{1}{\Gamma(\varepsilon_2)} \frac{1}{k^2} \int_v^\infty \left(\frac{k}{\sigma}\right)^{\varepsilon_2+2} \cdot e^{-\frac{k}{\sigma}}\; d\sigma. \tag{B.3}$$

Under the change of variable $u = \frac{k}{\sigma}$, the previous integral takes the form:

$$p_\infty(v) = \frac{1}{\Gamma(\varepsilon_2)} \frac{1}{k} \int_0^{\frac{k}{v}} u^{\varepsilon_2}\; e^{-u}\; du. \tag{B.4}$$

Equation (13) is then obtained using the definition of the incomplete Gamma function:

$$\Gamma(\varepsilon, x) = \frac{1}{\Gamma(\varepsilon)} \int_0^x u^{\varepsilon-1} e^{-u} du. \tag{B.5}$$

Note that equation (13) can also be derived as a limiting case of equation (12) for $\varepsilon_1 \gg 1$.

Appendix C: Data description

In the following we provide the list of identification numbers of the 100 stocks used in the paper:

T1301 T1332 T1333 T1351 T1352 T1501 T1503 T1515 T1518 T1601 T1603 T1661 T1701 T1801 T1802 T1804 T1806 T1808 T1810 T1812 T1813 T1814 T1815 T1816 T1817 T1818 T1819 T1820 T1821 T1822 T1823 T1824 T1826 T1827 T1829 T1833 T1834 T1835 T1836 T1837 T1838 T1839 T1847 T1851 T1852 T1855 T1858 T1860 T1861 T1863 T1865 T1866 T1881 T1882 T1883 T1884 T1885 T1886 T1888 T1889 T1890 T1893 T1895 T1896 T1898 T1917 T1920 T1921 T1922 T1923 T1924 T1925 T1926 T1928 T1941 T1942 T1943 T1945 T1948 T1949 T1950 T1951 T1954 T1955 T1957 T1958 T1959 T1960 T1961 T1962 T1963 T1967 T1968 T1969 T1970 T1971 T1972 T1974 T1975 T1979.

References

1. R. Vigfusson, *International Journal of Finance and Economics* **2**, 291 (1997)
2. A. Kirman, *Quarterly Journal of Economics* **108**, 137 (1993)
3. M. Gilli, P. Winker, Computational Statistics and Data Analysis **42**, 299 (2003)
4. S. Alfarano, T. Lux, F. Wagner, Computational Economics **26**, 19 (2005)
5. S. Alfarano, T. Lux, F. Wagner, in *Funktionsfähigkeit und Stabilität von Finanzmärkten*, edited by W. Franz, H. Ramser, M. Stadler (Mohr Siebeck, Tübingen, 2005), pp. 241–254
6. B.J. De Long, A. Shleifer, L.H. Summers, R.J. Waldmann, Journal of Political Economy **98**, 703, 738 (1990)
7. A.L. Lewis, *Option Valuation under Stochastic Volatility* (Finance Press, California, 2000)
8. D. Ahn, B. Gao, Review of Financial Studies **12**, 721 (1999)