Journal of Empirical Finance 16 (2009) 804–815

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/econbase

Applying the method of simulated moments to estimate a small agent-based asset pricing model[☆]

Reiner Franke

University of Kiel, Germany

ARTICLE INFO

Article history:
Received 9 December 2008
Received in revised form 16 June 2009
Accepted 18 June 2009
Available online 25 June 2009

JEL classification:
C14
C15
F31

Keywords:
Simulation-based estimation
Daily returns
Autocorrelation patterns
Fundamentalists and speculators

ABSTRACT

The paper takes a recent agent-based asset pricing model by Manzan and Westerhoff from the literature and applies the method of simulated moments to estimate its six parameters. In selecting the moments, the focus is on the fat tails and autocorrelation patterns of the daily returns of several stock market indices and foreign exchange rates. It is argued that it may be meaningful to abandon the econometrically optimal weighting matrix in the objective function and instead invoke the moments' *t*-statistics in an intuitively appealing way. This modification gives rise to estimations whose moment matching, given the model's parsimony, can be largely considered to be satisfactory. Also the parameter estimates across different markets make good economic sense.

1. Introduction

Behavioural asset pricing models often include latent variables in a nonlinear framework.[1] An estimation by maximum likelihood or the generalized method of moments becomes infeasible then since the likelihood function and the moment restrictions do not have analytically tractable forms in terms of the unknown parameters. Instead of suitable approximations of these models that might simplify the objective functions, there are now several simulation-based methods available that can serve as an alternative approach. The present paper is concerned with the method of simulated moments (MSM) to test a small model of speculative dynamics from the literature.[2] In brief, this approach seeks to identify numerical parameter values such that several summary statistics of interest, or ‘moments’, that are computed from the simulations of the model come close to their empirical counterparts. It thus amounts to what has also been called ‘empirical validation’ of a model (in a narrower sense; cf. Fagiolo et al., 2007, pp. 190f).

MSM is frequently criticized because of its arbitrary choice of moments in the criterion function. For this reason methods have been developed, such as indirect inference or the efficient method of moments, where in effect a set of moment conditions is

☆ I wish to thank Dick van Dijk, Thomas Lux, Frank Westerhoff as well as three anonymous referees for their many helpful comments. Financial support from EU STREP Complex Markets, contract number 516446, is gratefully acknowledged.
E-mail address: franke@iksf.uni-bremen.de.

[1] For recent surveys of the burgeoning field of research on agent-based models, see Chiarella et al. (2009), Hommes and Wagener (2009), Lux (in press) and Westerhoff (in press), among others.

[2] Opinions seem to be divided over calling this method the ‘simulated method of moments’ or the ‘method of simulated moments’ (at least if there is no conflict with the acronym of Markov switching models).

0927-5398/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jempfin.2009.06.006

endogenously defined by means of an auxiliary model.[3] It may, however, be argued that this does not completely resolve the problem of arbitrariness, or judgement; it is only shifted from the choice of moments to the choice of the auxiliary model. Taking into account that structural models are necessarily misspecified, one may furthermore prefer MSM to the other methods because of its higher transparency. After all, a model is not meant to resemble or mimic any actual economy or financial market, but the best we can expect from it is that it matches *some* of the 'stylized facts'. While MSM requires the researcher to make up his or her mind about the dimensions along which the model should be most realistic, the implications of a decision for a particular auxiliary model are usually less clear. Besides the fact that in certain cases MSM is possibly faster or more robust than other methods, we consider its explicitness and easy interpretation as the main argument to try this estimation procedure (in the sense that we prefer to try it first, which does by no means rule out the subsequent application of other methods).

Regarding the estimation of agent-based asset pricing models, it seems that so far MSM has only been applied by a research group around M. Gilli and P. Winker.[4] The latest and most elaborate paper is Winker et al. (2007). In their work the authors are, however, mainly taken up by their objective function and do not endeavour to make use of the existing econometric theory on MSM, from which one seeks to get information on the precision of the parameter estimates and the goodness-of-fit of the selected moments in the objective function (and possibly of other moments that were not included). It is one intention of the present paper to discuss these points in detail.

Another aspect is that the structural models to which Winker et al. apply MSM are as yet not well suited to reproduce the stylized facts of the daily returns on the financial markets (at least not with the authors' exogenous setting of the other model parameters that were not estimated). Hence the major economic outcome of the analysis is a clear rejection of the models, which in the end does not seem too satisfactory. By contrast, with Manzan and Westerhoff (2005) we choose an agent-based model from the literature that, although it is small and highly stylized, seems more promising concerning the properties of its returns.[5]

While it will be no great surprise that this model, too, is rejected by the standards of econometric theory, we are able to establish that the mismatch of the moments is not overly dramatic. For a more detailed study, we then abandon the econometrically optimal weighting matrix and invoke the *t*-statistics of the moments to modify the estimation's objective function in an intuitively appealing way. In consequence, the new estimations yield a moment matching that, given the model's parsimony, we consider to be satisfactory. We have thus a reasonable basis to judge the explanatory power of the model. Accepting the limitation that our moments only include returns but not the level of the time-varying fundamental value, it turns out that the role of the fundamentalists is minimal, unless dispensable. By contrast, the switching mechanism of the group of the so-called speculators is fully confirmed. Having this established, we subsequently estimate the model on several different stock market indices and foreign exchange rates. These results will give further support to the switching mechanism since we find that the corresponding parameter estimates are fairly robust and where there are greater differences, they make economic sense.

The remainder of the paper is organized as follows. The next section recapitulates the basic econometric facts of the method of simulated moments. Section 3 presents the Manzan–Westerhoff model, selects the moment functions, simulates the model, and then re-estimates it on the artificial data. The controlled experiment gives us some hints where there are problems and what at best could be expected from empirical estimations of this model. Section 4 turns to the real-world data of daily returns. Beginning with the Standard and Poor's 500 stock market index as our reference series, we discuss the merits of three different minimization criteria for the estimations and justify our decision for one of them which, as just indicated, is based on the *t*-statistics of the moments. Subsequently, we present and compare the estimation results for a few additional stock market indices and foreign exchange rates. Section 5 concludes.

2. The method of simulated moments

2.1. The objective function

In confronting the model with the empirical data, MSM focusses on a set of n_m descriptive statistics.[6] To begin with the empirical side, let L be the greatest lag involved in forming these statistics, y_t^{emp} a vector of (stationary) empirical variables observed in period t (whose dimension does not matter), and suppose observations are available over a time span $t=1-L,\ldots,1,\ldots,T$. The summary statistics derive from n_m moment functions $m_i(\cdot)$, which are generally defined on L-stretches of a variable y_t. Writing more compactly $z_t := (y_t, y_{t-1},\ldots, y_{t-L})$, the ith empirical moment is computed as the time average

$$\widehat{m}_{T,i} := (1/T)\sum_{t=1}^{T} m_i(z_t^{\text{emp}}),\ i = 1,\ldots,n_m \tag{1}$$

The index T makes the effective length of the sample period explicit. The caret over 'm' distinguishes the summary statistic from a single contribution to period t and also indicates that it is only an estimate of the true unconditional moment of the real-world stochastic process (as the actual sequence $\{y_t^{\text{emp}}\}_{t=1-L}^{T}$ is just one realization of it).

[3] Carrasco and Florens (2002) provide a succinct overview of MSM and the other two methods just mentioned.
[4] For a recent overview of the application of this and other estimation methods to asset pricing models see Chen et al. (2008; especially Table 3 on p. 22).
[5] In another paper, Manzan and Westerhoff (2007) estimate their model on monthly foreign exchange rate data. There the authors can basically get along with OLS since they specify the fundamental value by invoking the Purchasing Power Parity. With the daily changes considered in their earlier paper, however, the fundamental value is a more volatile, and unobservable, variable.
[6] The following exposition is based on Lee and Ingram (1991) and Duffie and Singleton (1993), which are two standard references for MSM.

Turning to the structural model under study, let $\boldsymbol{\theta}$ be the $n_p \times 1$ vector of the parameters that are to be estimated. Given numerical values for its components, an extended simulation of the model is run, such that after discarding a sufficiently long initial history to wash out transient effects we have model-generated observations over a horizon of sT periods, where s is some number greater than one. Possibly besides several other variables that one may not be interested in or that may not be observable in the real world, for each set of parameters $\boldsymbol{\theta}$ a unique series $\{y_t(\boldsymbol{\theta})\}_{t=1-L}^{sT}$ is produced in this way.

All simulations are required to employ the same sequence of pseudo-random numbers for the standardized stochastic terms in the model.[7] Hence the differences between two series $\{y_t(\boldsymbol{\theta}^1)\}$ and $\{y_t(\boldsymbol{\theta}^2)\}$ can be attributed to the differences in the parameters $\boldsymbol{\theta}^1$ and $\boldsymbol{\theta}^2$, and not to the random noise from the stochastic part of the model. As the artificial data $y_t(\boldsymbol{\theta})$, or $z_t(\boldsymbol{\theta})$ for that matter, are the counterparts of the empirical data y_t^{emp} and z_t^{emp}, it makes sense to compute the $n_m \times 1$ vector of the simulated moments,

$$\widehat{\mu}_{sT}(\theta) := (1/sT) \sum_{t=1}^{sT} m[z_t(\theta)] \tag{2}$$

Here the caret over 'μ' indicates that Eq. (2) is an estimate of the expected value $\mu(\boldsymbol{\theta}) := E_{\boldsymbol{\theta}}[m(z)]$, which are the model's unconditional moments with respect to $\boldsymbol{\theta}$. In most structural models, especially if they contain some nonlinearities, the latter will be analytically untractable, so that one has to resort to the estimates $\widehat{\mu}_{sT}$.

Ideally, by a suitable choice of $\boldsymbol{\theta}$, the expected moments of the model come close to the moments of the stochastic process that rules the real world. As both sets of moments are unknown, a comparison has to replace them with the sample counterparts $\widehat{\mu}_{sT}(\boldsymbol{\theta})$ and $\widehat{\boldsymbol{m}}_T$, respectively. Accordingly, a vector $\boldsymbol{\theta}$ has to be found that minimizes the distance between the two descriptive statistics. To formalize this idea, let $\widehat{\boldsymbol{m}}_T$ be the $n_m \times 1$ vector of the empirical moments $\widehat{\boldsymbol{m}}_{T,i}$ in Eq. (1) and define

$$\boldsymbol{g}_{s,T}(\theta) := \widehat{\mu}_{sT}(\theta) - \widehat{\boldsymbol{m}}_T \tag{3}$$

Furthermore, make use of an $n_m \times n_m$ weighting matrix $\mathbf{W}$ (which is to be discussed in a moment). The minimization problem then reads,

$$\mathbf{Q}_{s,T}(\theta) := \boldsymbol{g}_{s,T}(\theta)' \mathbf{W} \boldsymbol{g}_{s,T}(\theta) = \min_{\theta}! \tag{4}$$

and its solution yields the estimate $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}_{s,T}$ for the parameter vector (throughout, the prime denotes transposition of vectors and matrices).

2.2. The weighting matrix

The weighting matrix $\mathbf{W}$ in the objective function (Eq. (4)) is optimal if it provides the smallest asymptotic covariance for the estimator. Under these circumstances it is also ensured that the solution $\hat{\boldsymbol{\theta}}_{s,T}$ is a consistent estimate of the pseudo-true value $\boldsymbol{\theta}^o$ of the model's parameters.

In the following it is presupposed that the number of moments is not less than the number of parameters, i.e. $n_m \geq n_p$. To obtain the optimal weighting matrix we have to have recourse to an estimate of the long-run covariance matrix $\hat{\boldsymbol{\Omega}}$ of the empirical moments. It derives from the $n_m \times n_m$ sample autocovariance matrices for lags $j \geq 0$,

$$\Gamma_j := (1/T) \sum_{t=j+1}^{T} [m(z_t^{\mathrm{emp}}) - \widehat{\boldsymbol{m}}_T][m(z_{t-j}^{\mathrm{emp}}) - \widehat{\boldsymbol{m}}_T]'.$$

Typically, $\widehat{\Omega}_{sT}$ is a kernel-weighted average of these matrices, truncated at a suitable lag p. Using the Bartlett kernel with its linearly declining weights gives us the popular Newey–West estimator,

$$\widehat{\Omega} = \Gamma_0 + \sum_{j=1}^{p} \left(1 - \frac{j}{p+1}\right)(\Gamma_j + \Gamma'_j) \tag{5}$$

(which is heteroskedasticity and autocorrelation consistent, HAC; cf. Davidson and MacKinnon, 2004, pp. 362ff). Several choices for the bandwidth p are discussed in the literature, which are either rules of thumb or quite complicated expressions. We content ourselves with following a common practice that specifies p as the smallest integer greater than or equal to $(T^{1/4})$ (see Greene, 2002, p. 267, fn 10).

[7] If, for example, a term ε_t in the model is normally distributed as $N(\mu,\sigma^2)$ and μ and/or σ are among the parameters to be estimated, the expression 'standardized' means that for each simulation run at time t the same random number $\tilde{\varepsilon}_t$ is drawn from the standard normal distribution $N(0,1)$, and ε_t is set as $\varepsilon_t = \mu + \sigma\tilde{\varepsilon}_t$.

If in addition a number of standard regularity conditions are taken for granted, an optimal choice of the weighting matrix in Eq. (4), which we denote as $\mathbf{W}=\mathbf{W}^{o}$, is given by[8]

$$\mathbf{W}^{o} = \widehat{\Omega}^{-1} \tag{6}$$

Neglecting the covariances, a more straightforward idea of a weighting matrix is a diagonal matrix $\mathbf{W}=\mathbf{W}^{d}$.[9] It takes direct account of the fact that some empirical moments may be less precisely estimated than others, from which it is concluded that a minimization procedure should pay less attention to the errors from the 'imprecise' moments. Accordingly, the squared deviations of the simulated from the empirical moments are weighted by the reciprocal of the empirical variances. The weighting matrix $\mathbf{W}^{d}$ is thus given by the elements

$$w_{ij}^{d} = \begin{cases} \dfrac{1}{(1/T)\sum_{t=1}^{T}[m_i(z_t^{emp})-\hat{\boldsymbol{m}}_{T,i}]^2} & \text{if } i=j \\ 0 & \text{if } i \neq j \end{cases} \quad (i,j = 1,\ldots n_m). \tag{7}$$

Employing this matrix in the objective function (Eq. (4)) may equivalently be conceived as minimizing the sum of the squared t-statistics of the single moments (which will be shown shortly below).

2.3. The basic test statistics

Given that normally there are more moments than parameters to be estimated, the so-called J test is the standard statistic to test these overidentifying restrictions. It uses the fact that asymptotically the optimal value $\mathbf{Q}_{s,T}$ from Eq. (4) is distributed as a chi-square random variable with $(n_m - n_p)$ degrees of freedom. Specifically,

$$J := sT\mathrm{Q}_{s,T}(\widehat{\theta}_{s,\boldsymbol{T}})\,/\,(1+\boldsymbol{s}) \overset{\text{as.}}{\approx} \chi^2(n_m - n_p) \tag{8}$$

('as' stands for asymptotically, meaning for s fixed and T tending to infinity). A model is considered to be valid with regard to the chosen moments if J is less than the value of $\chi^2(n_m - n_p)$ corresponding to, say, a 95% significance level. If the J-statistic exceeds that critical value this can, but need not necessarily, be an indication that the model fails to mimic the empirical data in several dimensions. The problem that might generally arise is that the finite-sample distribution of the test statistic can happen to differ substantially from its asymptotic distribution, especially if, as in the case of the optimal weighting matrix $\mathbf{W}=\mathbf{W}^{o}$, a HAC covariance matrix is involved (Davidson and MacKinnon, 2004, p. 368).[10]

To locate the potential failures of a model one may use the distribution of the moment vector, again the asymptotic one. It is given by (see Duffie and Singleton, 1993, p. 945)

$$\widehat{\mu}_{sT}(\theta^{o}) - \widehat{\boldsymbol{m}}_T \overset{\text{as.}}{\approx} N(0,\mathrm{U}),\ \ \mathrm{U} := \frac{1+1/s}{T}\widehat{\Omega}. \tag{9}$$

Therefore, referring to $\widehat{\mu}_{sT}(\hat{\boldsymbol{\theta}}_{s,T})$ instead of $\widehat{\mu}_{sT}(\boldsymbol{\theta}^{o})$, a good hint for an evaluation of the model's ability to reproduce the ith empirical moment is its t-statistic. Denoting this measure by t_i^{o} in order to make explicit that it is based on the covariance matrix $\hat{\boldsymbol{\Omega}}$ in Eq. (5) and thus compatible with the optimal weighting matrix $\mathbf{W}^{o}$, and now also omitting the time indices s and T, we have

$$t_i^{o} = [\widehat{\mu}_i(\widehat{\theta}) - \widehat{\boldsymbol{m}}_i]\,/\,\sqrt{u_{ii}};\ \ u_{ii}\ \text{from}\,(9),\ i = 1,\ldots n_m. \tag{10}$$

For the discussion in Section 4 it proves useful to consider, in addition, the t-statistics associated with the diagonal weighting matrix $\mathbf{W}^{d}$ from Eq. (7). In this case the matrix $\mathbf{U}$ in Eq. (9) is determined by $\mathbf{U}=(1+1/s)(\mathbf{W}^{d})^{-1}/T$ and the squared t-statistics result as

$$(t_i)^2 = \frac{sT}{1+s}\Big[\frac{\widehat{\mu}_i(\widehat{\theta}) - \widehat{\boldsymbol{m}}_i]^2}{(1/T)\sum_{t=1}^{T}[m_i(z_t^{emp}) - \widehat{\boldsymbol{m}}_{T,i}]^2} \tag{11}$$

Note that in this alternative setting the J-statistic in Eq. (8) coincides with $\Sigma_i(t_i)^2$. Hence, as mentioned above, a set of parameters minimizing the sum of the squared t-statistics of (11) also minimizes the objective function $\mathbf{Q}_{s,T}$ in Eq. (4).

[8] Lee and Ingram (1991, p. 202) include the factor $(1+1/s)$ in the specification of the weighting matrix, i.e., they put $\mathbf{W}=[(1+1/s)\hat{\boldsymbol{\Omega}}]^{-1}$. The influence of the sampling uncertainty on the statistical fit is, however, clearer if this factor is here omitted and instead made explicit in Eqs. (8) and (13).

[9] There are also applications of MSM that adopt a diagonal weighting matrix because a complete covariance matrix would be ill-conditioned. For example, this problem arises when one seeks to match the model-generated impulse-response functions with those of a vector autoregression; see Boivin and Giannoni (2006, p. 453, fn 43).

[10] Two elaborate studies on this subject are Kocherlakota (1990) and Mao (1990).

To assess the accuracy of the parameter estimates, recall that $\mathbf{g}_{s,T}$ is the vector of the deviations of the simulated from the estimated moments in Eq. (3) and define the $n_m \times n_p$ matrix $\mathbf{B}$ by its partial derivatives with respect to the parameter changes,

$$b_{ij} := \partial g_{s,T}(\hat{\theta}_{s,T})_i / \partial\theta_j \quad (1 \le i \le n_m, 1 \le j \le n_p) \tag{12}$$

(these values have to be computed numerically). The estimated parameters are asymptotically normally distributed around their pseudo-true values with an $n_m \times n_p$ covariance matrix $\mathbf{V}$ that is generally a complicated expression involving $\mathbf{B}$ and the weighting matrix (see Lee and Ingram, 1991, p. 202). It simplifies if reference is made to the optimal weighting matrix $\mathbf{W}^o$ from Eq. (6). We then obtain

$$\hat{\theta}_{s,T} - \theta^o \overset{\text{as.}}{\sim} N(0, \mathbf{V}), \; V = \frac{1 + 1/s}{T} (\mathbf{B}'\mathbf{W}^o\mathbf{B})^{-1} \tag{13}$$

That is, if T is sufficiently large, the expressions $\sqrt{v_{ii}}$ provide a decent idea of the standard error that is associated with the estimation of the parameter $\boldsymbol{\theta}_i$.

The term $1 + 1/s$ in the covariance matrix is a (direct) measure of the sampling uncertainty in the simulations. Infinitely large s would correspond to a GMM estimation, where one can work with the model's population means $\mu(\boldsymbol{\theta})$. However, the loss in efficiency that results from estimating them with $\hat{\mu}_{sT}(\boldsymbol{\theta})$ is limited: for $s = 5$ or $s = 10$ it is only 20% or 10% relative to the GMM benchmark.

The covariance matrix $\mathbf{V}$ also plays a role if a Wald test is employed. In this way it can be checked if the model would be able to reproduce additional moments that were not included in the original estimation. The details of this test and its application in our estimations are discussed in an extended version of this paper, Franke (2008).

3. The Manzan–Westerhoff model and a first test of MSM

3.1. Formulation of the model

On a market for a risky asset, the model by Manzan and Westerhoff (2005) distinguishes fundamentalist traders and so-called speculators. The daily market orders of the former are inversely related to the percentage deviations of the price from the fundamental value. Speculators form expectations of tomorrow's price and buy (sell) if it exceeds (falls short of) the current price. The inventive feature of the model is that these expectations are based on the daily random walk changes η_t of the (log of the) fundamental value, such that the expected price is higher (lower) than the current price if the fundamental value has increased (decreased). Quantitatively the perceptions of the news are, however, biased. In the determination of the expected price more than η_t is added to the current (log) price, and in this sense the speculators overreact to the news, if the historical volatility of the asset exceeds a certain threshold. Otherwise they underreact, which means less than η_t is added to the current price. The volatility itself is specified as the mean of the absolute returns over a rolling sample period ('history') of length H. So at day t it is given by

$$\mathbf{V}_t(H) := (1/H) \sum_{\tau=1}^{H} |S_{t-\tau+1} - S_{t-\tau}| \tag{14}$$

where S_t denotes the log prices. A market maker is supposed to change the price in proportion to the sum of the demand of fundamentalists and speculators. In this way the dynamics of the log price can be reduced to the equation,

$$S_{t+1} = S_t + x(F_t - S_t) + \begin{cases} y\eta_t & \text{if } \mathbf{V}_t(H) \ge K \\ z\eta_t & \text{else} \end{cases} \quad \eta_t \sim N(0, \sigma_\eta^2). \tag{15}$$

Here x, y, z are positive coefficients that are composed of the structural parameters of the model (with $y > z$), K is the value of the threshold just mentioned, F_t is the log of the fundamental value, and the stochastic innovations η_t are normally distributed with standard deviation σ_η (i.i.d., of course). As indicated, the latter are just the changes in the random walk of F,

$$F_t = F_{t-1} + \eta_t. \tag{16}$$

In sum, the Manzan–Westerhoff (MW) model is compactly described by the three equations (Eqs. (14)–(16)). Obviously, if (represented by the parameter x) the direct influence of the fundamentalists is low, as it will turn out to be, the dynamics of the returns $S_t - S_{t-1}$ from Eq. (15) essentially amounts to a switching process between two normal distributions (akin to a reformulated GARCH-in-mean model). While this would be a technical description, Manzan and Westerhoff (2005) motivate it by a little, experimentally supported story of investor psychology, which especially allows them to speak of the speculators' responsiveness to news (the parameters y and z), a time horizon, or memory, for the relevant returns (H), and a threshold (K) for the under- and overreactions.

Table 1
Re-estimation of a model-generated return series ($T=6760$).

	ϕ	y	z	ρ	J^o
1	0.100	1.350	0.650	0.750	8.14
2	2.138	1.352	0.650	0.742	2.57
	(2.554)	(0.081)	(0.032)	(0.052)	
3	0.100	1.356	0.655	0.745	2.84
	–	(0.047)	(0.020)	(0.031)	
4	10.000	1.360	0.644	0.740	3.00
	–	(0.088)	(0.044)	(0.063)	
5	0.000	1.353	0.653	0.744	2.80
	–	(0.042)	(0.020)	(0.029)	

Note: J^o is the J test statistics determined by Eqs. (4) and (8) with weighting matrix $\mathbf{W}=\mathbf{W}^o$ from Eq. (13) and $s=10$. The parameters $\sigma_\eta=0.010$ and $H=10$ are fixed at their true values, the other true values are given in the first row, with $\phi:=100\cdot x$ and $\rho:=100\cdot K$. Numbers in parentheses are the standard errors from Eq. (13), while a long hyphen indicates that ϕ was exogenously fixed in that estimation.

3.2. Setting up the moment functions

Agent-based models are built to explain certain stylized facts of financial markets. There are four stylized facts that, referring to returns, have received the most intensive attention, namely: absence of autocorrelations, volatility clustering, long memory, and fat tails (see Chen et al., 2008, p. 19). The choice of the moments for our application of MSM will be quite in line with this concern. So let us express returns in percent and denote them as

$$r_t := 100\cdot(S_t - S_{t-1}),\ v_t := |r_t|. \tag{17}$$

Squared returns, which are also often referred to, behave very similar to the absolute returns and so can be neglected here. The first moments the model should be able to match are the mean and variance of r_t and v_t. In setting up the moment functions that are to capture the autocorrelations, reference is made to the autocovariances (ACV). To this end we first compute the mean values of the empirical data r_t^{emp} and v_t^{emp} over their time horizon T. Designating them $\bar{r}^{emp}$ and $\bar{v}^{emp}$, respectively, and writing $z_t=(r_t,r_{t-1},\ldots,r_{t-100}, v_t,v_{t-1},\ldots,v_{t-100})$ for the relevant L-stretch of the observations for day t, we may generally consider the following functions $m_i=m_i(z_t)$ (for suitable indices i when stacking some of them in the moment vector):

$$\begin{array}{lll} \underline{q\,mean} & : \; m_i(z_t) = q_t & (q=r,v) \\ \underline{q\,ACV\,k} & : \; m_i(z_t) = (q_t - \bar{q}^{emp})(q_{t-k} - \bar{q}^{emp}) & (q=r,v;\, k=0,\ldots,100) \end{array}. \tag{18}$$

Certainly, 'q ACV 0' represents the variance of $q=r,v$ and 'q ACV k' the autocovariances with lag k. Note that also the model-generated series of r_t and v_t will refer to the empirical means in Eq. (18). If instead the simulated means were used, the entire time series, and not only its period-t contributions, would enter the definition. It would then read $m_i=m_i[z_t(\boldsymbol{\theta}), \{z_t(\boldsymbol{\theta})\}_{t=1}^{sT}]$, which slightly deviates from the econometric setting above. Though one will expect that normally the propositions of MSM continue to hold true in this case, this would still need to be demonstrated.[11]

Fat tails are usually measured by the Hill estimator $\hat{\alpha}_H$ of the absolute returns. With respect to a threshold value v^o specifying the tail (to which all $v_t \ge v^o$ are assigned), it is given by $\hat{\alpha}_H=1/\hat{\gamma}_H$, where $\hat{\gamma}_H=(1/\tau)\Sigma_t \max\{0, \ln v_t - \ln v^o\}$ and τ is the number of entries in this tail. Hence lower values of $\hat{\alpha}_H$ and higher values of $\hat{\gamma}_H$ indicate a fatter tail of the data. The definition of $\hat{\gamma}_H$ motivates us to put forward the following two additional moment functions, which are based on a given value v^o that specifies a 5% tail of the absolute empirical returns:

$$\underline{v\,Hill\,k}:\ m_i(z_t) = \begin{cases} \ln v_t - \ln v^o & \text{if } v_t > k v^o \\ 0 & \text{else} \end{cases} \quad (k=1,2) \tag{19}$$

If $k=1$, the function corresponds to the single terms in the sum that yields $\hat{\gamma}_H$. In a simulation of the model, however, the moment 'v Hill 1' results from dividing the sum of these $m_i(z_t)$ through sT, and since we do not know in advance how many v_t will exceed the predetermined value of v^o, this expression cannot be suitably rescaled to compute the statistic $\hat{\gamma}_H$ proper.

Whereas this kind of scaling does not matter if also roughly 5% of the model-generated v_t exceed v^o, it is possible that the empirical and simulated moments 'v Hill 1' are approximately equal although the model puts only, say, 3% in the tail specified by v^o. This might happen if a larger part of these 3% attains more extreme values than the empirical data. It is the idea of the second

[11] At least we know of no corresponding treatment in the literature.

Table 2
Specification of three parameter sets.

	ϕ	y	z	ρ	H
Set 1	2.342	1.350	0.658	0.737	21
		(0.078)	(0.029)	(0.038)	
Set 2	2.362	1.861	0.759	0.956	13
		(0.193)	(0.028)	(0.069)	
Set 3	3.297	1.960	0.777	0.993	13
		(0.192)	(0.026)	(0.050)	

Note: Numbers in parentheses are the standard errors from Eq. (13) for the estimations on S&P 500 described in the text.

function 'v Hill 2' to avert such a misleading match, since in this case the simulated moment would be noticeably larger than its empirical counterpart.

3.3. Re-estimation of the model

Before applying MSM with the moment functions (Eqs. (18)–(19)) to empirical data, it should first be tested if the method is reliable enough to identify the model itself. To this end we adopt the parameter values from Manzan and Westerhoff (2005) and try to re-estimate them from an artificial sample series.

Regarding the moments to be included in the estimation, the most obvious candidates are the means and variances of r_t and v_t. A zero mean of the raw returns will be automatically brought about by the model, so it only makes sense to include 'v mean'. The functions 'r ACV 0' and 'v ACV 0' cannot be simultaneously included because this leads to a (nearly) singular covariance matrix $\hat{\mathbf{\Omega}}$. Hence we only check the matching of the variance 'r ACV 0'.

As the correlograms in Manzan and Westerhoff (2005, p. 686) show, the raw returns have near-zero autocorrelations at all lags—even at the first lag, where several other models in the literature have difficulty avoiding significance. Since it can be reasonably expected that if the coefficients are close to zero at short lags, they will also tend to vanish at longer lags, it will do to include the moment function 'r ACV 1'; functions with higher lags k would provide no additional information.

In accordance with their empirical counterparts, the autocorrelations of the absolute returns in MW are significantly positive and quite steadily declining as the lag length increases. It thus suffices to select one short, one medium, and one longer lag, which are represented by the moments 'v ACV 1', 'v ACV 5', and 'v ACV 50'. Lastly, the two Hill functions in Eq. (19) are utilized, so that underlying the estimation is a total of 8 moment functions ($n_m = 8$). In the objective function $\mathbf{Q}_{s,T} = \mathbf{Q}_{s,T}(\boldsymbol{\theta})$ in Eq. (4) we use the optimal weighting matrix $\mathbf{W} = \mathbf{W}^{\mathrm{o}} = \hat{\mathbf{\Omega}}^{-1}$ from Eq. (6); the corresponding J test statistic in Eq. (8) is designated J^{o}.

After these preparations we can take the parameter values from Manzan and Westerhoff (2005, p. 683), generate a sequence of (pseudo) random numbers and use them to simulate the model over more than 7000 days. We discard a transient phase of several hundred days and retain a return series over a period of $T = 6760$ days, which is of a similar length to some of the empirical series below. This series represents our empirical data, on which the model is now to be re-estimated. Of course, the simulations of the moments in the estimation procedure are based on a different random number sequence (the same for each parameter set, then). The corresponding time horizon is taken ten times longer than the empirical series, $s = 10$.

The true numerical parameters are a standard deviation $\sigma_\eta = 0.010$ for the innovations to the fundamental value, $H = 10$ for the length of the rolling sample period, and the values given in the first row of Table 1. In order to avoid unnecessary digits the table

Table 3
Moments, t^{o}-statistics and values of the objective functions for the three parameter sets in Table 2.

	Moments				t^{o}-statistics		
	Empirical	Set 1	Set 2	Set 3	Set 1	Set 2	Set 3
r ACV 0	0.971	0.826	0.939	0.937	−2.88	−0.63	−0.66
r ACV 1	0.021	0.007	0.002	0.003	−0.93	−1.30	−1.22
v mean	0.706	0.680	0.714	0.715	−1.78	0.54	0.61
v ACV 1	0.068	0.063	0.087	0.084	−0.37	1.37	1.15
v ACV 5	0.095	0.062	0.086	0.084	−2.44	−0.67	−0.83
v ACV 50	0.058	0.053	0.061	0.060	−0.56	0.26	0.14
v Hill 1	1.431	0.974	1.370	1.311	−2.71	−0.36	−0.71
v Hill 2	0.361	0.068	0.351	0.377	−3.31	−0.11	0.19
Hill est.	3.43	4.23	3.10	3.11	Values of objective function		
J^{o}					**36.21**	55.64	58.40
J^{d}					68.06	**7.83**	8.23
$\mathbf{Q}^{\mathrm{bt}}$					42.66	1.53	**1.19**

Note: The t^{o}-statistics are given by Eq. (10), where all three parameter sets have the same matrix $\mathbf{U}$ determined in Eq. (9) underlying. 'Hill est.' is the Hill estimator $\hat{\alpha}_{\mathrm{H}}$.

Table 4
The empirical series to be estimated.

Series	Begin of series	Observations
S&P 500^{pc}	Jan 1980	6867
S&P 500	Jan 1980	6864
Dow Jones^{pc}	Jan 1980	6867
Dow Jones	Jan 1980	6864
DAX	Nov 1990	4115
Nikkei	Jan 1984	5712
USD/DEM	Jan 1980	6862
JPY/USD	Jan 1980	6857

Note: All series are daily and end around March 20, 2007. Superscript ‘pc’ stands for ‘purged of crash’, that is, the returns of the three days October 19–21, 1987, are excluded. The stock market indices are the close price adjusted for dividends and splits. Following euro adoption, USD/DEM are pseudo rates imputed by applying the euro locking rate to the current euro exchange rate.

refers to $\phi := 100 \cdot x$ and $\rho := 100 \cdot K$ instead of x and K (‘ϕ’ may help to recall that the coefficient is associated with the deviations of the current price from the fundamental value, and ‘ρ’ may be indicative of the overreaction threshold for the absolute returns in the price equation). The first row in Table 1 also shows that the simulation with the true parameters and the new random shocks yields a J-statistic of $J^o = 8.14$. Hence if the four parameters in that row were the outcome of the estimation, they would not be rejected since the critical 95% significance value is $\chi^2(n_m - n_p) = \chi^2(8-4) = 9.49$.

Owing to the sampling variability there will certainly be other parameter values that give rise to lower values of J^o and the objective function $\mathbf{Q}_{s,T}$, respectively. For these estimations it has first to be noted that the three parameters y, z and σ_η cannot be independently identified, because only the products $y\sigma_\eta$ and $z\sigma_\eta$ enter the determination of the price in Eq. (15). We may therefore maintain Manzan and Westerhoff's standard deviation $\sigma_\eta = 0.010$ as a normalization, which amounts to an annual volatility in the fundamental value of 15.8% (on account of $0.158 / \sqrt{250} = 0.010$). This issue will nonetheless be later reconsidered. For the present estimations we also fix the integer parameter H at its correct value $H = 10$.[12]

This being understood, the four parameters ϕ, y, z, ρ have to be estimated. In our first attempt, we choose an initial set of these parameters with, in particular, $\phi = 2$, start the Nelder–Mead simplex search algorithm from there (see Press et al., 1986, pp. 289–293), and restart it upon convergence several times until no more noteworthy improvement in the minimization occurs. As seen in the second row of Table 1, a decrease of the J test statistic down to $J^o = 2.57$ is achieved in this way, with only marginal deviations of the other parameters y, z and ρ from their true values. According to the standard errors in parentheses, these estimates are fairly precise. The estimated ϕ, by contrast, remains at the initial order of magnitude of this coefficient, and its high standard error indicates that this estimate is not very reliable.

Actually, the function J^o is rather insensitive to changes in the parameter ϕ. This is exemplified in Table 1 by fixing ϕ at such extreme values as $\phi = 0.10$ and $\phi = 10$ and estimating the remaining y, z. ρ. These estimates are not very different from the original ones, and also the minimal J^o is not much higher. The last row in the table, where ϕ is fixed at zero, demonstrates that the fundamentalists may be even completely relegated from the model.

We can conclude from this evidence that the estimation of the model-generated returns can recognize the switching mechanism of the speculators in quite a precise manner, but it fails to capture an interplay between the two groups of traders. This, however, comes as no great surprise since there is simply no moment which could notice the excessive deviations of the price from the fundamental value that would be caused by an insufficient influence of the fundamentalists. While it would be desirable to complement our set of moments by a measure of misalignment, this is not easily done because the model's daily fluctuations of the fundamental value are unobservable (which was another reason why we opted for MSM). We will therefore have to accept that the parameter ϕ cannot be properly identified.

Finally the matching of the moments should be checked. The low J test statistics reported in Table 1 suggest that there is no problem with the moments included in the estimations. In finer detail we can refer to the t^o-statistics from Eq. (10). Computing them for the second estimation in Table 1, they are indeed less than one in modulus for all of the eight moments.

4. Estimation on empirical data

4.1. A comparison of three minimization criteria

After applying MSM and our specification of the moments to artificial data that was generated by the model itself, we are now ready to work with empirical data. Although Manzan and Westerhoff present their model as a behavioural exchange rate model, their story could quite as well relate to a stock market or stock market index. We will in this section consider several empirical exchange rates and stock market indices, but let us first concentrate on the Standard & Poor's stock market index, the S&P 500. The series to be estimated comprises $T = 6864$ days from January 1980 to March 2007. However, as it is occasionally done in other studies, we choose to purge the return series of the October 1987 crash; that is, the three days of October 19–21 are discarded. As will be seen a little later, this is a harder test for the model than the full series.

[12] Since H is an integer number, including it in the parameters to be estimated means that ϕ, y, z, ρ are obtained from minimizing J^o under alternative values of H. It was checked that the minimal J^o for $H \neq 10$ generally exceed the values reported in Table 1. Of course, no standard error can be computed for this parameter.

Table 5
Estimation of the series listed in Table 4.

	ϕ	$\tilde{y}$	$\tilde{z}$	$100\cdot\tilde{\sigma}_\eta$	ρ	H	$\mathbf{Q}^{bt}$	$\mathbf{Q}^{bt}$	ϕ
S&P 500^{pc}	0.50	1.43	0.57	1.37	0.99	13	1.34	1.19	3.30
S&P 500	0.50	1.42	0.58	1.28	0.92	13	0.18	0.19	2.78
	0.50	1.46	0.54	1.39	0.99	10	0.02	0.02	2.78
Dow Jones^{pc}	0.50	1.41	0.59	1.35	1.00	13	2.57	2.25	2.83
Dow Jones	0.50	1.44	0.56	1.42	1.02	13	0.61	0.62	2.18
	0.50	1.48	0.52	1.49	1.08	8	0.02	0.01	2.11
DAX	0.50	1.48	0.52	1.82	1.24	13	1.48	1.47	2.07
Nikkei	0.50	1.42	0.58	1.69	1.25	9	0.07	0.06	2.79
USD/DEM	0.50	1.31	0.69	0.80	0.63	16	4.23	3.55	2.88
JPY/USD	0.50	1.38	0.62	0.85	0.66	11	7.95	7.66	2.0

Note: The estimations are based on the function $\mathbf{Q}^{bt}$ as it is defined in Eq. (20), with $t^b=0.75$. The values of $\tilde{y}$, $\tilde{z}$, $\tilde{\sigma}$ are derived from $\sigma=0.01$ and the estimates of y, z as stated in Eq. (22). In the left part of the table, ϕ was exogenously set at $\phi=0.50$, the last two columns report the estimated value of ϕ and the value of the objective function thus brought about.

We begin the discussion with contrasting three sets of coefficients, which result from the three minimizations described below. The values are given in Table 2. Interestingly, apart from ϕ (whose estimates as we know are not very precise) and the length of the rolling sample period H, the parameters of set 1 are fairly close to Manzan and Westerhoff's tentative values. Simulating the model with the three parameter sets and employing the same moments as in the previous section (again putting $s=10$), the statistics compiled in Table 3 are obtained.

It is a little caprice of the data that the first-order autocovariance of the absolute returns is lower than 'v ACV 5', and one will not expect a model to mimic this feature.[13] We have indeed found that all model parameters of any relevance produce a slow monotonic decay of the autocovariances of v_t, though possibly at different levels. Furthermore, the three parameter sets dealt with in Table 3 keep the autocovariances distinctly positive even at a lag of 50 days. This brief overview confirms that the model is basically compatible with at least a certain subset of the stylized facts.

The first set of parameters originates with a minimization of $\mathbf{Q}_{s,T}$ under the optimal weighting matrix $\mathbf{W}^o$ (some details of the technically somewhat involved minimization procedure across the five parameters can be found in Franke, 2008, Section 4.2). Unsurprisingly, the associated J test statistic $J^o=36.21$ rejects the model as the true data generation process. We may nevertheless still ask which of our 8 moments can be reproduced by the simulations, and to what extent. To this end the table reports the levels of the moments and the t^o-statistics from Eq. (10). Clearly, an unsatisfactory match is observed for the variance of the raw returns and the two Hill moments (as well as for the Hill estimator $\hat{\alpha}_H$ itself).

Taking the estimation with the first parameter set as a point of departure, it is remarkable that the second parameter set, in spite of its considerably higher value $J^o=55.64$, achieves a largely superior collection of t^o-statistics. Unless a high priority is assigned to 'r ACV 1' and 'v ACV 1' with their t_i^o-values of -0.93 and -0.37, respectively, a direct comparison of the t^o-statistics leads one to prefer the moment matching of parameter set 2 to that of set 1. Note also the far superior match of the Hill estimator.

As a matter of fact, the second parameter set has been obtained from neglecting the cross-correlations among the moments and replacing the optimal weighting matrix $\mathbf{W}^o$ with the diagonal matrix $\mathbf{W}^d$ from Eq. (7) in the objective function (Eq. (4)). This is also evidenced by the minimal value of J^d in bold face, which denotes the corresponding J-statistic (Eq. (8)). It may here be recalled that J^d has already been noted to coincide with the sum of the squared t_i-statistics in Eq. (11), which equivalently can thus be said to be minimized by parameter set 2. Although these t_i are numerically different from the t_i^o that were obtained from Eq. (10) and used in Table 3 as a common basis to compare the parameter sets, the two statistics convey a very similar information.[14]

It follows that if one acknowledges the falsity of the model and rather aims at a close match of the selected moments that is directly recognizable, then (J^d or, equivalently) the sum of the squared t_i-statistics is a more appropriate, or at least more intuitive, minimization criterion. We may then even take one little step further. It captures the idea that two moments having both moderate t-statistics may be preferred to one moment with a very low and one with a large statistic. Accordingly the t-statistics may count as a 'loss' only if they exceed a certain positive benchmark value t^b, which leads us to put forward the following objective function:

$$\mathbf{Q}^{bt} = \mathbf{Q}^{bt}_{s,T} := \sum_{i=1}^{n_m} \{max[0, |t_i| - t^b]\}^2; \qquad t_i \text{ from } (11). \tag{20}$$

Parameter set 3 is the outcome of a minimization of this function, where we decided on a benchmark $t^b=0.75$. While Table 3 still records the t^o-statistics of the moments, the effect of the modification in the objective function is still clearly visible: the two

[13] Besides, including the 1987 crash raises 'v ACV 1' and 'v ACV 5' to 0.102 and 0.118, respectively, which is a nonnegligible change in the time averages of such a long series.

[14] Using J^d in the re-estimation of the artificial data discussed in Section 3.3, this function, too, was able to recognize the true parameters (apart from ϕ, of course).

largest values 1.30 and 1.37 of the second parameter set for 'r ACV 1' and 'v ACV 1' are reduced to 1.22 and 1.15, respectively. In exchange, the absolute values of the t^{o}-statistics of the Hill moments, for example, increase from 0.36 and 0.11 to 0.71 and 0.19, but a deterioration of this order of magnitude has just been defined as negligible. Hence function (20), as documented in the last row of the table, yields a lower value for set 3 than for set 2, although set 3 is inferior in terms of the two other functions J^{o} and J^{d}.

The relatively low t^{o}-statistics for parameter set 3 or 2 are remarkable for two reasons. First, they are associated with the optimal weighting matrix $\mathbf{W}^{o}$, whereas the objective functions $\mathbf{Q}^{bt}$ and J^{d} yielding the two parameter sets derive from different principles. Second, these statistics are obtained in a highly stylized structural model with essentially one volatility mechanism. Although the model is definitely rejected by the J test (see the row J^{o} in Table 3), its matching of the empirical moments can therefore still be regarded as fairly satisfactory.

Comparing the standard errors of the three parameter sets in Table 2 with those in the second row of Table 1, which was concerned with the estimation of the model-generated return series, it is seen that the estimations of the coefficients y, z and ρ are also quite precise. The estimates of ϕ are, of course, as unreliable as in the artificial re-estimations of the model; its standard error is therefore omitted. There is no standard error for H since it is an integer parameter which we fixed at alternative values and finally chose the one that gave rise to the best estimation result.

4.2. Estimation results for different financial markets

In the remainder of the paper we prefer the intuitive interpretation of the objective function $\mathbf{Q}^{bt}$ specified in Eq. (20). We use it now to estimate the model on the daily returns from a few other financial markets. First, the 1987 October crash should be included in the S&P 500. Second, we wish to check how similar an estimation of the Dow Jones Industrial Average Index may be to the S&P 500. We also consider the German stock market index DAX and the Japanese Nikkei index (in both of which the 1987 crash had less dramatic effects). Finally, we are interested in the question if the model can differentiate between stock and foreign exchange markets. To this end, the model is tested with the exchange rates of the US dollar (USD) against the Deutsche Mark (DEM), and the Japanese Yen (JPY) against the US dollar. The sample periods of these empirical series and the number of observations are given in Table 4.

The feature that, as discussed above, the coefficient ϕ cannot be precisely identified prompts us to set it exogenously at a common value for all series. Since within a certain range one value is as good as any other, we quite arbitrarily choose $\phi = 0.50$. The estimations resulting from this are reported in the left part of Table 5.[15] Alternatively, we have also included ϕ in the estimation, where our extensive search procedure started out from $\phi = 2$. Here the table, in its right part, only reports the improvement in the objective function $\mathbf{Q}^{bt}$ that (if ever) could thus be achieved.

The common value of ϕ is a specification that makes the results better comparable across the financial markets. An additional issue is the scaling of the standard deviation σ_η of the random walk of the fundamental value, which we had so far normalized at $\sigma_\eta = 0.010$. This choice also determines the levels of y and z. However, in accordance with their interpretation of the demand of the speculators, Manzan and Westerhoff centred the two parameters around unity (see the first row in Table 1). In Table 5 we have achieved this convenient and theoretically motivated property by a suitable rescaling of σ_η. That is, having estimated y and z on the basis of $\sigma_\eta = 0.010$, we specify new values $\tilde{\sigma}_\eta$, $\tilde{y} > 1$ and $\tilde{z} < 1$ obeying the relationships

$$\tilde{y}\,\tilde{\sigma}_\eta = y\sigma_\eta, \quad \tilde{z}\,\tilde{\sigma}_\eta = z\sigma_\eta, \quad \tilde{y} - 1 = 1 - \tilde{z}. \tag{21}$$

The solution to these equations is

$$\tilde{y} = 2y/(y+z), \quad \tilde{z} = 2z/(y+z), \quad \tilde{\sigma}_\eta = (y+z)\sigma_\eta/2. \tag{22}$$

In this way the model has regained some of its structural interpretation. The values of $\tilde{y}$ and $\tilde{z}$ allow us a direct assessment of the speculators' over- and underreactions to the news on the different markets, and $\tilde{\sigma}_\eta$ gives us an impression of the relative volatilities of the fundamental news underlying these markets.[16]

The last two entries in the first row of Table 5 reproduce the value of ϕ of parameter set 3 in Table 2 and the minimal value of $\mathbf{Q}^{bt}$ in the last column of Table 3. The first row in Table 5 shows that the constraint $\phi = 0.50$ leads to a certain deterioration in the objective function. The changes in the coefficients are, however, only marginal, which once again indicates the almost negligible role of ϕ (with $\sigma_\eta = 0.010$, the other coefficients are estimated as $y = 1.954$, $z = 0.776$, $\rho = 0.991$, which may be compared to parameter set 3 in Table 2).

Including the three crash days of October 1987, while still fixing the rolling sample period at $H = 13$, improves the fit of moments. A free choice of H, where $H = 10$ turned out to be optimal, even decreases practically all t-statistics below their benchmark $t^{b} = 0.75$. The corresponding changes in the parameter estimates remain nevertheless fairly limited. The model's increased performance in this case confirms its potential for explaining bubbles.

The estimation of the Dow Jones produces results that are not much different from the first three rows in Table 5. Since the segments covered by the two stock market indices are similar, this attests to a reasonable robustness of the model.

As concerns the DAX and Nikkei, the over- and underreactions are about as strong as in the two American indices. In contrast, from the estimation of the German and Japanese markets we infer that their fundamental news process exhibits a higher volatility.

[15] The standard errors of the estimated parameters are of a similar order of magnitude to those in Table 2 and are so no longer discussed.

[16] If one does not share Manzan and Westerhoff's theoretical motivation and has an idea of a credible standard deviation s^{c}_{η}, one can easily compute the corresponding value of y and z from $\tilde{y}$, $\tilde{z}$, $\tilde{\sigma}_\eta$ reported in Table 5 and the relationships $\tilde{y}\tilde{\sigma}_\eta = y\sigma^{c}_\eta$, $\tilde{z}\tilde{\sigma}_\eta = z\sigma^{c}_\eta$.

It makes good sense that in these cases also the threshold $\rho = 100 \cdot K$ in Eq. (15) for the recent history of the returns, above which the speculators overreact to the news, is higher than for the S&P 500 and Dow Jones (although the relationship between $\tilde{\sigma_\eta}$ and ρ is not strictly monotonic).

The estimation of the two foreign exchange rates in the last two rows of Table 5 are notable for three reasons. First, the matching of the moments is inferior to the stock markets. Second, the overreaction is weaker than on the stock markets. And third, the fundamental values on the foreign exchange markets appear to be far less volatile. To this conforms a lower threshold ρ, which is also brought out by the estimations.[17] From this and the results above we can summarize that the model jointly with the MSM estimation has some meaningful explanatory power.[18]

5. Conclusion

Any structural model is false and therefore, trivially, cannot be expected to reproduce reality in all facets. Given that a model is meant for estimation at all, it will fulfill its purpose if it captures reality along certain dimensions, for which it was also more or less designed. Besides the treatment of unobservable variables and nonlinearities in the model dynamics, it is one advantage of the method of simulated moments that it obligates the researcher to make these dimensions explicit, in the form of moments. Our estimations have emphasized this point in that, after some explorations, we decided to abandon the econometrically optimal weighting of the moments in the objective function. We instead postulated directly the minimization of the *t*-statistics of the moments, or more precisely, of the sum of their squared deviations from a benchmark value. As discussed at the beginning of Section 4, this gives us a more intuitive measure of how well the model-generated moments are able to match the empirical moments.

Applying this approach to a minimal agent-based asset pricing model by Manzan and Westerhoff (2005), it was found that although the model was rejected by the standards of econometric theory, it nevertheless performs reasonably well in terms of our criterion, where the selected moments themselves relate to the daily raw and absolute returns and specify the stylized facts of their autocorrelation patterns and fat tails. We may also emphasize that whichever of our three versions of the objective function is chosen, it provides a measure that can be directly compared across different models. The present goodness of moment fitting can thus be regarded as a challenge to other models of similar simplicity, which they still have to stand.

The estimations of the model can identify all of the parameters except the one that reflects the relative importance of the fundamentalists. This failure was seen to be due to the unobservability of the daily fluctuations of the fundamental value, so that we had no empirical moment (readily) available that could take account of a possible misalignment if the influence of the fundamentalists is too weak. Dealing with this problem may be one issue for future research.

On the other hand, the coefficients characterizing the behaviour of the group of speculators come out meaningful and are practically not affected by the unidentified parameter. They are furthermore quite robust across similar markets. In particular, the estimations offer us an interpretation of why stock and foreign exchange markets are different: on the stock markets the model's speculators are more responsive to the arrival of news, and the fundamental value itself or its news process is more volatile. Given its simplicity, the estimation results lend the model a remarkable explanatory power. Nevertheless, there is certainly scope for structural improvement, and for applying similar versions of MSM to these enhanced models.

References

Alfarano, S., Franke, R., 2007. A simple asymmetric herding model to distinguish between stock and foreign exchange markets. Working Paper, University of Kiel. http://www.bwl.uni-kiel.de/gwif/downloads_papers.php?lang=en.

Boivin, F., Giannoni, M.P., 2006. Has monetary policy become more effective? Review of Economics and Statistics 88, 445–462.

Carrasco, M., Florens, J.-P., 2002. Simulation-based method of moments and efficiency. Journal of Business and Economic Statistics 20, 482–492.

Chen, S.-H., Chang, C.-L., Du, Y.-R., 2008. Agent-based economic models and econometrics. Paper presented at the Econophysics Colloquium, Kiel, August 2008.

Chiarella, C., Dieci, R., He, X.-Z., 2009. Heterogeneity, market mechanisms, and asset price dynamics. In: Hens, T., Schenk-Hoppé, K.R. (Eds.), Handbook of Financial Markets: Dynamics and Evolution. InNorth-Holland, Amsterdam, pp. 277–344.

Davidson, R., MacKinnon, J.G., 2004. Econometric Theory and Methods. Oxford University Press, Oxford.

Duffie, D., Singleton, K.J., 1993. Simulated moments estimation of Markov models of asset prices. Econometrica 61, 929–952.

Fagiolo, G., Birchenhall, C., Windrum, P., 2007. Empirical validation in agent-based models: introduction to the Special Issue. Computational Economics 30, 189–194.

Franke, R., 2008. Applying the method of simulated moments to estimate a small agent-based asset pricing model: extended version. Working Paper, University of Kiel. http://www.bwl.uni-kiel.de/gwif/downloads_papers.php?lang=en.

Greene, W.H., 2002. Econometric Analysis, 5th edition. Prentice-Hall, Upper Saddle River, N.J.

Hommes, C., Wagener, F., 2009. Complex evolutionary systems in behavioral finance. In: Hens, T., Schenk-Hoppé, K.R. (Eds.), Handbook of Financial Markets: Dynamics and Evolution. InNorth-Holland, Amsterdam, pp. 217–276.

Kocherlakota, N.R., 1990. On tests of representative consumer asset pricing models. Journal of Monetary Economics 26, 285–304.

Lee, B.-S., Ingram, B.F., 1991. Simulation estimation of time series models. Journal of Econometrics 47, 197–205.

Lux, T., in press. Applications of statistical physics in finance and economics. In: Rosser, J.B., Jr. (Ed.), Handbook of Research on Complexity. Edward Elgar, Cheltenham.

Manzan, S., Westerhoff, F., 2005. Representativeness of news and exchange rate dynamics. Journal of Economic Dynamics and Control 29, 677–689.

Manzan, S., Westerhoff, F., 2007. Heterogeneous expectations, exchange rate dynamics and predictability. Journal of Economic Behavior and Organization 64, 111–128.

[17] It is interesting to compare the latter two points with the distinguishing feature of stock and foreign exchange markets that was revealed by an empirical analysis of a totally different model by Alfarano and Franke (2007). There the stock markets were characterized by a consistently higher population share of noise traders *vis-à-vis* the stabilizing fundamentalists.

[18] Although the fit of the exchange rates is worse than the fit for the stock market indices, the working paper (Franke, 2008) contains a discussion from which it is concluded that, generally speaking, this fit is still not so bad after all.

Mao, C.S., 1990. "Hypothesis testing and finite sample properties of generalized method of moments estimators: A Monte Carlo study", Federal Reserve Bank of Richmond, Working Paper 90–12.
Press, W.H., et al., 1986. Numerical Recipes: The Art of Scientific Computing. Cambridge University Press, Cambridge.
Westerhoff, F., in press. Exchange rate dynamics: a nonlinear survey. In: Rosser, J.B., Jr. (Ed.), Handbook of Research on Complexity. Edward Elgar, Cheltenham.
Winker, P., Gilli, M., Jeleskovic, V., 2007. An objective function for simulation based inference on exchange rate data. Journal of Economic Interaction and Coordination 2, 125–145.