ISSN: 2039-4004

department of economics
"S. Cognetti de Martiis"

Via Po, 53 – 10124 Torino (Italy)
Tel. (+39) 011 6704043 - Fax (+39) 011 6703895
URL: http//www.de.unito.it

WORKING PAPER SERIES

ESTIMATING MICROMOTIVES FROM MACROBEHAVIOR

Jakob Grazzini

Dipartimento di Economia “S. Cognetti de Martiis”

Working paper No. 11/2011

Università di Torino

Estimating Micromotives from Macrobehavior

Jakob Grazzini*

University of Turin, Department of Economics, via Po 53, 10124 Torino.

November 25, 2011

Abstract

In large economic systems the observed data are usually macro-data; an agent-based model is a set of individual decision makers that influence the macro behavior of the system both trough their direct actions and trough the influence they have on the other decision makers. The complexity of the model impedes an explicit analytical relation between the micro-parameters and the macro-behavior, and in turn the use of "traditional" econometric tools. This paper develops an estimation procedure that allows the estimation of the agents' micromotives in an agent-based model using the emerging behavior of the system. Starting from a simple stock market model, in which agents learn by interacting, the properties of simulation-based estimations are investigated. As a first step in the direction of extending and adapting the econometric literature to agent-based models, the paper is focused on the consistency and on the bias of the estimates.

JEL classifications: C15, C53,C63

Keywords: Agent-based models, Estimation, Simulation, Stock Market

1 Introduction

The economic system is composed by many different autonomous agents that interact with each other and with the environment. The result is a system that exhibits emergent properties: the properties at the macro level cannot be explained directly by the properties at the micro level (Gilbert 2001). Agent-based modeling is a tool used to overcome the limitations of a purely mathematical analysis and it allows the construction of more realistic models; unfortunately this happens at a cost. Agent-based models are more difficult to understand,

*Email: `jakob.grazzini@unito.it`

to generalize and to explain. A model consisting of algebraically solved equations can easily be interpreted and generalized using formal proofs. Despite the fact that it can be considered as a well-defined set of equations (Leombruni & Richiardi 2005), an agent-based model suffers from the different (smaller) degree of knowledge about the functions that are at the base of the model. While analytical results are conditional only in relation to the specific hypothesis about the model, simulation results are conditional in relation to both the specific hypothesis of the model and to the specific values of the parameters used in the simulation runs: each run of an agent-based model yields a sufficiency theorem, but a single run does not provide any information on the robustness of such theorems (Axtell 2000). To treat the "sufficiency problem", a sensitivity analysis over the parameter space has to be performed in order to assess the robustness of the results (Axtell 2000). An estimation procedure in agent-based models is crucial to compare the model with empirical data but also because it can be used to reduce the sufficiency problem by reducing the space of the parameters to the neighborhood of the "empirical relevant" parameters. Estimation of agent-based models is important in the interpretation of the model and possibly in the validation of it (Windrum et al. 2007, Bianchi et al. 2007), but it is still largely missing in the literature (Alfarano et al. 2005, Richiardi et al. 2006, Leombruni & Richiardi 2005). The interaction between the agents in a stock market, for example, is largely accepted as fundamental in shaping the properties of the markets. This led to building (e.g. Arthur et al. (1997), Kirman (1993), Lux & Marchesi (1999), Brock & Hommes (1998), Cross et al. (2005)) and rarely to estimating complex financial models Boswijk et al. (2007), Alfarano et al. (2005), Gilli & Winker (2003), for a survey see Chen et al. (2009).

This paper describes a method for using empirical data in agent-based models; by using observed data about the system under analysis it is possible to select the values for the parameters so that the artificial data and the observed data are as similar as possible, i.e. minimizing an objective function. The results of such minimization will crucially depend on the properties of the model. The starting point for the estimation of the parameters is the simulation-based econometrics literature using the method of simulated moments [1]. The smaller degree of knowledge about the model is a problem also in the estimation procedure. The properties of the model are not known *a priori*, this means that in order to know how to interpret the parameters resulting from simulation-based econometrics methods, the model has to be tested. In particular it is necessary to perform a sensitivity analysis to understand the behavior of the moments of the model with changing parameters and to choose the moments that allow a sufficient characterization of the model behavior with

[1] Other simulation-based econometrics techniques could also be useful such as the indirect inference and simulated maximum likelihood.

different parameters. The choice of the moment is crucial for the identification problem and influences the efficiency of the estimator. Once the moments are chosen it is necessary to understand whether they are stationary and ergodic, the test that will be used are described in Grazzini (2012). The tests are performed on the artificial data, and using the well-specification hypothesis, the results are extended to the real data. If the moments are well behaved then it is possible to consistently estimate the agent-based model. The model used to show the estimation procedure is a simple stock market model in which the agents learn the equilibrium price trough the information provided by the actions of the other traders. The behavior of the emergent price is used to estimate a behavioral parameter that governs the learning process.

2 The model

The model used to show the estimation procedure is an agent-based stock market model proposed in Cliff & Bruten (1997) to reproduce the experimental results obtained by Smith (1962). In the experimental market the traders are divided into buyers and sellers. Each subject receives a card containing an induced value for the fictitious commodity. The trade is conducted through a continuous double auction over a sequence of periods. The agents are free to bid and offer at any time and they withdraw from the market for the given period when they successfully close a deal. The induced value for each trader is the same in all periods. The aim in Smith (1962) was to study the behavior of the price in a controlled situation where demand and supply schedule were well defined over a unit of time. Figure 1 shows the market environment, determining the supply and demand schedules and the theoretical equilibrium. The experiment shows how a small number of inexperienced traders converge rapidly to a competitive equilibrium under the double auction mechanism (Smith 1962, p.157). The result is interesting because it shows how the interaction between the traders allows the emergence of the equilibrium price and how, in this simple environment, the equilibrium price is predictable by the classical microeconomic theory (Cliff & Bruten 1997). The private profit-seeking incentives allow the market to reach the equilibrium and even if the market environment (demand and supply schedule) is unknown the traders learn it in a few periods [2].

The model proposed by Cliff & Bruten (1997) reproduces explicitly the continuous double auction, therefore there is a book where the "active" traders can ask or bid. The behavior of the traders is very simple, they use the observed orders in the book to improve their trading strategy. In particular the bid or ask depends on their predefined role and they adapt the

[2] In Smith (1962), note 5: "It is only through some learning mechanism of this kind that I can imagine the possibility of equilibrium being approached in any real market."

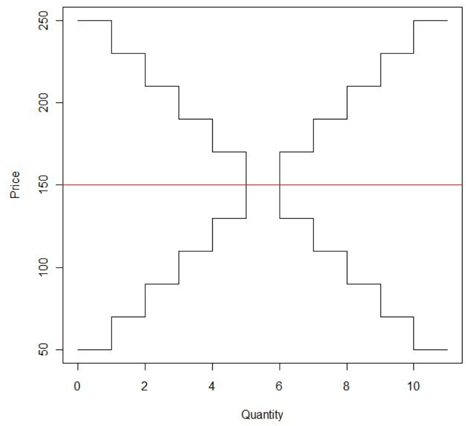


Figure 1: Supply and demand functions, defined by the private values of the traders.

price through a simple learning mechanism. The price proposed by the agent i is

$$p_i(t) = v_i(1 + \mu_i(t)) \tag{1}$$

where v_i is the certain induced value and $\mu_i(t)$ is the profit margin, positive for sellers and negative for buyers. Note that a budget constrain is imposed: no bid or offer can be made with a loss. The profit margin evolves over time following a very simple heuristic. The pseudo algorithms for a seller and a buyer are Algorithm 1 and Algorithm 2 [3]. The agents observe the book and use the information about the last proposal to understand the market.

Algorithm 1 The basic behavior of a Seller: adapting the profit margin

Seller
if the last shout was accepted at price q **then**
  1. any seller s_i for whom $p_i \leq q$ should raise the profit margin
  2. if last shout was a bid and $p_i \geq q$, any active seller s_i should lower the margin
else
  if the last shout was an offer and $p_i \geq q$ any active seller s_i should lower the margin
end if

Algorithm 2 The basic behavior of a Buyer: adapting the profit margin

Buyer
if the last shout was accepted at price q **then**
  1. any buyer b_i for whom $p_i \geq q$ should raise the profit margin
  2. if last shout was an offer and $p_i \leq q$, any active buyer b_i should lower the margin
else
  if the last shout was a bid and $p_i \leq q$ any active buyer b_i should lower the margin
end if

[3] On http://www.jakob.altervista.org/Python-model1.rar it is possible to download the python files of the model

The algorithm is well explained in Cliff & Bruten (1997). In the Smith (1962) experiment the time was divided into periods and each trader had the opportunity to trade only once [4] during each period. The traders start the period as active traders and become "non-active" after having agreed on a contract. The aim of the traders is to trade at the best possible condition, i.e. with the maximum possible profit margin. The seller s_i might for example start offering at a given price $\bar{p}_i(t)$. From equation 1 we know that the offered price depends on the private (constant) value and on the profit margin. If s_i observes that the last order was accepted at a price q greater than $\bar{p}_i(t)$, the incentive to maximize the profits will induce the seller to increase the profit margin. The observed order tells seller i that there are buyers willing to buy at a higher price. If on the contrary the last accepted order was an offer with a price q lower than $\bar{p}_i(t)$, the incentive to trade induces s_i to lower the profit margin (if it is greater than zero). The seller s_i will use the information contained in the last order to infer on the behavior of the other sellers. To be able to trade she must reduce the selling price (reducing the profit margin) to undercut the competition. It is important to note that the reduction of the profit margin by a seller is triggered only by offers: the traders undercut their competitors. If on the other hand the sellers react also to very low bids, the buyers could coordinate and artificially reduce the price. For the buyers the algorithm works symmetrically. The crucial point is to understand how the traders adapt and that only some bids and offers influence the market. Extra-marginal traders and exceptional bids and offers (very low bids and very high offers) have no effect on the market. This simple algorithm allows the traders to understand the optimal pricing strategy by adapting the profit margin. In order to adapt there is the need for some form of updating rule. Cliff & Bruten (1997) propose the Windrow-Hoff "delta rule":

$$A_{t+1} = A_t + \Delta_t \tag{2}$$

where A_{t+1} is the output after the update, A_t is the current output and Δ_t is the change in output in time t and depends on the difference between actual A_t and the desired output D_t and a learning rate coefficient β:

$$\Delta_t = \beta(D_t - A_t) \tag{3}$$

The traders want to update the proposed price by updating the profit margin. Given $p_i(t)$ and a target price $\tau_i(t)$ it is possible to compute $\Delta(t)$ from equation 3,

[4] In some experimental sessions the traders were able to trade more than once, but always for a given number of times. This procedure is useful as it provides a definition of demand and supply schedule.

$$\Delta_i(t) = \beta_i(\tau_i(t) - p_i(t)) \tag{4}$$

and the new profit margin rearranging equation 1:

$$\mu_i(t+1) = \frac{p_i(t) + \Delta_i(t)}{v_i} - 1 \tag{5}$$

The target price is defined using the price of the last shout $q(t)$ in the following way:

$$\tau_i(t) = R_i(t)q(t) + e_i(t) \tag{6}$$

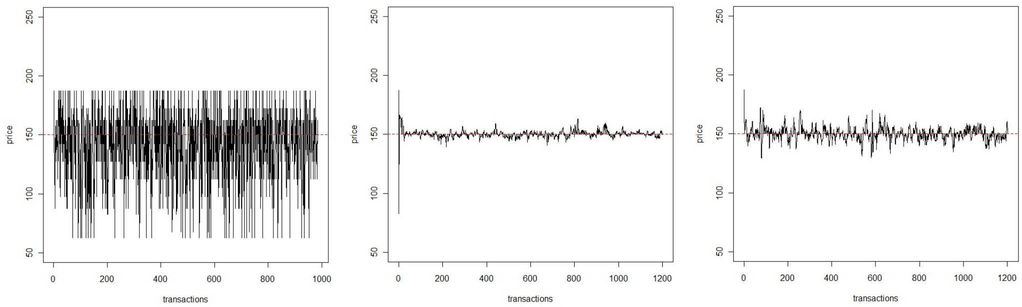
where R_i is a random coefficient and e_i is a random perturbation. If the aim is to increase the last shout, $R_i > 1$ and $e_i > 0$, if the aim is to decrease the last shout $0 < R_i < 1$ and $e_i < 0$. The agents learn about their environment using the orders observed in the market [5]. During each simulated period every agent will issue an order - if active - on average every 20 seconds. One period lasts 500 seconds and trading normally takes place in the first part of the period [6]. The timing has some relevance, as the traders use the last proposal to gain information about the market; if the agents acted simultaneously they would use less orders and the learning mechanism would be more instable and slow. This is one example of how important it is to explicitly model the price formation mechanism since it shapes the market behavior. The asynchronicity of actions is crucial as it allows the traders to understand the market environment. The timing parameter is assumed known and constant, therefore the interesting parameter is β_i, the parameter that formalizes how the traders are learning. The aim of the paper is to check whether it is possible to estimate the learning parameter and in particular define a set of procedures that tell whether it is possible. Note that in the model the β_i parameter is heterogeneous, it means that in a market with 22 traders there would be the need to estimate 22 parameters. Since the number of parameters do not change the overall procedure, to simplify the problem the traders are assumed to learn using the same parameter, i.e. $\beta_i = \beta$. This is a huge simplification if we consider the model as a representation of reality, but it is a non-crucial simplification in the light of estimation techniques. The aim here is to show that estimation of micro-parameters using macro-observations is possible when a set of conditions is satisfied. The conditions do not change with more parameters, the only difference is that it might be more difficult to satisfy them.

[5] In Cliff & Bruten (1997) a further learning procedure was implemented taking into account past Δ_i; since it has a minimal effect on the results it has been eliminated

[6] In Smith's experiment the traders could freely offer and bid; within the time limit of a trading period this procedure was continued until bids and offers were no longer leading to contracts.

2.1 Simulations

The supply and demand curve are shown in figure 1. The initial profit margin are set to $\mu_0 = 0.25$ for the sellers and $\mu_0 = -0.25$ for the buyers. The number of traders is 22, 11 buyers and 11 sellers.



(a) $\beta = 0$ (b) $\beta = 0.5$ (c) $\beta = 1$

Figure 2: Artificial stock market behavior with different values of the learning parameter.

The theoretical equilibrium is 150. In figure 2 three runs of the model are shown with different values for β. The model has an initial converging phase where the agents learn the environment watching the actions of the other agent, the price converges toward the theoretical equilibrium due to the double incentive acting on the agents: increasing their profit margin and at the same time increasing the probability of trade. The length of the converging phase and the behavior around the equilibrium price depends on β. Note that $\beta = 0$ means no learning and has a very different behavior from a learning situation with $\beta > 0$. In the next section the simulation-based econometrics is introduced as a starting point for the estimation of the agents based model.

3 Simulation Based Econometrics

Agent-based modeling is an instrument used to model complex phenomena that involve interactions between the elements of the system under analysis, interaction between the elements and the environments, heterogeneity and so on. Agent-based models are thus used in situations in which the analytical approach is too restrictive to have a good representation of the system. From the previous statement it follows that it is not possible to use standard econometric tools to compare artificial data and real data. Indeed, the complexity of an agent-based model impedes the writing of an analytical condition to find the parameters that minimize a given distance between real data and the model. To overcome this difficulty it is possible to refer to the literature on simulation-based econometric methods. A good reference in this framework is Gourieroux & Monfort (1996) where the most important methods are described: Methods of Simulated Moments (MSM), Indirect Inference and Simulated Maximum Likelihood. In the following, the method of simulated moments will be used; it is an intuitive way to extend simulated econometrics to agent-based models. The simulated methods of moments was introduced by Fadden (1989) and Pakes & Pollard (1989). Duffie & Singleton (1993) apply the method of simulated moments to a markovian process, Lee & Ingram (1991) apply it to time series models.

Suppose that we have a set of observations y_t , a vector of B explanatory variables $\{\mathbf{z_t}\}$ and a vector of K instruments $\{\mathbf{x_t}\}$. Supposing that a well specified model is available, that the data generator process is well behaved, i.e. that $(y_t, \mathbf{z_t}, \mathbf{x_t})$ are jointly ergodic and jointly stationary and that the orthogonality conditions are satisfied, the generalized method of moments estimates the parameters by minimizing $J(\beta, W)$; $J(.)$ is the quadratic form that represents the distance between the theoretical moments and the observed moments (that is between the orthogonality conditions and the sample counterparts), β is the vector of parameters and W is a weighting matrix (see Hayashi (2000)). The method of moments or the general method of moments requires the possibility of computing analytically the theoretical moments; unfortunately such a condition significantly limits the applicability. If the model is complex, it may be impossible to find an analytical form of the conditional moments and thus it may be impossible to find an analytical expression of the quadratic form and of its derivatives. This means that it may be impossible to minimize analytically the objective function. The solution is to simulate the model: if the analytical theoretical moments conditional to the parameters cannot be found, it is possible to simulate the model and compute the moments from the artificial data. The method of simulated moments thus extends the method of moments by replacing the theoretical moments with its simulated counterpart calculated with simulated data (Duffie & Singleton 1993). To estimate the parameters it is sufficient to choose the value of the parameters that minimizes the distance between the simulated moments and the observed moments. The general expression of the objective function to be minimized can be found in Gourieroux & Monfort (1996, p.27), where also the asymptotic properties of the estimator are shown. In particular when the number of observations (n) tends to infinity and the number of simulations (S) is fixed, the estimator is strongly consistent and its distribution tends towards a Normal under regularity conditions in Hansen (1982). The variance of the simulated moments estimator (given the weighting matrix W) decreases when S increases, and tends to be equal to the variance of the GMM estimator when $S \to \infty$ (Gourieroux & Monfort 1996). The extension to agent-based models is straightforward: the artificial data produced by the simulation model are used to compute the simulated moments to be compared with the observed moments.

The aim is to evaluate the estimation procedure applied to a simple agent-based model.

The model used is the one described in the previous section, where the agents are traders that trade for a profit in a continuous double auction. The behavior of the traders is essentially described with the proposed price, that is the maximum price they are willing to pay to buy one asset or the minimum price the are willing to accept to sell one asset. The proposed price changes depending on the behavior of the other agents and the learning mechanism is governed by one parameter. There are no exogenous variables. Despite the simplicity, there is no way of writing any analytical expression of the emergent data (the price) as a function of the behavioral parameter. The objective function to be minimized is:

$$J(\beta, W) = (\mu^R(\beta_0, \epsilon) - \mu^S(\beta))' W (\mu^R(\beta_0, \epsilon) - \mu^S(\beta)) \tag{7}$$

where μ^R is the vector of dimension M containing the chosen M moments computed over the observed data, μ^S is the vector of dimension M containing the M moments computed over the simulated data. The observed moments depend on the "true parameters" β_0 and on a random error due to the sample. The simulated moments depend on the parameters β used in the simulation. The simulated moment is:

$$\mu_m^S = \frac{1}{S} \sum_{s=1}^{S} m(\{y_t\}^n)_s \tag{8}$$

where $m(.)$ is the moment estimator using $\{y_t\}^n$ observations, n is the number of observed data and S is the number of simulations. Equation 8 can be used also with biased moment estimators and when the model is non-ergodic it gives information about the ensemble moments of the model. If $S \to \infty$, the simulated moments tend to the theoretical moments and the MSM estimator tends to the GMM estimator. The variance of the estimator is thus reduced if the number of simulations increases.

The estimated set of parameters is the solution of the minimization of $J(\beta, W)$. Under the regularity conditions defined in Hansen (1982), the values of the parameters resulting from the minimization of equation 7, $\hat{\beta}$, are consistent (Gourieroux & Monfort 1996). The crucial issue for agent-based estimation is to know whether the regularity conditions are actually met. In particular this paper will focus on the consistency and bias of the estimator and thus on the identification problem [7] and on stationarity and ergodicity properties.

4 Properties of the model

As noted in the previous section the crucial part in agent-based models estimation is to analyze the properties of the model itself. The lack of any analytical expression that links

[7] The number of moments has to be greater than the number of parameter, and the chosen moments have to characterize the model using different parameters.

the parameters of the model to the behavior of the emergent property of the system implies the lack of knowledge about the fundamental properties of the model. The tools to be used to increase our knowledge about the behavior of the model are the sensitivity analysis, often used in agent-based modeling literature and statistical tests for stationarity and ergodicity. To show the consistency properties, the sensitivity analysis and the estimation will be made for the case in which 200 days and 1000 days are observed.

4.1 Sensitivity Analysis

The sensitivity analysis is a crucial step in understanding agent-based models. As noted in the introduction, the lack of any analytical representation of the model hides the properties of the system in the system. The sensitivity analysis allows to increase the knowledge about a selected set of properties of the model in response to the change of a set of parameters in the model. The price behavior has to characterize the system using different parameters. The aim of the sensitivity analysis is thus to search for the moments to be used in the objective function. The problems to be faced are the identification problem and the efficiency problem. The moments have to be monotonic in the change of the parameter, i.e. they have to be different for different parameters values. Given this property, the more the moment changes when the parameter changes the more efficient is the estimation, the more different parameters are distinguishable.

To simplify the problem the parameter is considered discrete, the learning parameter β is assumed to be between 0 and 1 and to take only discrete values with 0.05 steps. This is a simplifying hypothesis to reduce the computational burden and it does not change the main properties of the estimator. The resolution of the parameter to be estimated is a choice of the modeler depending on the type of model. Higher resolution needs more computation. The sensitivity analysis has been done by running the model for 100 times for every possible value of the parameter. The tested moments are the mean, the variance, the skewness and the kurtosis. In figure 3 the four moments for the model run over 200 trading days (corresponding to about 1000 transactions). To compute meaningful moments it is necessary to select the stationary part of the process, this can be done using both a qualitative judgmental procedure, looking at the output of the model and selecting the part of it which is stationary and by using the stationarity and ergodicity tests described below. The important role of the moment is to characterize unambiguously the model with a given set of parameters. Thus the moment conditional on the parameters has to be stationary and ergodic, i.e. it has to come always from the same distribution. The moments in the figures are computed taking the observations from 200 onward to allow the process to converge toward its stationary part.

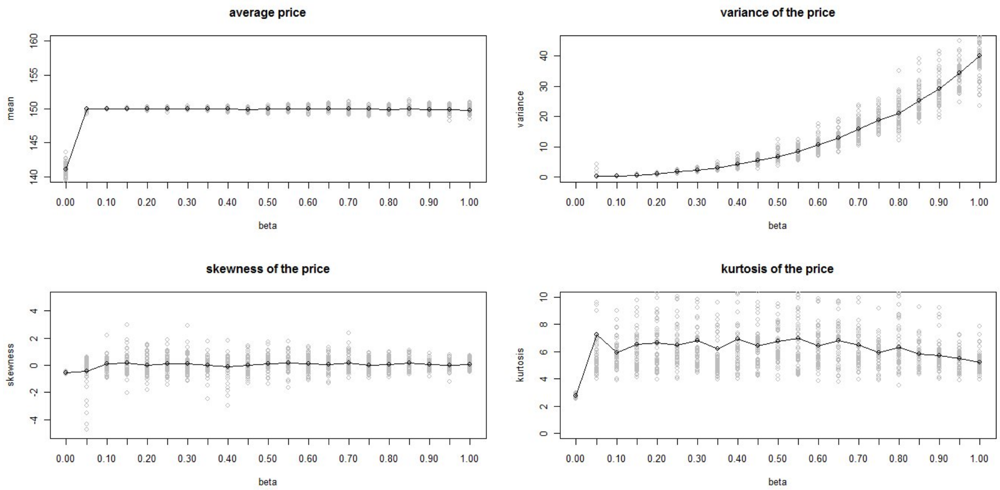


Figure 3: Sensitivity analysis, the effect of changing β on the moments.

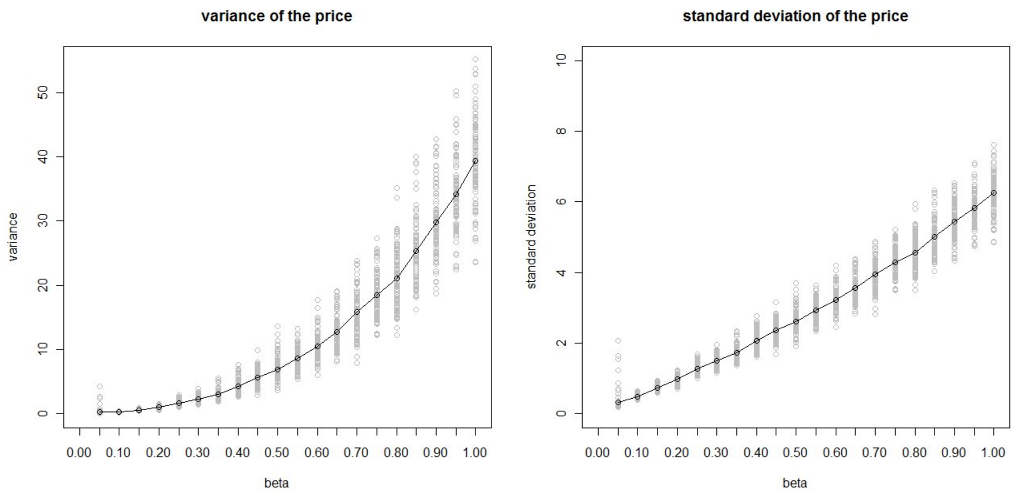
The first important information about the model is that when learning is present, i.e. $\beta > 0$, the equilibrium value toward which the model converges is always the same. The stationarity and ergodicity test in the next section will indeed confirm that for every possible value of β the model produces a stationary and ergodic mean value (i.e. the model has a unique equilibrium). Evidently the mean value of the model cannot characterize the behavior with different values of the parameters. The same argument can be made on skewness and kurtosis. The model is quite symmetric for any value of β and has a slightly decreasing kurtosis from slowest learning, $\beta = 0.05$, to fastest learning, $\beta = 1$. The case of no learning is very different from the rest of the results. The only moment that seems to actually discriminate between different learning parameters is the variance; this will be the moment used in the objective function.

A problem arise regarding the functional form of the chosen moment with respect to the parameter: the variance is non-linear in the parameter and this might create a bias in the estimates.

4.2 Small sample bias

The non linearity of the relation between the moments used for the estimation and the value of the parameters, the "moment function", can create biased estimations. Suppose that $\gamma(\theta)$ is the moment function of the model. It represents the relation between a given moment (chosen for the estimation procedure) and the value of the structural parameter θ in the model. Suppose that the true parameter is θ_0. The observed moment will thus be

$$\gamma(\theta_0, \epsilon)^{obs} = \gamma(\theta_0) + \epsilon \tag{9}$$



(a) Sensitivity on variance (b) Sensitivity on standard deviation

Figure 4: The variance conditional on the parameter value (a) and the standard deviation conditional on the parameter value (b).

The observed moment is the "true moment" plus an error ϵ depending on the sample. Supposing that the underlying process is ergodic and stationary and that the moment estimator is unbiased ($E(\epsilon) = 0$), by applying the ergodic theorem we know that the expected value of the observed moment is equal to the true value, $E[\gamma(\theta)^{obs}] = m(\theta_0)$. It is possible to express the variance of the error term σ_ϵ^2 as a function of the number of observations, with $lim_{n\to\infty}\sigma_\epsilon^2(n) = 0$. The moment estimator, given the assumptions, is unbiased and consistent. The method of simulated moments minimizes the distance between the observed moment and the simulated moment with respect to the parameter:

$$\hat{\theta} = argminJ(\theta) = (\gamma(\theta_0, \epsilon)^{obs} - \gamma(\theta))^2 \tag{10}$$

where $\hat{\theta}$ is the estimated parameter and $\gamma(\theta)$ is the theoretical moment computed from the model [8]. Supposing that the moment function and the parameter are continuous, the method of simulated moments set $\hat{\theta}$ such that[9]:

$$\gamma(\theta_0, \epsilon)^{obs} - \gamma(\hat{\theta}) = 0 \tag{12}$$

using equation 9:

[8] In the simulation based econometric framework, $\gamma(\theta)$ is simulated and might contain an error, this is omitted for simplicity.

[9] In a traditional setting, supposing differentiability of the moment function, $\hat{\theta}$ is selected by setting the first derivative of the objective function to zero:

$$\frac{dJ(\theta)}{d\theta} = 2(\gamma(\theta_0, \epsilon)^{obs} - \gamma(\theta))\frac{d\gamma(\theta)}{d\theta} \tag{11}$$

Since the moment function has to be monotonic to be able to identify the parameters, it is supposed that $\frac{d\gamma(\theta)}{d\theta} \neq 0$ for all θ, the minimization condition thus is equal to equation 12

$$\gamma(\theta_0) + \epsilon - \gamma(\hat{\theta}) = 0 \tag{13}$$

and

$$\gamma(\hat{\theta}) = \gamma(\theta_0) + \epsilon \tag{14}$$

By assumption the moment estimator is unbiased ($E(\epsilon) = 0$), therefore

$$E(\gamma(\hat{\theta})) = \gamma(\theta_0) \tag{15}$$

The moment estimator is unbiased and consistent. The estimation procedure select the estimated parameters such that the expected value of the theoretical/simulated moment is equal to the true moment (given the assumptions above). If the moment function is non-linear, this condition does not imply that the expected value of the estimates is equal to the true parameter. If the moment function is convex:

$$\gamma(E(\hat{\theta})) \leq E(\gamma(\hat{\theta})) = \gamma(\theta_0) \tag{16}$$

which implies that $E(\hat{\theta}) \neq \theta_0$. The direction of the bias depends on the first derivative of the moment function. If $\gamma'(\theta) > 0$, equation 16 implies that $E(\hat{\theta}) \leq \theta_0$, i.e. a downward bias. On the contrary if $\gamma'(\theta) < 0$, the equation 16 implies that $E(\hat{\theta}) \geq \theta_0$, i.e. an upward bias. In the same way it is possible to show that if the moment function is concave, the estimated parameter is upward bias if the moment function is increasing and downward bias if the moment function is decreasing.

The bias can be can be solved by knowing the moment function (which is not the case here, since we are dealing with an agent-based model) or reduced either by applying a monotonic transformation or by increasing the number of observations (i.e. by reducing the variance of the error). To prove the possible presence of a bias, the error has been assumed with zero mean and symmetric. The former assumption is verified when the observed sample moment is not biased with respect to the simulated moment. The latter depends on the observed moment distribution and is not guaranteed. If the the error is non-symmetric the bias can be reduced or increased depending on the direction of the asymmetry. The symmetry of the moment can be tested with the sensitivity analysis.

As evident from figure 4, the moment function is convex, from the sensitivity analysis the skewness is slightly positive. As it will be shown below, taking the square root of the variance (i.e. the standard deviation, see figure 4b) reduces the convexity of the moment function, and this will reduce the bias in the estimates. To define the convexity and to compute the reduction of it after the transformation, a convexity measure is defined. The aim is only to check whether the transformation did actually reduce the non linearity of the moment function.

A function is convex on an interval if for any x and y in the interval and for any $\lambda \in [0, 1]$:

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y) \tag{17}$$

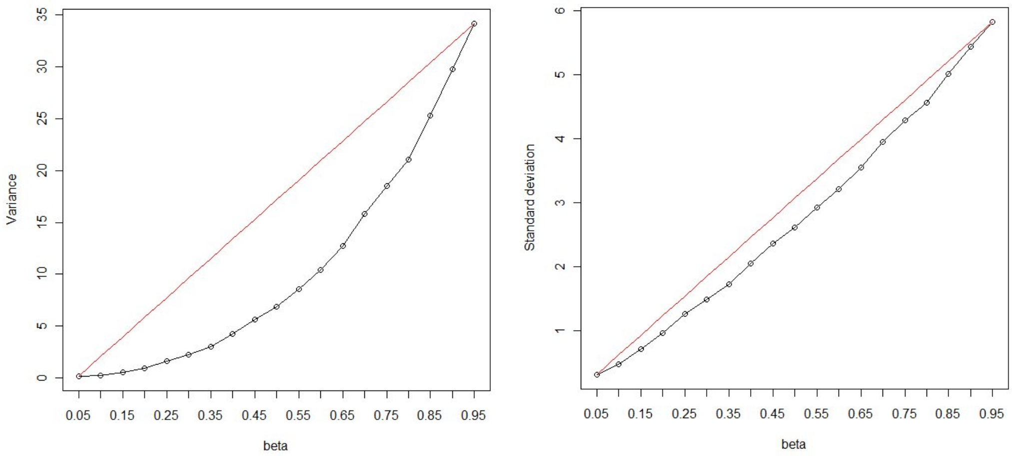
For a concave function the inequality is reversed[10]. Supposing that the chosen interval is $[x_1, x_2]$, a convexity index can be defined as the ratio between the area below the straight line joining $f(x_1)$ and $f(x_2)$ (denoting it $g(x)$) and the area below the function $f(x)$ in the in the interval $[x_1, x_2]$. If $f(x)$ is discrete (as the moment function), the convexity index can be computed as:

$$I = \frac{\sum_{x_1}^{x_2} g(x)}{\sum_{x_1}^{x_2} f(x)} \tag{18}$$

Given the definition in equation 17, the index is $I < 1$ if the function is strictly convex, $I > 1$ if the function is strictly concave and $I = 1$ if the function is linear. Since the average moment for every parameter is a random variable, to rule out the linearity it is possible to use the non-parametric fitness test used in Chapter 2 (Grazzini 2012, Wald & Wolfowitz 1940, Gibbons 1985). The null hypothesis is that the moment function is randomly distributed around the straight line, i.e. that $I \neq 1$ is random. If the null is rejected than it is possible to consider the $I < 1$ or $I > 1$ as a systematic deviation that denotes the shape of the moment function.

In figure 5 the moment function is shown in the interval $\beta \in [0.05, 1.0]$ for variance and standard deviation. The index is $I = 1.615$ for the variance and $I = 1.106$ for standard deviation. Both moment functions are convex but with a significant reduction using the standard deviation. As it will be seen in the next section the concave transformation allows to reduce the bias by a significant amount. The magnitude of the bias depends on the local behavior of the moment function in the neighborhood of the true parameter, where the actual definition of "neighborhood" depends on the variance of the observed moment. In a realistic setting the real parameter is unknown, the sensitivity analysis on the moments is crucial for understanding the results. Once an estimate of the parameter is available it is possible to try to understand the local properties in the neighborhood of the estimate.

[10]Another method to check for the concavity, convexity or linearity of the moment function is to compute the numerical second difference (second derivative).



(a) Variance

(b) Standard deviation

Figure 5: The properties of the estimates depend on the behavior of the moment function in the neighborhood of the moment function. On the left the behavior of the variance in the interval $\beta \in [0.05, 1.0]$ and the straight line joining the 0.05 and 1.0. On the right the behavior of the standard deviation in the same interval. The function is consistently below the straight line denoting convexity.

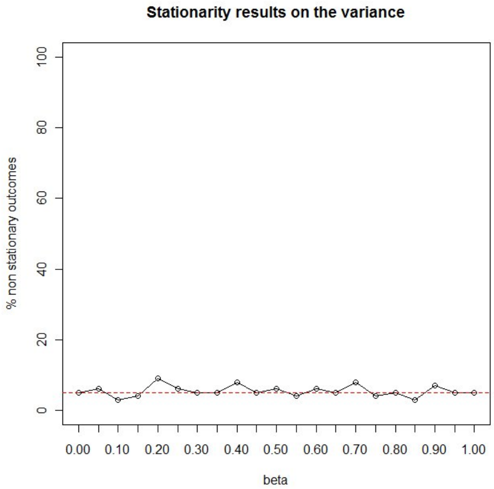
4.3 Tests

The stationarity test and the ergodicity test are fundamental to have a better understanding of the model and to make a statistical statement about the behavior of the moments chosen for the objective function. It possible for example to test the stationarity and ergodicity of the mean of the model to know whether the model has an equilibrium price (stationarity) and whether it is unique (given the value of the parameters) using the ergodicity test Grazzini (2012). During the estimation procedure it is crucial to know whether the moment chosen to estimate the model are ergodic and stationary. The parameter to be estimated is one, therefore there is the need for at least one moment. The sensitivity analysis made in section 4.1 has shown that among the tested moments only the variance is discriminating between different values of the parameters. The test will therefore be carried out on the variance to understand whether it is stationary and ergodic, thus if it possible to use the variance to obtain consistent estimates. The tests are described in (Grazzini 2012) and are essentially an application of the Wald-Wolfowitz non-parametric runs test to test whether the moments are constant in time for stationarity and constant between different runs of the same process for ergodicity. The aim is to test whether the observed moment is a good estimator of the true moment characterizing the model.

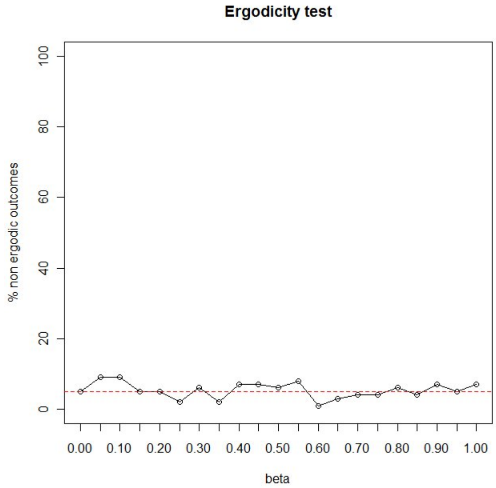
The stationarity test tests whether the given moment is constant in time, the test has been done on the variance, but given the nature of the tests the results can be extended also to the standard deviation. Suppose that a time series with n observations has been created from the stock market model and that observations from transaction 200 have been selected to eliminate the convergence part of the series. The first stationarity test to be done is on the first moment. The time series is divided in w windows, on each windows the first moment is computed. If the time series is stationary in the first moment then the overall mean is a good estimator of the series of windows' first moments. The test is made using the Wald-Wolfowitz non-parametric test that tests whether a given set of observations (the windows moments) are randomly distributed around the fitness function (in this case the overall mean). If the process is strictly stationary then the test cannot reject the stationarity null-hypothesis even if each window have just one observation (if the process is strictly stationary then each observation comes exactly from the same distribution). If the process is not strictly stationary (but still stationary), the windows need to be longer to have a good estimate of the windows moments. The test thus gives also an indication of the type of stationarity. In order to carry out a powerful test the number of windows should be higher than 50 (Grazzini 2012). In the artificial stock market model, the time series are stationary for all possible values of the parameters, but not strictly stationary; this is due to the presence of learning by using the last transaction price. The test has to be done for each value of the parameters since the agent-based model might change its properties when the value of the parameters change. In the case in which the number of parameters is high and/or the parameters have too many possible values, the same test can be done using the estimated value of the parameters and possibly in the neighborhood of it.

The test on the mean is useful to understand the model but not for the estimation procedure, since the moment used in the objective function is the variance [11]. The stationarity test for the variance is done exactly in the same way. Using a type I error $\alpha = 0.05$, the results for all possible values of β are shown in figure 6a. The test is not rejecting the null hypothesis of stationarity. The results were obtained using a time series produced with 3500 trading days (eliminating the first 200 transactions), 55 windows with 400 observations each. The fact that the stationarity test was not rejected using 400 observations tells us that 400 observations are actually enough to compute a good estimation of the true moment. If the observed time series has more than 400 observations it is possible to use it for estimation. The non-parametric test is complementary to the parametric stationarity test Augmented Dickey Fuller. The idea of the non-parametric test is to acknowledge the fact that the understanding of the model is limited due to its complexity. The non-parametric tests have lower power than parametric tests and are therefore difficult to use on the observed time series (that are usually not long enough to ensure full power) and the test is therefore carried

[11] Since the variance estimator is biased for non strictly stationary series, the overall moment has be computed as the mean of the windows moments for theoretical consistency. The result in the particular case analyzed was the same.



(a) Stationarity test results



(b) Ergodicity test results

Figure 6: Stationarity and Ergodicity test for the variance. The test was made on every possible value of β. The dashed line is the theoretical rejection probability ($\alpha = 0.05$). The results are showed as percentage of rejected null hypothesis over 100 experiments.

out on artificial data. The aim is to increase the knowledge about the model, and at the same time, supposing that the model is well specified, the knowledge about the real system.

The stationarity test is needed to test whether the process shows the same behavior *within* the series, while the ergodicity test is done to test whether the model shows the same behavior *between* different series. Given different initial condition and the same parameters, the ergodicity property implies that the time series do not change their properties depending on the initial condition. Taking for example the ergodicity test on the mean, the test tells whether it is possible to reject the null hypothesis that the time series converge always to the same average transaction price, regardless the of initial conditions. The stock market under examination is ergodic in the mean. Ergodicity is thus important to understand the equilibrium properties of the model (whether, given the parameters, the equilibrium is unique) and, like the stationarity test, it is fundamental to understand the behavior of the moment used to estimate, thus to understand the basic properties of the estimate. The results of the ergodicity test in variance are shown in figure 6b; the test is not rejecting the null hypothesis of ergodicity. For details about the tests see Grazzini (2012).

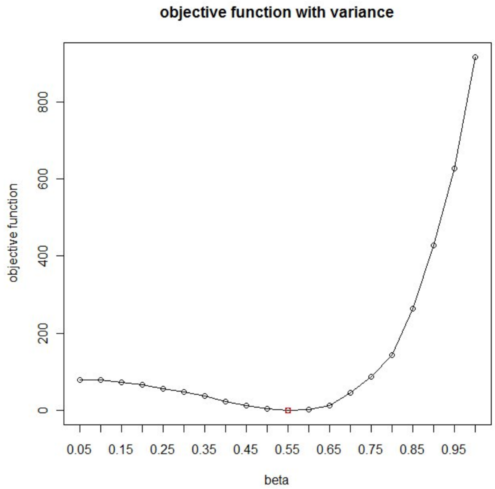
The tests performed on the moment used for estimation is needed to understand the properties of the resulting estimates. If the moment is stationary and ergodic and if the moment is able to identify the parameters then the estimates are consistent; by increasing the number of observations the variance of the simulated moment estimator will decrease. This is the consequence of the ergodic theorem, if a process is stationary and ergodic and if its moments exist then the sample moments will converge toward the true moments as the observations increase.

The preliminary tests in this section and in section 4.2 have shown that biased and consistent estimates have to be expected from the simulation based estimator. The bias might be a problem, but it has been argued that it is possible to reduce it. Moreover the direction of the bias is known despite the limited analytical knowledge available about the model. In the next section these results are confirmed using the model described in section 2 by estimating the behavioral parameter. The Monte Carlo experiments will show that the bias can be reduced using an appropriate transformation of the moment and that the estimates are consistent.

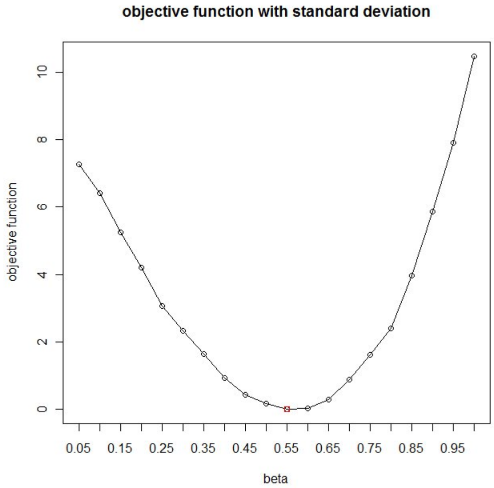
5 Estimates

Given the analysis made in the previous section on the properties of the model it is possible to start the estimation. The estimation will be done as a Monte Carlo experiment to check the properties of the estimates. The simulated moments are computed using the results of the sensitivity analysis. To check the performance of the estimation method a set of runs of the model with $\beta = 0.55$ will be used as "pseudo-real" observations. For every run, the estimation method will produce an estimate of the behavioral parameter. Using the Monte Carlo experiments it is possible to asses the expected value of the estimator and its variance. The aim is to show that the estimation of the micromotives of a system using the macrobehavior of the system is possible, that the estimates are consistent and that using the information about the model it is possible also to build non-parametric confidence interval for the estimated parameters.

A crucial technical point is how to minimize the objective function. The objective function is defined in equation 7, the vector of moments is composed by only one moment and the weighting matrix is simply equal to 1. The lack of an analytical form for the conditional theoretical moment impedes an analytical minimization. One alternative is to use a genetic algorithm, which is efficient and simple to understand. The genetic algorithm was introduced by Holland (1975) and has been widely used in economics in Axelrod (1987), Arifovic et al. (1997), Arifovic (1995), Vriend (2000) as a learning algorithm and in Grazzini (2011) as an optimization heuristic. For details about genetic algorithms see Holland (1975), Goldberg & Holland (1988), Gilber & Troitzsch (2005). Another option is described in Gilli & Winker (2003), who use a combination of the Nelder-Mead simplex direct search method and the threshold accepting optimization heuristic. The advantage of using an optimization heuristic is that it does not need strong hypotheses about the optimization problem, apart from the assumption that a global minimum actually exists and helps in saving computation



(a) Objective function with variance



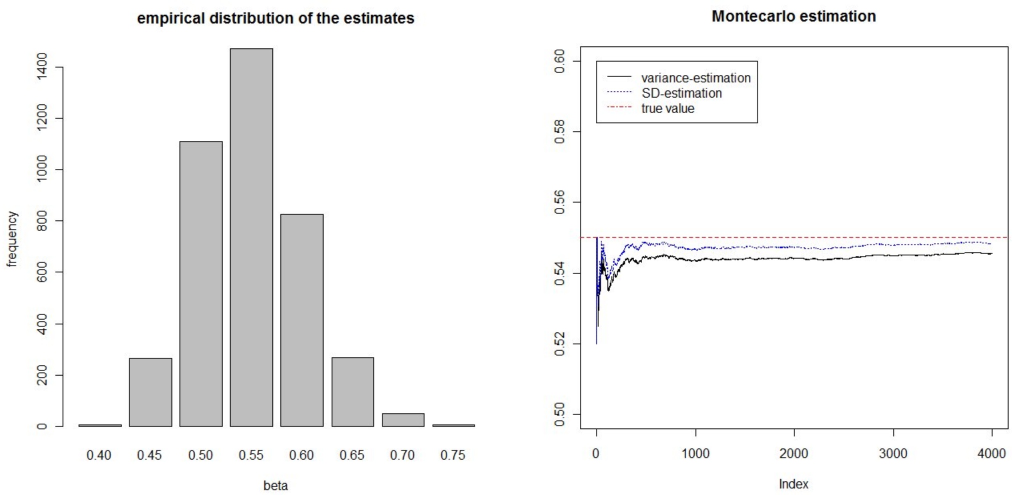
(b) Objective function with standard deviation

Figure 7: Brute force minimization allows to draw the objective function. In panel (a) the objective function using the variance. In panel (b) the objective function using the standard deviation. Both objective functions are drawn using the same observed moment.

time. On the other side heuristics cannot produce high-quality solutions with certainty (Gilli & Winker 2008).In the following the brute force approach is used in order to obtain a good evaluation of the estimation procedure. The brute force method simply means to compute the simulated moments for any possible value of the parameters (note that the parameters are discrete by definition in a computer) and then select the parameters that minimize the defined distance between observed and simulated moments. The (not small) disadvantage is that the computational time increases with the increase of the number of parameters (all the combinations are needed) and with the resolution of the parameter. The advantage is that the objective function can be drawn since its value is known for any combination of the parameters. This is useful in order to avoid local minima, check for uniqueness of the global minimum and to be certain that the selected values for the parameters actually globally minimize the objective function. When the computational time is reasonable, the brute force is the best minimization method. In figure 7 examples of an objective function drawn using the variance (on the left) and the standard deviation (on the right) using a random run of the model with $\beta = 0.55$ and using as simulated moments the average moments computed during the sensitivity analysis.

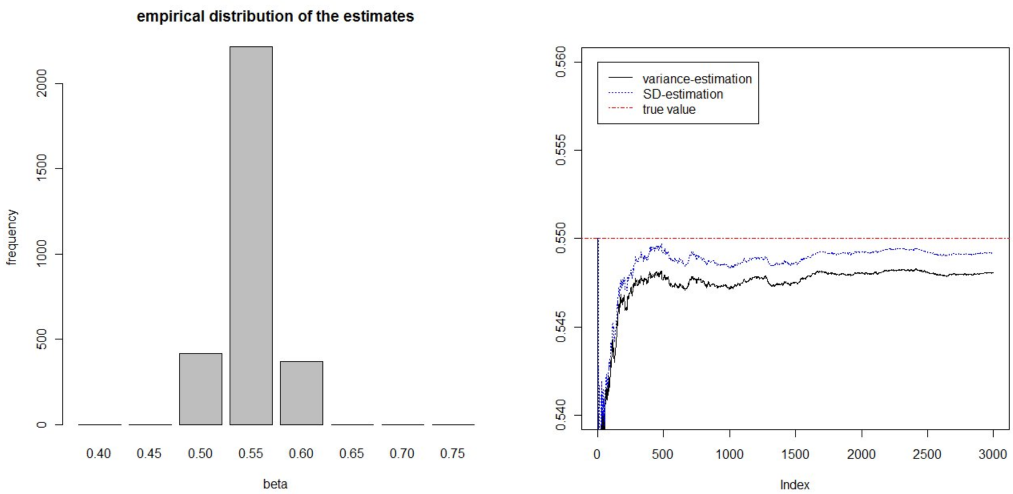
The base assumption, which makes the estimation meaningful, is that the model is well-specified. During the following experiments this assumption will be true by construction since a run of the model is used to produce the pseudo-real observations. Once the hypothesis of well specification is satisfied, the analysis in the previous sections allows an interpretation of the results as consistent and possibly unbiased estimations. The pseudo real parameter used is $\beta = 0.55$, the experiments show the behavior of a set of estimates. Given the sensitivity analysis the simulated moments can be computed. The Monte Carlo experiments consist in creating a number of time series using the model and the pseudo-real parameter, computing the pseudo-real moment chosen in the objective function, and estimating the parameter by choosing the value that minimizes the objective function. Since the parameter is known it is possible to analyze the properties of the method, weaknesses and strengths. Figure 8b shows the result of the Monte Carlo experiments with a pseudo real time series of 200 days (the number of observations is about 1000 transactions). The transformation of the moment discussed in the previous section actually reduces the bias. The average value of 4000 estimations is 0.5455 (the variance of the estimates is 0.00291) using the variance in the objective function, and 0.5485 (the variance is 0.00292) using the standard deviation. The difference between the estimations using the different moment is entirely due to the different objective function. In each experiment the same time series is used both to estimate the parameter with the variance and with the standard deviation, therefore the difference between the results has no noise. In figure 8a the empirical distribution of the estimates using the standard deviation is shown. The mode of the distribution corresponds to the "pseudo-real" parameter, the distribution is slightly skewed to the left. The empirical distribution can be used to build a confidence interval that exclude the 5% tail-observations without the need to know the exact distribution of the estimates. With 4000 estimates the confidence interval is built by excluding 100 observations on the left and 100 observations on the right leaving $\beta_{int} = [0.45, 0.65]$. In this particular case the true parameter is known and the Monte Carlo experiments are performed by using the true parameter to generate the observed data. In a realistic setting it is possible to estimate the parameter of interest and build the empirical distribution by performing the estimation Monte Carlo running the model with the estimated parameter.

The consistency of the estimator can be seen from figure 9, where the Monte Carlo experiments have been performed using a pseudo-real time series with 1000 trading days, which in the model correspond to about 6000 transactions. The average estimation over 3000 estimates is 0.5481 (the variance is 0.00067) using the variance in the objective function, and 0.5492 (0.00066) using the standard deviation in the objective function. The bias is reduced both by transforming the moments reducing the convexity and by increasing the number of observations. The estimates are consistent, by observing longer time series it is possible to estimate the observed moment with smaller noise, thus reducing the variance of the estimates. The empirical distribution of the estimates is shown in figure 9a



(a) Empirical distribution of the estimates (b) Cumulative mean of the estimates

Figure 8: The estimation properties with 200 observed trading days. In figure (a) the empirical distribution of 4000 estimations exercise, the true parameter is $\beta = 0.55$. In figure (b) the cumulative mean after each experiment is shown. The average estimated value converges quickly toward its mean. The estimates are biased.



(a) Empirical distribution of the estimates (b) Cumulative mean of the estimates

Figure 9: The estimation properties with 1000 observed trading days. In figure (a) the empirical distribution of 3000 estimations exercise, the true parameter is $\beta = 0.55$. In figure (b) the cumulative mean after each experiment is shown. The average estimated value converges quickly toward its mean. The estimates have a smaller bias and variance.

6 Conclusions

The paper shows that provided that a set of properties are satisfied by the model (and in turn by the real system supposing well specification) it is possible to consistently estimate parameters at the micro level of the system using the information from the macro-behavior. Agent-based models are built imposing a set of behavioral hypotheses on the agents and a set of institutional hypotheses that govern the way the agents interact. Given such a structure by running the model it is possible to observe how the system behaves and how the emergent properties emerge. The crucial point is that given the series of tests it is known whether the estimates are consistent, biased and even the direction of the bias. Agent-based models are used to relax at least some of the many hypotheses needed to build analytically solvable models, allowing for more realistic models. Unfortunately more realistic models often share with reality non-ergodic (and/or non stationary) behaviors. This would impede the estimation of the model, not due to a flaw in the model itself but to the ontology of the system under analysis. Even in such a case it might be useful to bring the model to the data using the method of simulated moments as a calibration method. The minimization of an objective function is always possible, the tests are important in order to obtain a correct interpretation of the resulting parameters and of the model itself. In this paper the method of simulated moments has been investigated. One interesting alternative is the Indirect Inference method (Gourieroux & Monfort 1996), especially in the view of comparing computational models with analytical models. Supposing that a model reasonably explains the real system, it is possible to define a meta-model and estimate the meta model using both the artificial data and the real data [12]. The results of the estimation of the meta model on the artificial data will depend on some structural parameters in the computational model. By minimizing the distance between the meta model estimates on artificial and observed data by systematically change the structural parameters in the computational model, it is possible to have consistent estimates of the structural parameters. The meta model is not well specified by definition, and the efficiency depends on the chosen meta model; the method of the simulated moment was therefore chosen in this paper. The indirect inference method can be very useful when a data set is not available (e.g for propriety reasons), but other studies using that data set are available in the literature. It is possible to use the model estimated in the literature as a meta model and try to estimate the computational model using indirect inference. This method can be very useful also to compare the computational model with other analytical models and to create stronger links with the literature.

For future research it is essential to keep extending tests, methods and theories from

[12]To be clear: it is possible to use the same (for example) linear model both on the artificial data and on the observed data.

the vast econometric literature to computational econometrics. The empirical foundation of agent-based models is an important step for building new and more complex/ realistic models. The econometric literature is an immense source of ideas that can be carefully adapted to the new more complex environment.

References

Alfarano, S., Lux, T. & Wagner, F. (2005), 'Estimation of agent-based models: The case of an asymmetric ', *Computational Economics* **26**, 19–49.

Arifovic, J. (1995), 'Genetic algorithms and inflationary economies', *Journal of Monetary Economics* **36**, 219–243.

Arifovic, J., Bullard, J. & Duffy, J. (1997), 'The transition from stagnation to growth: An adaptive learning approach', *Journal of Economic Growth* **2**(2), 185–209.

Arthur, W. B., Holland, J. H., LeBaron, B., Palmer, R. & Tayler, P. (1997), *The Economy as an Evolving Complex System II*, Addison-Wesley, chapter Asset Pricing under Endogenous Expectations in an Artificial Stock Market.

Axelrod, R. (1987), *Genetic Algorithms and Simulated Annealing*, : Pitman e Los Altosy, chapter The Evolution of Strategies in the Iterated Prisoner's Dilemma.

Axtell, R. (2000), 'Why agents? on the varied motivations for agent computing in the social sciences', *Center on Social and Economic Dynamics Working Paper* (17).

Bianchi, C., Cirillo, P., Gallegati, M. & Vagliasindi, P. A. (2007), 'Validating and calibrating agent-based models: A case study', *Computational Economics* **30**, 245–264.

Boswijk, H. P., Hommes, C. H. & Manzan, S. (2007), 'Behavioral heterogeneity in stock prices', *Journal of Economic Dynamics and Control* **31**, 1938–1970.

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

Chen, S.-H., Chang, C.-L. & Du, Y.-R. (2009), 'Agent-based economic models and econometrics', *The knowledge Engineering Review* pp. 1–46.

Cliff, D. & Bruten, J. (1997), 'Minimal-intelligence agents for bargaining behaviors in market based environments', *HP Laboratories Bristol* (HPL-97-91).

Cross, R., Grinfeld, M., Lamba, H. & Seaman, T. (2005), 'A threshold model of investor psychology', *Physica A* **354**, 463–478.

Duffie, D. & Singleton, K. J. (1993), ‘Simulated moments estimation of markov models of asset prices’, *Econometrica* **61**(4), 929–952.

Fadden, D. M. (1989), ‘A method of simulated moments for estimations of discrete response models without numerical ’, *Econometrica* **57**(5), 995–1026.

Gibbons, J. D. (1985), *Nonparametric Statistical Inference*, second edn, Marcel Dekker Inc., New York.

Gilber, N. & Troitzsch, K. G. (2005), *Simulation for the social scientist*, second edn, University of Michigan Press.

Gilbert, N. (2001), *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*, Elsevier, chapter Holism, Individualism and Emergent Properties. An Approach from the Perspective of Simulation.

Gilli, M. & Winker, P. (2003), ‘A global optimization heuristic for estimating agent based models’, *Computational Statistics and Data Analysis* **42**, 299–312.

Gilli, M. & Winker, P. (2008), ‘Review of heuristic optimization methods in econometrics’, *Comisef Working Paper Series* .

Goldberg, D. E. & Holland, J. H. (1988), ‘Genetic algorithms and machine learning’, *Machine Learning* **3**(2-3).

Gourieroux, C. & Monfort, A. (1996), *Simulation-Based Econometric Methods*, Oxford University Press, New York.

Grazzini, J. (2011), Consistent estimation of agent based models, LABORatorio R. Revelli Working Papers Series 110, LABORatorio R. Revelli, Centre for Employment Studies.

Grazzini, J. (2012), ‘Analysis of the emergent properties: Stationarity and ergodicity’, *Journal of Artificial Societies and Social Simulation* **forthcoming**.

Hansen, L. P. (1982), ‘Large sample properties of generalized method of moments estimator’, *Econometrica* **50**(4), 1029–1054.

Hayashi, F. (2000), *Econometrics*, Princeton University Press, Priceton.
URL: *http://www.worldcat.org/isbn/0691010188*

Holland, J. H. (1975), *Adaptation in Natural and Artificial Systems*, Open University Press, Ann Arbor.

Kirman, A. (1993), 'Ants, rationality, and recruitment', *The Quarterly Journal of Economics* **108**(1), 137–56.
URL: *http://ideas.repec.org/a/tpr/qjecon/v108y1993i1p137-56.html*

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

Leombruni, R. & Richiardi, M. (2005), 'Why are economists sceptical about agent-based simulations?', *Pshysica A* **3559**, 103–109.

Lux, T. & Marchesi, M. (1999), 'Scaling and criticality in a stochastic multi-agent model of a financial market', *Nature* **397**, 498–500.

Pakes, A. & Pollard, D. (1989), 'Simulation and the asymptotics of optimization estimators', *Econometrica* **5**, 1027–1057.

Richiardi, M., Leombruni, R., Saam, N. J. & Sonnessa, M. (2006), 'A common protocol for agent-based social simulation', *Journal of Artificial Societies and Social Simulation* **9**(1), 15.
URL: *http://jasss.soc.surrey.ac.uk/9/1/15.html*

Smith, V. L. (1962), 'An experimental study of competitive market behavior', *The Journal of Political Economy* **70**(2), 111–137.

Vriend, N. (2000), 'An illustration of the essential difference between individual and social learning', *Journal of Economic Dynamics and Control* **24**, 1–19.

Wald, A. & Wolfowitz, J. (1940), 'On a test whether two samples are from the same population', *The Annals of Mathematical Statistics* **11**(2).

Windrum, P., Fagiolo, G. & Moneta, A. (2007), 'Empirical validation of agent-based models: Alternatives and prospects', *Journal of Artificial Societies and Social Simulation* **10**(2), 8.
URL: *http://jasss.soc.surrey.ac.uk/10/2/8.html*