Working Paper no. **118**

**Indirect estimation of agent-based models.
An application to a simple diffusion model.**

Jakob Grazzini
University of Turin

Matteo G. Richiardi
University of Turin, and LABORatorio R. Revelli, Turin

Lisa Sella
Ceris - CNR, Turin

January, 2012

Indirect estimation of agent-based models. An application to a simple diffusion model.

Jakob Grazzini^{*a}, Matteo Richiardi^{†a,b}, and Lisa Sella^{‡c}

^aUniversity of Turin, Department of Economics, via Po 53, 10124 Torino.

^bLABORatorio R. Revelli, Collegio Carlo Alberto, via Real Collegio 30, 10024 Moncalieri, Torino

^cCeris - CNR, via Real Collegio 30, 10024 Moncalieri, Torino

January 12, 2012

Abstract

Starting from an agent-based interpretation of the well-known Bass innovation diffusion model, we perform a Montecarlo analysis of the performance of a method of simulated moment estimator. We show that nonlinearities of the moments lead to a small bias in the estimates in small populations, and prove that our estimates are consistent and converge to the true values as population size increases. Our approach can be generalized to the estimation of more complex agent-based models. However, a trade-off emerges between model inadequacy and data inadequacy. This is particularly severe when only aggregate information is available, as common with diffusion data.

Keywords: diffusion model, method of simulated moments, estimation

JEL codes: C15 (Statistical Simulation Methods: General), C53 (Forecasting Models; Simulation Methods), C63 (Computational Techniques; Simulation Modeling), D12 (Consumer Economics: Empirical Analysis), O33 (Technological Change: Choices and Consequences; Diffusion Processes)

1 Introduction

In this paper we show how simulation-based econometric techniques can be used for the estimation of agent-based (AB) models. The well-known Bass model of innovation diffusion (Bass,

^{*}Email: jakob.grazzini@unito.it

[†]Email: matteo.richiardi@unito.it

[‡]Email: lisa.sella@unito.it

1969) is reinterpreted as an AB model where each agent's probability of adoption depends on the number of linked agents that have already adopted. In this perspective, the Bass model turns out to be a special case of the AB model, where the underlying network of agents is fully connected, and the literature on the estimation of the Bass model can be used as a benchmark to evaluate the estimation procedure of the AB model.

The structural estimation of AB models is an open area of research. Having an analytical solution is of course a desirable property of a theoretical model, but it often limits its specification to assumptions that are too unrealistic, or rules out the analysis of empirically relevant dynamics. As Albert Einstein once put it, 'a model should be kept as simple as possible, but not simpler than that'. When the simplifications needed for obtaining a closed form solution are considered to be too strong, researchers have to go for a simulation model. In particular, the higher freedom in model specification of AB models with respect to analytical models allows to take into account to a higher extent the complexity of the real world (Tesfatsion and Judd, 2006; Miller and Page, 2006). More complex models can account for bounded rationality, learning mechanisms, networks, interactions, and feedbacks. Such models can reproduce the emergence of system-level properties as a consequence of the actions and interactions of the agents at the microlevel.

However, a common critique addressed towards AB models is that they often remain at a theoretical level, and lack a sound empirical grounding (Gallegati and Richiardi, 2009). When present, this is often limited to some ad-hoc calibration of the relevant parameters. However, estimation is crucial for the empirical validation of the model, for comparing the model with other available models, and for policy analysis. Only a handful of examples exist on the structural estimation of AB models. Winker and Gilli (2001) and Gilli and Winker (2003) estimate respectively 2 and 3 parameters of an AB model of the foreign exchange market introduced by (Kirman, 1991, 1993), by employing the method of simulated moments (MSM). Their focus is on optimization heuristics. In Winker et al. (2007) they deal with the problem of moments selection, and propose a set of statistics on exchange rate returns to estimate models of exchange rate. Boswick et al. (2007) estimate, by means of a non-linear least square method, a dynamic asset pricing model characterized by agents with heterogeneous beliefs. Dancik et al. (2010) use a Gaussian Process emulator of scalar computer model output for sensitivity analysis, (Bayesian) calibration, and model comparison. Their methodology is relevant for models that are expensive to run, in money or time, and for which the number of possible evaluations is therefore limited.

Finally, Alfarano et al. (2005, 2006) estimate agent-based models that are simple enough to derive a closed form solution for the distribution of relevant statistics.

The main reason for this paucity is that, even if AB models can be regarded as a set of mathematical equations (Leombruni and Richiardi, 2005), their properties remain hidden in the complexity of the relations among the many elements in the model. The lack of an analytical formalization linking the behavior of the agents with the outcome of the system impedes a traditional approach to model estimation, and calls for computational methods. These methods, known as simulation-based estimation techniques (Stern, 1997, 2000), have been originally developed in the econometric literature to deal with analytical models leading to criterion functions without simple analytical expression (for instance because of integrals of large dimensions in the probability density function or in the moments). Their application to AB models, however, is not straightforward and, to the best of our knowledge, no simple examples exist of the structural estimation of an AB model.

The purpose of this work is to illustrate the many issues that have to be dealt with when performing such an exercise. The choice of our illustrative example is motivated by three main reasons. First, innovation and diffusion dynamics are a central area of application of AB models (Bonabeau, 2002). Second, there is a simple analytical benchmark against which to evaluate our computational estimation procedure. Third, the standard estimation approach suffers however from a number of shortcomings, which makes the exercise far from trivial. In our sample application, we reduce the complexity of the problem by estimating two parameters with a traditional, maximum likelihood (ML) approach, conditional on a third parameter which is then estimated by MSM. We show how a proper choice of the estimation procedure leads to consistency (which is *not* obtained in most estimates of the standard Bass model available in the literature), and how nonlinearities in the moments lead to a small sample bias, which is however of predictable direction.

The paper is structured as follows. Section 2 and 3 describe the AB diffusion model, vis-à-vis the original Bass model. Section 4 reviews the existing empirical representations of the Bass model, and the related estimation strategies. Section 5 describes our approach to the estimation of the influence parameters p and q when the market potential m is known. Section 6 presents our MSM estimator for m and shows the results of a Montecarlo analysis of its performance. Section 7 discusses the origin and direction of the bias. Section 8 offers our conclusions.

2 The model

The diffusion model takes into consideration the dynamics of adoption of a given technology or product among a population of agents. The probability of adoption depends on peer pressure and increases as the number of peers (linked agents) that have adopted increases:

$$h_i(t) = p_i + q_i A_i(t) \quad (1)$$

where p_i is the probability of an independent adoption and q_i measures the strength of the peer pressure, with $p_i + q_i \leq 1$. Defining $N_i(t)$ as the number of peers of agent i who have adopted at time t , and m_i as the total number of peers of agent i , $A_i(t) = \frac{N_i(t)}{m_i}$ is the share of peers that have adopted. The parameters shaping the diffusion dynamics are p_i, q_i and m_i , i.e. the three parameters defining the probability of adoption for each agent. Potentially, agents can be characterized by heterogeneous parameters, but the consequently high number of parameters would impede estimation, therefore in the following we impose $p_i = p$ and $q_i = q$. The number of peers for each agent depends on the network structure underlying the model. In each period (time is discrete) each agent adopts with the probability defined in eq. 1. At the beginning the adoption is slow since the number of agents that have already adopted is small and therefore the interaction term is negligible. Once the number of adopters starts to increase, the probability of adoption for those who have not already adopted (the population at risk) increases and the diffusion gets faster. As the population at risk gets smaller, the number of new adopters decreases until the diffusion process is completed. The diffusion dynamic follows a typical S-curve (figure 1). The model described here is very similar to the original Bass model, but for the fact that it allows more general network structures. In fact, in the Bass model a fully connected network is assumed, i.e. every agent is connected to all other agents. This assumption implies that all agents have the same peer group; therefore, $N_i(t) = N(t)$ and $m_i = m$. Hence, the Bass model can be considered as a special case of this AB model, where all agents have the same adoption probability. Equation 1 can then be rewritten as:

$$h(t) = p + qA(t) \quad (2)$$

where $A(t)$ corresponds to $F(t)$ in the Bass model, i.e. the cumulative share of adopters in the whole network. In this special case, the parameters to be estimated are p , q , and m . In the

Figure 1: The diffusion dynamic in a fully connected network with 1,000,000 agents, $p = 0.03$ and $q = 0.4$.

more general case, additional parameters governing the network structure have to be estimated.

3 More on the Bass Model

The Bass model (Bass, 1969), which provides a mathematical explanation of the different stages of product adoption described in the seminal work by Everett Rogers (Rogers, 1962) (innovators, early adopters, early majority, large majority, and laggards), and formalizes the crucial distinction between innovators and imitators, is considered as one of the most important empirical generalization in marketing, and it is widely used in sales and technology adoption analysis.

The model is an example of early epidemic models of innovation diffusion (Geroski, 2000). It consists of a differential equation that specifies the rate of adoption $h(t)$ as a function of an external force and an internal (endogenous) one. The external influence is constant over time and represents the effects of advertisement, while the internal influence depends on how many others have already adopted at time t and formalizes word-of-mouth:

$$h(t) = p + qF(t) \tag{3}$$

where $F(t) = N(t)/m$ is the c.d.f. of adopters, that is the ratio of those who have already adopted ($N(t)$) over the number of potential adopters (the market potential m). p is the

parameter for external influence and q is the parameter for internal influence, with $p + q < 1$.¹

As pointed out by Boswijk and Franses (2005), the importance of the Bass model is twofold. First, it provides a tool for forecasting new adoptions, with parameters that are estimated either on an early stage of the diffusion process, or on data from similar products or from other geographical areas. Second, it allows to study the correlation of the key parameters estimated at a country level with country characteristics, in order to gain insights on the determinants of the diffusion process.

4 Estimation of the Bass model

Being a deterministic macro-model cast in continuous time, the Bass model requires some adaptation before being taken to the data, requiring different empirical translations leading to different statistical behavior, and therefore to different choices and properties of the estimators. The related literature has converged on the approach of generalizing the original Bass differential equation into a stochastic differential equation, which is then discretized into a stochastic difference equation and used for estimation, generally by means of nonlinear methods. The underlying idea is that (a discrete-time version of) the Bass equation well describes an idealized aggregate adoption curve, to which a stochastic error term ϵ_t is added in order to take into account the imperfections of the real world, of the data, and of the model specification. In most applications the Bass model is manipulated in order to derive theoretical implications on the number of observed adoptions (say: sales) in each period. The empirical specification is therefore in the form:²

$$n_t \equiv N_t - N_{t-1} = g(N_{t-1}) + \epsilon_t \quad (4)$$

where n_t is the observed number of new adoptions between $t-1$ and t and g is some (nonlinear) function of the observed number of cumulative adoptions N_{t-1} and the parameters, which are generally estimated by some nonlinear technique as nonlinear least squares (NLS).

This approach however runs into two main problems. First, as acknowledged by Bass (1969) himself, a bias is introduced because discrete time series data are used for estimat-

¹This specification of the hazard function had already been introduced to characterize innovation diffusion processes prior to Bass' work (Coleman, 1964; Mansfield, 1961). However, empirical applications were scant, because knowledge of the number of potential (and ultimate) adopters m was required to compute $F(t)$. Bass contribution was to express the adoption ceiling as a parameter, which could be estimated together with p and q using aggregate sales data.

²We use $x(t)$ to denote a continuous variable in a theoretical relationship, and x_t its discrete counterpart used in the empirical specifications.

ing a continuous-time model. Second, the use of NLS generally leads to biased estimates, and the characteristics of the bias are local, i.e. they depend on the true but unknown values of the parameters (Ivanov, 1997; van den Bulte and Lilien, 1997).

Moreover, consistency is also questioned. Of course, the market potential m is always estimated consistently, given that the diffusion process is finite and all potential adopters eventually adopt: if one waits until the process has terminated, the market potential can simply be observed in the data. However, the other parameters do not converge to the true value, no matter the length of the observation period or the frequency of the observations. In fact, the finiteness of the process imposes an intrinsic limit on the amount of information that can be collected: extending the observation period beyond market saturation brings no additional benefit (Boswijk and Franses, 2005). In addition, when the estimation of the market potential is concerned, it is not possible to know in advance if the flow of new adoptions has effectively and completely dried out, hence the period when market saturation is attained is identifiable only after some time has passed. Moreover, it is generally of more use to estimate the parameters at an earlier stage of the diffusion process.³

4.1 Empirical representations of the Bass model

From eq. 3, Bass (1969) derived

$$\begin{aligned}
\frac{dN(t)}{dt} &= mf(t) = h(t)[m - N(t)] \\
&= [p + qF(t)][m - N(t)] \\
&= \left[p + q \frac{N(t)}{m} \right] [m - N(t)] \\
&= pm + (q - p)N(t) - \frac{q}{m}N^2(t)
\end{aligned} \tag{5}$$

on which he based his empirical formulation:

$$n_t = \alpha_1 + \alpha_2 N_{t-1} + \alpha_3 N_{t-1}^2 + \epsilon_t \quad (\text{Bass})$$

assuming homoscedastic and independent errors: $Var(\epsilon_t) = \sigma^2$; $Cov(\epsilon_t, \epsilon_s) = 0 \forall t \neq s$. Bass proposed to estimate this equation by OLS.

³To this regard, a problem that is reported with the estimation approaches proposed in the literature is that the estimated ceiling is often close to the cumulative number of adopters in the last observation period, and it keeps increasing as the observation period is extended, thus pointing at a systematic underestimation of this parameter (van den Bulte and Lilien, 1997). This reverberates on the estimates of the influence parameters: in particular, for any given flow of adoptions $\{N(t)\}$ the estimated coefficient of social influence q decreases as the size of the population at risk m increases, and therefore it tends to decrease as later observations are added.

A nice feature of this approach is that it is not necessary to know the exact timing of the introduction of the new product/technology. Only knowledge of cumulative sales for the estimation period is required. On the other hand, Schmittlein and Mahajan (1982) identified a main shortcoming of this approach in the time-interval bias, which originates because discrete time series data are used for estimating a continuous-time model: rather than the derivative of $N(t)$, its discrete counterpart n_t is used. This substitution causes a problem in that n_t , as defined, will underestimate $dN(t)/dt$ for time intervals before the maximum adoption rate is reached, and it will overestimate after that point.⁴

The subsequent literature has focused on improving on the OLS estimator by means of nonlinear techniques. Srinivasan and Mason (1986) and Jain and Rao (1990) apply NLS to estimate two slightly different empirical formulations of the Bass model:

$$n_t = m(F_t - F_{t-1}) + \epsilon_t \quad (\text{Srinivasan-Mason})$$

$$n_t = (m - N_{t-1})(F_t - F_{t-1})/(1 - F_{t-1}) + \epsilon_t \quad (\text{Jain-Rao})$$

The error term is still white noise.

In a more recent paper, Boswijk and Franses (2005) allow random deviations in n_t from its mean path in order to display mean-reverting serial correlation and incorporate heteroscedastic errors, accounting for the fact that uncertainty on the diffusion path is not constant over time. In fact, they argue that it is lower at the beginning and toward the end of the diffusion process, while it is higher in the middle. They propose the following specification:

$$dn(t) = \alpha[n^*(t) - n(t)]dt + \sigma n(t)^\gamma dW(t) \quad (6)$$

where $n^*(t) = p(m - N(t)) + \frac{q}{m}N(t)(m - N(t))$ is the expected number of new adopters at time t , given the actual number of cumulative adopters⁵, $\alpha > 0, \sigma > 0, \gamma \geq \frac{1}{2}$ and $W(t)$ is a standard

⁴Bass recognized this in his article, and suggested a method for reducing the problem. “However, a substantial amount of bias is likely to remain since his correction procedure makes the unrealistic assumption that the evolution of sales over time resembles the exponential density function” (Srinivasan and Mason, 1986). It should be noted that all the empirical specifications based on discrete data proposed in the literature, including those of Srinivasan and Mason (1986), Schmittlein and Mahajan (1982) and ours, incur in this problem.

An additional problem of the Bass approach that is not shared by more recent approaches is that it does not directly provide standard errors for the estimated parameters.

⁵This differs from the expected number of new adopters at time t , given the expected number of cumulative adopters, which corresponds to the deterministic $\frac{dN(t)}{dt}$ in eq. 5

Brownian motion. This leads to the empirical specification:

$$n_t - n_{t-1} \approx \alpha \left[p(m - N_{t-1}) + \frac{q}{m} N_{t-1}(m - N_{t-1}) - n_{t-1} \right] + n_{t-1} \epsilon_t$$

(Boswijk-Franses)

where $\epsilon_t = \sigma(W_t - W_{t-1}) \sim i.i.d.N(0, \sigma^2)$.

Unfortunately, this approach leads again to biased and inconsistent estimates. The reason ultimately lies in the combination of an additive error term with a nonlinear estimation strategy, which requires consistency to be looked for in a time-series perspective, for $t \rightarrow \infty$. Given that the process is by definition finite, consistency is unattainable.

The approach that is closer to ours is that of Schmittlein and Mahajan (1982). They also interpret equation 3 in terms of the distribution of adoption times: individual adoption times are considered as independent draws from that distribution.⁶ If potential adopters are i.i.d. actors, then the number $n(t)$ of new adoptions in each period follows a multinomial distribution whose probabilities are specified by the Bass model. It is then possible to estimate the parameters of this distribution by ML, and the estimates tend to the true value of the parameters for increasing population size.

The advantage of a distribution approach like Schmittlein and Mahajan (1982) and ours is that consistency of the estimates depends on population size, and not on the length of the observation period. The disadvantage is that it only assumes sampling errors. However, the model might be at least partly misspecified, for instance because some explanatory variables (be they individual –observed or unobserved– or aggregate variables) are omitted. In this case an empirical specification derived from such a theoretical model would result in standard errors of the coefficients that are too low in order to reproduce the real uncertainty regarding the evolution of the diffusion processes (Srinivasan and Mason, 1986). By including an *ad hoc* additive error term, however, the standard approach departs in no minor way from the theoretical model, resulting in biased and inconsistent estimates of the structural parameters.⁷

Our simulation approach involves a three-stage procedure. First, we estimate the external influence parameter p , conditional on a specific value of the market potential parameter m .

⁶The assumption of independence becomes less and less appropriate as the time intervals grow bigger. Hence, this approach too is subject to the time-interval bias.

⁷To reconcile the model with the evidence that other disturbances might be at work in addition to pure sampling error, the distribution approach might be extended by including individual or aggregate covariates in the specification, something that has not been done so far and that we also leave for future research.

Second, we estimate the internal influence parameter q , conditional on the market potential and the estimated internal influence parameter. Third, we implement a MSM estimator of m by iterating the first two steps with different values of the market potential parameter, until we find the set of estimates $(\hat{m}, \hat{p}, \hat{q})$ which minimize the distance between the observed and simulated moment (the mean adoption time for the adopters). As in Schmittlein and Mahajan (1982), our three-stage estimator converges to the true values of the parameters as the population of potential adopters increases. In small samples, the estimates of the market potential are upward biased, while those of the influence parameters are downward biased.

5 Estimation of p and q when m is known

Let's consider an homogeneous population of m individuals, where the individual hazard of adoption is given by eq. 3. As standard in this literature, we assume that the individuals act independently of each other within each time interval. In every period therefore we observe a number of adoptions which is a random draw from a binomial distribution $B(m - N_t, h_t)$, though the parameters of this distribution –the population at risk $m - N_t$ and the probability of adoption $h_t = p + qN_t/m$ – change over time. The natural (ML) estimator for h_t is therefore:

$$\hat{h}_t = \frac{n_t}{m - N_{t-1}} \quad (7)$$

which is unbiased, consistent and asymptotically normal for large populations m .

Given that $N_0 = 0$, data on the first period of introduction of the new product can be used to estimate p :

$$\hat{p} = \frac{N_1}{m} \quad (8)$$

with variance

$$Var(\hat{p}) = \frac{\hat{p}(1 - \hat{p})}{m} = \frac{N_1(m - N_1)}{m^3} \quad (9)$$

The estimator is unbiased, consistent and asymptotically normal for large populations m . We can then exploit knowledge of $\hat{p}$ to estimate q . By plugging $\hat{p}$ in eq. 7 we obtain:

$$\hat{p} + q \frac{N_t}{m} = \frac{n_t}{m - N_{t-1}} \quad (10)$$

Figure 2: Top panel: simulated adoption data, with $p = 0.03, q = 0.4, m = 100,000$. Bottom panel: estimates of q .

which gives:

$$\hat{q}(t) = \frac{m}{N_t} \left[\frac{n_t}{m - N_{t-1}} - \frac{N_1}{m} \right] \quad (11)$$

with variance, given $\hat{p}$,

$$Var(\hat{q}(t)|\hat{p}) = Var(\hat{h}) = \frac{h(t)(1 - h(t))}{m - N_{t-1}} = n_t \frac{m - N_{t-1}}{m - N_t} \quad (12)$$

In every period after the first one, an estimate of q can thus be obtained, which is also unbiased, consistent and asymptotically normal, for large populations m . As observed by Boswijk and Franses (2005), at the beginning and toward the end of the diffusion process the number of new adoptions is smaller than in the middle, which leads to more volatile estimates. Figure 2 shows the different estimates of q that are obtained from one simulated process with parameters $p = 0.03, q = 0.4, m = 100,000$.⁸

⁸In empirical applications of the Bass model, the average value of p has been found to be 0.03, and it is often less than 0.01; the average value of q has been found to be 0.38, with a typical range between 0.3 and 0.5 (Mahajan et al., 1995).

As our final estimator of q we therefore use a weighted average of $\hat{q}$, using the inverse of the variance of the estimator (eq. 12) as weight w :

$$\hat{q}(T) = \sum_{i=2}^T w_i \hat{q}(i) \quad (13)$$

This estimator is unbiased, consistent and asymptotically normal, for large populations m . We now turn to the case when the market potential m is not known and has to be estimated in the data together with the influence parameters p and q .

6 Estimation of m

To estimate the market potential, we adopt a MSM estimator (McFadden, 1989; Pakes, 1989; Lee and Ingram, 1991; Duffie and Singleton, 1993), and minimize the distance between the observed moment τ_r (which is given) and the simulated moment $\tau_s(m)$, obtained by simulating the adoption time of m individuals with $h_t = \hat{p}(m) + \hat{q}(m)N_{t-1}/m$:⁹

$$\begin{aligned} \hat{m} &= \operatorname{argmin}_m [\tau_r - \tau_s(m)]^2 \\ \hat{p} &= \hat{p}(\hat{m}) \\ \hat{q} &= \hat{q}(\hat{m}) \end{aligned} \quad (14)$$

The moment we use is the mean adoption time for those who have adopted in the observation period:

$$\tau(T, m) = \frac{1}{N_t} \sum_{t=0}^T (tn_t) \quad (15)$$

On the simulated data, this moment is computed by using the estimated parameters $\hat{m}, \hat{p}$ and $\hat{q}$. Figure 3 shows how the moment responds to changes in m , for fixed values of the other parameters.

For each value of m , $\tau(T, m)$ is a random variable. Figure 4, which depicts its skewness, shows that it is not significantly different from 0. The distribution is therefore (almost) symmetric, a property that will turn out to be important in understanding the direction of the small sample bias.

Table 1 reports the results of a Montecarlo analysis on our estimation procedure, for $p =$

⁹Minimization is performed in Matlab with the `fmincon()` function. The code is available upon request.

Figure 3: Average adoption time for the adopters ($\tau(T, m)$), different values of m . Other parameters: $p = 0.03, q = 0.4, T = 10$. Ten artificial adoption sequences are simulated for each value of m . For each sequence, 10 replications of the estimation procedure are performed, with different pseudo-random numbers. For each set of estimated parameters, $\tau(T, m)$ is computed. The graph reports average values.

0.03, $q = 0.4$, different values of m and an observation period $T = 10$. 1,000 different pseudo-real adoption sequences are simulated for each experiment. The table also reports the standard deviations of the estimated parameters, which –given that the estimates are almost centered around the true values of the parameters– can be interpreted as the standard errors of the estimated parameters.¹⁰

		m=10,000	m=100,000	m=1,000,000
true value		mean over 1,000 replications (std.dev.)		
$\hat{m}$		10,049 (325)	100,118 (1,063)	999,932 (3,666)
$\hat{p}$	0.03	0.02988 (0.0013)	0.02998 (0.0083)	0.03001 (0.0001)
$\hat{q}$	0.4	0.39954 (0.0244)	0.39960 (0.0004)	0.40006 (0.0028)

Table 1: Estimated parameters, Montecarlo results. 1,000 different pseudo-real adoption sequences are simulated for each value of m . Estimation is based on the first $T = 10$ periods of each sequence.

There is a small sample bias, although not large: in the first experiment ($m = 10,000$), m is overestimated by 0.49%. This is reflected in the estimates for p and q , that are underestimated by 0.39% and 0.12% respectively. However, the bias for $\hat{m}$ drops to 0.12% with a population of

¹⁰In a real application, we would get only one set of estimates. To obtain the standard errors, we would bootstrap the estimation procedure by simulating artificial adoption sequences for the estimated values of the parameters.

Figure 4: Skewness of the $\tau(T, m)$ distribution, different values of m . Other parameters: $p = 0.03, q = 0.4, T = 10$. Ten artificial adoption sequences are simulated for each value of m . For each sequence, 10 replications of the estimation procedure are performed, with different pseudo-random numbers. For each set of estimated parameters, $\tau(T, m)$ is computed. The graph reports the skewness of the conditional distributions $\tau(T, m|m)$.

potential adopters of 100,000 individuals, and becomes negligible with 1 million agents. Correspondingly, the bias for $\hat{p}$ and $\hat{q}$ drops to -0.08% and -0.10% respectively when $m = 100,000$ and vanishes for $m = 1,000,000$ (+0.03% for $\hat{p}$ and +0.01% for $\hat{q}$). Figure 5 shows the distributions of the 1,000 estimates of the three parameters obtained in the Montercarlo experiment for $m = 1,000,000$. Normality is refused only for $\hat{p}$ (the values of the Shapiro-Wilk test are 0.089 for $\hat{m}$, 0.024 for $\hat{p}$, and 0.475 for $\hat{q}$).

Even for lower values of m , the bias is small enough that we might consider estimation on a smaller number of periods –something that is highly valuable when the estimates are used for forecasting. Figure 6 shows the cumulated mean of the estimated parameters over 1,000 simulated adoption series, for 1 million agents. The estimates quickly converge to values very close to the true values.

7 Small sample bias

Where does the small sample bias come from? Figure 3 contains the answer. The theoretical moment is not linear in m . If the observed moment, which is a random variable, is symmetric and centered around the theoretical model (see figure 4), we have

$$E[\tau^{-1}(m) \neq m] \tag{16}$$

Figure 5: Distribution of estimated coefficients. True values of the parameters: $p = 0.03$, $q = 0.4$, $m = 1,000,000$. The estimates are based on observations on the first $T = 10$ periods.

The direction of the bias depends on the sign of the first and second derivatives of the moment, at the true value of the parameter. For example, if the first derivative is positive, a positive second derivative implies that the moment is accelerating in m : it is less steep at the left than at the right of the true value of the parameter. Therefore, a low realization of the moment τ_L leads to a very low inferred value of the parameter $\hat{m}_L = \tau^{-1}(\tau_L)$, while a high realization τ_H leads to a not-so-high inferred value $\hat{m}_H = \tau^{-1}(\tau_H)$, with $E[\hat{m}_L, \hat{m}_H] < m$. We get a downward bias. Figure 7 illustrates the possible cases.

Given the shape of the mean adoption time for the adopters (figure 3), an upward bias is expected for $\hat{m}$ in small samples, that is exactly what we get from the Montecarlo analysis in table 1 and figure 6. The bias in the other parameters is consequential: an upward bias in $\hat{m}$ implies a downward bias in $\hat{p}$ and $\hat{q}$, given that the simulated penetration rate $F(t)$ is lower than the true (but unobserved) one (the same point is made in Grazzini, 2011).

The bias vanishes as the population of potential adopters increases because with a higher number of adopters the uncertainty over their mean adoption time reduces: the mean adoption

Figure 6: Cumulated mean of the estimated parameters, Montecarlo results. 300 different pseudo-real adoption sequences are simulated with $p = 0.03$, $q = 0.4$, $m = 1,000,000$. Estimation is based on the first $T = 5$ periods of each sequence.

Figure 7: Distribution of estimated coefficients. True values of the parameters: $p = 0.03, q = 0.4, m = 1,000,000$. The estimates are based on observations on the first $T = 10$ periods.

time converges to its theoretical value. Therefore, any extraction of the real data would produce the same mean adoption time, and the problem outlined above disappears.

8 Conclusions

In this paper we have shown an application of simulation-based econometric techniques to the estimation of AB models. The model chosen for the demonstration is important both because it represents a possibly very interesting application of AB models, and because it has been widely studied in its analytical form. The simplification of the model to a situation in which the diffusion dynamics is analytically described by the well known Bass model of innovation diffusion, allows to compare the simulation-based procedure with the estimation results obtained in that literature. Our simulation approach to estimation involves a three-stage procedure to estimate the parameters of the Bass diffusion model. First, we estimate the external influence parameter p , conditional on a specific value of the market potential parameter m . Second, we estimate the internal influence parameter q , conditional on the market potential and the estimated internal influence parameter. Third, we implement a MSM estimator of m by iterating the first two steps with different values of the market potential parameter, until we find the set of estimates $(\hat{m}, \hat{p}, \hat{q})$ which minimize the distance between the observed and simulated moment

(the mean adoption time for the adopters). Our three-stage estimator converges to the true values of the parameters as the population of potential adopters increases. In small samples, the estimates of the market potential are upward biased, while those of the influence parameters are downward biased. The bias problem is present also in the analytical estimation procedure and it is due to the non-linearity in the model. The micro representation of the diffusion dynamics allows to reduce the variance of the estimates (and the bias) by increasing the number of agents, leaving fixed the number of periods. Given that the Bass model is often applied to common products/technologies with a large market potential, the small sample bias should pose little concern.

Our estimation strategy has been carried out in the simple case of fully connected network, but it can be seen as a first step toward the estimation of diffusion model with more realistic network structures. In the more general version of our model, each individual has her own peer network, the individual hazards are no longer equal, and the simplification of obtaining ML estimates for the internal influence parameter q envisaged here is no longer applicable: all parameters, with the possible exception of the external influence parameter p for which ML estimation is still viable, should then be estimated by means of simulation-based techniques, along the lines we have described for the m parameter.

However, a richer specification in terms of individual heterogeneity and interaction structure has to confront with the paucity of empirical data, which in many cases are available only at a macro level. Even if we make parsimonious parametric assumptions about the structure of the network –for instance, we might assume that the network is a random graph– additional parameters have to be estimated –in the example, the average number n of links for each agent. We would then have to look for at least three moments, e.g. the first, second and third order moment of adoption time for those who have adopted in the observation period. Although this is theoretically feasible along the lines we have illustrated in this paper, from an empirical perspective it is simply asking too much to the data, especially if an early estimation is required (e.g., as in the previous section, $T = 5$). A trade-off inevitably appears between model inadequacy and data inadequacy.

This is the well known problem of *equifinality*, or under-identification (Liu, 1960; Hendry, 2002). Equifinality is of particular importance for AB models due to their intrinsic non-linearities: different non-linear models, equally validated by the real data, are in general characterised by different out-of-sample dynamics. Thus, they cannot be considered equivalent in

order to explain the phenomenon of interest, and exploit this knowledge for interpreting new events or evaluating the effects of alternative policies. This is in sharp contrast with linear models, which exhibit only limited deviations for limited departures from the validation set.¹¹

Future research should then investigate to what extent richer and more realistic network structures can be estimated from aggregate diffusion data.

Acknowledgements

We thanks Sebastiano Alessio Del Re, with whom we originally discussed the project and who developed the initial Matlab code. Preliminary versions of the paper were presented at the 37th Annual Meeting of the Eastern Economic Association in New York, February 25-27 2011, and at the GSDP Agent-based modelling workshop in Paris, September 8-10 2011. We thanks the participants to those events for their comments.

¹¹Of course, estimating these models on data including variation in those policy variables which affect the behavior of the agents would remove the problem: unfortunately, such data are not always available. This introduces a bias against the use of (AB) models with richer micro-foundations that favors the *status quo* in terms of modelling choices. Note however that reality itself is often intrinsically under-identified (Liu, 1960).

References

- Alfarano, S., Wagner, T., and Lux, F. (2005). Estimation of agent-based models: The case of an asymmetric herding model. *Computational Economics*, 26:19–49.
- Alfarano, S., Wagner, T., and Lux, F. (2006). Estimation of a simple agent-based model of financial markets: An application to australian stock and foreign exchange data. *Physica A*, 370(1):38–42.
- Bass, F. (1969). A new product growth for model consumer durables. *Management Science*, 15:215–227.
- Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *PNAS*, 99(Suppl. 3):7280–7287.
- Boswick, H., Hommes, C., and S., M. (2007). Behavioral heterogeneity in stock prices. *Journal of Economic Dynamics and Control*, 31(6):1938–1970.
- Boswijk, H. and Franses, P. (2005). On the econometrics of the bass diffusion model. *Journal of Business and Economic Statistics*, 23(3):255–268.
- Coleman, J. (1964). *Introduction to mathematical sociology*. The Free Press of Glencoe, London.
- Dancik, G., Jones, D., and Dorman, K. (2010). Parameter estimation and sensitivity analysis in an agent-based model of leishmania major infection. *J Theor Biol.*, 262(3):398–412.
- Duffie, D. and Singleton, K. (1993). Simulated moments estimation of markov models of asset prices. *Econometrica*, 61:929–952.
- Gallegati, M. and Richiardi, M. (2009). Agent-based modelling in economics and complexity. In R.A., M., editor, *Encyclopedia of Complexity and System Science*. Springer.
- Geroski, P. A. (2000). Models of technology diffusion. *Research Policy*, 29(4-5):603–625.
- Gilli, M. and Winker, P. (2003). A global optimization heuristic for estimating agent based models. *Computational Statistics and Data Analysis*, 42(2):299–312.
- Grazzini, J. (2011). Estimating micromotives from macrobehavior. Department of Economics Working Papers Series 201111, University of Turin.

- Hendry, D. (2002). Model identification and non-unique structure. Economics Papers Series No. 2002-W10, University of Oxford.
- Ivanov, A. (1997). *Asymptotic Theory of Nonlinear Regression*. Kluwer, Dordrecht.
- Jain, D. and Rao, R. (1990). Effect of price on the demand for durables: Modeling, estimation, and findings. *Journal of Business and Economic Statistics*, 8:163–170.
- Kirman, A. (1991). Epidemics of opinion and speculative bubbles in financial markets. In Taylor, M., editor, *Money and Financial Markets*, page 354368. Macmillan.
- Kirman, A. (1993). Ants, rationality, and recruitment. *The Quarterly Journal of Economics*, 108:137–156.
- Lee, B. and Ingram, B. (1991). Simulation estimation of time series models. *Journal of Econometrics*, 47:197–205.
- Leombruni, R. and Richiardi, M. (2005). Why are economists sceptical about agent-based simulations? *Physica A*, 3559:103–109.
- Liu, T. (1960). Underidentification, structural estimation, and forecasting. *Econometrica*, (28):855–865.
- Mahajan, V., Muller, E., and Bass, F. (1995). Diffusion of new products: Empirical generalizations and managerial uses. *Marketing Science*, 14(3):G79G88.
- Mansfield, E. (1961). Technical change and the rate of imitation. *Econometrica*, 29:741–766.
- McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica*, 57(5):995–1026.
- Miller, J. H. and Page, S. E. (2006). *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton University Press, Princeton, NY.
- Pakes, Ariel & Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, 57(5):1027–57.
- Rogers, E. M. (1962). *Diffusion of innovations*. Free Press, New York.
- Schmittlein, D. C. and Mahajan, V. (1982). Maximum likelihood estimation for an innovation diffusion model of new product acceptance. *Marketing Science*, 1(1):57–78.

- Srinivasan, V. and Mason, C. (1986). Nonlinear least square estimation of new product diffusion models. *Marketing Science*, 5(2):169–178.
- Stern, S. (1997). Simulation based estimation. *Journal of Economic Literature*, 35(4):2006–2039.
- Stern, S. (2000). Simulation-based inference in econometrics: motivation and methods. In Mariano, R., Schuermann, T., and Weeks, M. J., editors, *Simulation-based inference in econometrics: methods and applications*. Cambridge University Press.
- Tesfatsion, L. and Judd, K., editors (2006). *Handbook of Computational Economics.*, volume Volume 2: Agent-Based Computational Economics of *Handbook in Economics 13*. North-Holland.
- van den Bulte, C. and Lilien, G. (1997). Bias and systematic change in the parameter estimates of macro-level diffusion models. *Marketing Science*, 16(4):338–353.
- Winker, P. and Gilli, M. (2001). Indirect estimation of the parameters of agent based models of financial markets. Technical Report 03/2001, School of Business Administration, International University in Germany, Bruchsal.
- Winker, P., Gilli, M., and Jeleskovic, V. (2007). An objective function for simulation based inference on exchange rate data. *Journal of Economic Interaction and Coordination*, 2:125–145.