Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Physica A 355 (2005) 103–109
PHYSICA A
www.elsevier.com/locate/physa

Why are economists sceptical about agent-based simulations?

Roberto Leombruni, Matteo Richiardi*

LABORatorio R. Revelli Centre for Employment Studies, via Real Collegio 30, 10024 Moncalieri (Torino), Italy

Received 31 October 2004; received in revised form 11 January 2005
Available online 4 May 2005

Abstract

Despite many years of active research in the field and a number of fruitful applications, agent-based modeling has not yet made it through to the top ranking economic journals. In this paper we investigate why. We look at the following problematic areas: (i) interpretation of the simulation dynamics and generalization of the results, and (ii) estimation of the simulation model. We show that there exist solutions for both these issues. Along the way, we clarify some confounding differences in terminology between computer science and economic literature.

Keywords: Agent-based simulation; Economic methodology

1. Introduction

As it is well known, agent-based models (ABM) are a methodology developed in order to investigate the interplay occurring at two different scales of a given system: the macro structure and the micro structure. In facts, many systems are characterized by the fact that their aggregate properties cannot be deduced simply by looking at how each component behaves, the interaction structure itself playing a crucial role.

*Corresponding author.
E-mail addresses: roberto.leombruni@laboratoriorevelli.it (R. Leombruni), matteo.richiardi@laboratoriorevelli.it (M. Richiardi).
URL: http://www.laboratoriorevelli.it.

0378-4371/$ - see front matter © 2005 Published by Elsevier B.V.
doi:10.1016/j.physa.2005.02.072

This is particularly the case with complex systems, which are often characterized by the so-called "self-organization" property: the ability of a system to "learn" the individual strategies that lead to some unexpected ("emergent") aggregate behavior. However, non-linear systems become easily analytically intractable, or simply very difficult to manipulate algebraically, and the traditional approach of simplifying them may often "throw the baby out with the wash water". On the contrary, ABM allow a flexible design of how the individual entities behave and interact, since the results are computed and need not be solved analytically. It should thus be no surprise that they started to command a growing attention as a new "science of complexity" arose in the last decades of the 20th century [1]. In the social sciences, the recognition that many social and economic phenomena are themselves "complex" and can be fruitfully analyzed with the tools of the new science is mainly linked with the work conducted at the Santa Fe Institute during the 1990s [2,3]. However, despite the upsurge in ABM research witnessed in the past 15 years (see the reviews [4–7]), the methodology is still left aside in a standard economist's toolbox. Among the top 20 economic journals we were able to find only eight articles based on ABM ([8–15]).[1] This number is to be compared with the 26,698 articles that were published since the seminal work of Arthur (1988) in the top 20 journals considered. Agent-based modeling thus counts for less than 0.03% of top economic research. It seems to be confined only in specialized journals like the Journal of Economic Dynamics and Control[2], ranking 23rd, the Journal of Artificial Societies and Social Simulation, and Computational Economics, which are not even ranked. A notable exception is the Journal of Economic Behavior and Organization, ranked 32, which sometimes publishes research in ABM.

No matter that non-standard approaches like econophysics, evolutionary economics and Austrian economics have fruitfully employed ABM to develop stand-alone models or to extend analytical results, the mainstream remains sceptical. Of course, it is not necessary to justify scepticism toward a new methodology: it is sufficient to ignore it. Consequently, there are no discussions on the side of traditional economics about the perceived limits of ABM.[3] So, the aim of this work is to offer an explanation of why ABM in economics are generally regarded with suspect. Our work could be valuable for making researchers from other fields and ABM practitioners in economics more aware of the kind of reaction their work may have, and to provide a few counterarguments that could be taken into consideration by the sceptical economist, who will often react only by saying that simulations

[1]We looked for journal articles containing the words "agent-based", "multi-agent", "computer simulation", "computer experiment", "micro-simulation", "genetic algorithm", "complex systems", "El Farol", "evolutionary prisoner's dilemma", "prisoner's dilemma and simulation" and variations in their title, keywords or abstract in the EconLit database, the American Economic Association electronic bibliography of world economics literature. Note however that EconLit sometimes does not report keywords and abstracts. We have thus integrated the resulting list with the references cited in the review articles cited above. The ranking is provided in Ref. [16].

[2]JEDC has a section devoted to computational methods in economics and finance.

[3]A journalist has actually taken the job: see the non-technical and rather critical view of the research on complex systems undertaken at the Santa Fe Institute through the mid-1990s in Ref. [17].

"do not prove anything". Actually, to a closer inspection this general concern boils down to the belief that agent-based simulations are (i) *difficult to interpret and generalize*, *and* (ii) *difficult to estimate*. In the next sections we briefly discuss each issue.

2. Generalization

A rather common misunderstanding about simulations is that they are not as sound as mathematical models. In particular, they do not offer a compact set of equations—together with their inevitable algebraic solution—which can easily be interpreted and generalized. Actually, simulations *do* consist of a well-defined (although not concise) set of functions. These functions, which may be either deterministic or stochastic[4], unambiguously define the macro-dynamics of the system. Moreover, the eventual unique equilibrium of the macro-dynamics is, in turn, a known function of the structural parameters and initial conditions of the simulation. We will show here that the only difference from a model consisting of an algebraically solved set of equations is in the degree of knowledge that we have about these functions.

Let us start from the following general characterization of dynamic micro models. Assume that at each time t an individual i, $i \in 1, \dots, n$, is well described by a state variable $x_{i,t} \in \Re^k$. Let the evolution of her state variable be specified by the difference equation:

$$x_{i,t+1} = f_i(x_{i,t}, x_{-i,t}; \alpha_i) , \tag{1}$$

where we assume that the behavioral rules[5] may be individual-specific both in the functional form of the phase line $f_i(.)$ and in the parameters α_i, and may also be based on the state x_{-i} of all individuals other than i. Once we have specified the behavior of each individual, we will typically be interested in some macro-feature of our economy, that we may represent as a statistic Y defined over the entire population:

$$Y_t = s(x_{1,t}, \dots, x_{n,t}) . \tag{2}$$

The crucial question now is whether it is possible to solve Eq. (2) for each t, regardless of the specification adopted for $f_i(.)$, and the answer is that a solution can always be found by iteratively solving each term $x_{i,t}$ in Eq. (2) using Eq. (1):

$$\begin{aligned}
Y_0 &= s(x_{1,0}, \dots, x_{n,0}) , \\
Y_1 &= s(x_{1,1}, \dots, x_{n,1}) \\
&= s(f_1(x_{1,0}, x_{-1,0}; \alpha_1), \dots, f_n(x_{n,0}, x_{-n,0}; \alpha_n)) \\
&\equiv g_1(x_{1,0}, \dots, x_{n,0}; \alpha_1, \dots, \alpha_n) \\
&\vdots \\
Y_t &= g_t(x_{1,0}, \dots, x_{n,0}; \alpha_1, \dots, \alpha_n) . \qquad (3)
\end{aligned}$$

[4]In what follows we will refer to the deterministic case. Generalization to the stochastic case requires some changes (mainly regarding the notation), but the idea remains the same.

[5]Here and in the following we use "behavioral rules" and similar terms in a loose sense that encompasses the actual intentional behaviors of individuals as well as other factors such as technology etc.

The law of motion (3) uniquely relates the value of Y at any time t to the initial conditions of the system and to the values of the parameters α_i. Sometimes[6], g_t may converge to a function not dependent on t[7], so that we also have an expression for the equilibrium value of Y, again as a function of the initial conditions and parameters:

$$Y^e = \lim_{t \to \infty} Y_t \equiv g(x_{1,0}, \ldots, x_{n,0}; \alpha_1, \ldots, \alpha_n) . \tag{4}$$

Notice that this formalization describes both "traditional" dynamic micro-models and agent-based simulations. Indeed, given this common framework, it is easy to discuss the alleged differences in terms of "mathematical soundness". To explore this point, let us consider how the framework is implemented in the two approaches. As an example of the "traditional approach" think of a model based on a representative agent. The behavioral rule (1), will be very simple in structure, since all subscripts i can be dropped, along with any reference to other individuals' behavior. In turn, any "macro" statistic considered will collapse on a transformation of the state variable of just one individual, and the resulting law of motion (3) will also be very simple. We thus end up with a simple formulation for all Eqs. (1)–(3), and usually also for Eq. (4). By "simple formulations" we mean that they can be manipulated algebraically, and general propositions about the model can be stated by computing derivatives, comparing different equilibrium solutions, and so on.[8]

Let us turn to the agent-based simulation approach. The critical factor rests in the formula for the macro-dynamics (3), the law of motion of Y. As t and n get higher, the expression for $g_t(.)$ can easily grow enormous, hindering any attempt at symbolic manipulation, i.e., any attempt to solve it algebraically.[9]

Nevertheless, the functions (3) are completely specified. It is thus possible to explore their local behavior, by computing the value of Y corresponding to different values of the parameters and the initial conditions. A way to extrapolate this point evidence, and thus to recover a local approximation of the shape of $g_t(.)$, is to specify a functional form $\hat{g}_t(x_{1,0}, \ldots, x_{n,0}, \alpha_1, \ldots, \alpha_n, \beta)$ to be fitted on the artificial data generated by the simulation runs, where β are the coefficients of $\hat{g}_t(.)$. For instance, if $\hat{g}_t(.)$ is assumed to be linear, there will be two coefficients β_0 and β_1 (the intercept and the slope) to be estimated in the artificial data. The use of econometric techniques to approximate $g_t(.)$, starting from a number of—somehow designed—artificial experiments is indeed common practice in the computer science literature. The resulting regression model is also known as *metamodel*, *response surface*, *compact model*, *emulator*, etc. [18].

[6]When the dynamic system has one (or more), stable equilibrium and the initial conditions lie in its (their) basin of attraction.

[7]Or even not dependent on the initial conditions.

[8]Note that the problem of deriving the equilibrium relation (4) from the law of motion (3) is often skipped altogether. Equilibrium conditions are externally imposed, and the dynamics towards the equilibrium is simply ignored: the system "jumps" to the equilibrium.

[9]This difficulty is the same experienced in game theory models, where games typically become intractable if they involve more than a handful of players.

A cause of concern with this procedure stems from the possibility that the artificial data may not be representative of all outcomes the model can produce. In other words, it is possible that as soon as we move to different values of the parameters, the behavior of $g_t(.)$ will change dramatically. The metamodel $\hat{g}_t(.)$ will then become a poor description of the simulated world. At a theoretical level, this issue can be answered with two observations. First, if it applies to what we know about the artificial world defined by the simulation model, it also applies to what we know about the real world. As the real data generating process is itself unknown, stylized facts could in principle go wrong at some point in time. Second, we should not worry too much about the behavior of a model for particular "evil" combinations of the parameters, as long as these combinations remain extremely rare.[10] If the design of the experiments is sufficiently accurate (often particular combinations of the relevant parameter can be guessed, and oversampled in the artificial experiments), the problem of how "local" the estimated local data generating process is becomes marginal.

3. Estimation

The approximation of the law of motion $g_t(.)$ of the aggregate variable Y cannot be used for further estimation on real data. It has no unknown coefficients. It simply describes how the simulation model behaves, for given values of the structural parameters and the initial conditions. As such, it can be used to assess whether the simulation model is able to mimic the phenomenon of interest, by imposing the same metamodel $\hat{g}_t(.)$ on the real data, and comparing the coefficient vector $\hat{\beta}$ estimated on the artificial data with the coefficient vector $\tilde{\beta}$ estimated on the real data. Now, different coefficient vectors $\hat{\beta}$ are obtained for different values of the structural parameters vectors α_i. Intuition may suggest that we are not far from being able to estimate the structural parameters themselves. For instance, we could compare the outcome of the simulation with the real data, and change the structural coefficients values until the distance between the simulation output and the real data is minimized. In the simulation literature, this is called *calibration*. In the econometrics literature, it is called *structural estimation*, and can also be performed by means of *simulation-based estimation* methods [19].

Another objection says that the richer specifications of simulation models often lead to underidentification, due to the lack of exclusion restrictions. This claim seems to suggest that algebraic models are characterized by lean specifications only to avoid the problem of underidentification, and not because of symbolic tractability.

[10] The relevant exception is when rare events are themselves the focus of the investigation, for instance as in risk management. Here, simulations may prove extremely useful, by dispensing from making assumptions—such as the Gaussian distribution of some relevant parameters—which may be necessary in order to derive algebraic results but have unpleasant properties—like excessively thin tails. In a simulation, the reproduction of such rare events is limited only by the computational burden imposed on the computer. However, techniques can be used in order to artificially increase the likelihood of their occurrence.

Moreover, underidentification should not be the fear number one in writing a model. Rather, the inability of a model to provide a good description of the underlying phenomenon is a much greater limit [20]. Simulation allows complex models. This must be considered as a positive rather than a negative feature, since it makes a more detailed description of the phenomena of interest possible. The risk of under-identification is often simply unavoidable, the structural coefficients being *really* indeterminate: algebraic models that claim to be immune are sometimes only poor models.

4. Conclusions

In this paper we have presented a mathematical formalism able both to describe traditional analytical models and agent-based ones. Within this framework we critically assess the pros and cons of agent-based simulations, vis-à-vis analytical models. We then argue that much of the criticism regarding (1) absence of generality and (2) difficulty of estimation of ABMs fails. Moreover, ABMs have some important advantages, like the richer specification they can support, a feature which allows the description of complex phenomena.

References

[1] J. Gleick, Chaos: Making a New Science, Penguin Books, New York, 1987.
[2] P.W. Anderson, K.J. Arrow, D. Pines (Eds.), The Economy as an Evolving Complex System, SFI Studies in the Sciences of Complexity, Addison-Wesley Longman, Redwood City, CA, 1988.
[3] W.B. Arthur, S.N. Durlauf, D.A. Lane (Eds.), The Economy as an Evolving Complex System II, Addison-Wesley Longman, 1997.
[4] L. Tesfatsion, Special issue on agent-based computational economics, J. Econ. Dyn. Control 25 (3–4) (2001).
[5] L. Tesfatsion, Special issue on agent-based computational economics, Comput. Econ. 18 (1) (2001).
[6] L. Tesfatsion, Agent-based computational economics: a brief guide to the literature, in: J. Michie (Ed.), Reader's Guide to the Social Sciences, vol. 1, Fitzroy-Dearborn, London, UK, 2001.
[7] H.A. Wan, A. Hunter, P. Dunne, Autonomous agent models of stock markets, Artif. Intell. Rev. 17 (2002) 87–128.
[8] J. Arifovic, Genetic algorithms and inflationary economies, J. Monetary Econ. 36 (1) (1995) 219–243.
[9] J. Arifovic, The behavior of the exchange rate in the genetic algorithm and experimental economies, J. Polit. Econ. 104 (3) (1996) 510–541.
[10] J. Andreoni, J. Miller, Auctions with adaptive artificial agents, Games Econ. Behav. 10 (1995) 39–64.
[11] B. Arthur, On designing economic agents that behave like human agents: a beavioral approach to bounded rationality, Amer. Econ. Rev. 81 (1991) 353–359.
[12] B. Arthur, Inductive reasoning and bounded rationality, Amer. Econ. Rev. 84 (1994) 406.
[13] D.K. Gode, S. Sunder, Allocative efficiency of markets with zero-intelligence traders: markets as a partial substitute for individual rationality, J. Polit. Econ. 101 (1993) 119–137.
[14] G. Weisbuch, A. Kirman, F. Herreiner, Market organization and trading relationships, Econ. J. 110 (2000) 411–436.
[15] G. Tullock, C.D. Campbell, Computer simulation of a small voting system, Econ. J. 80 (317) (1970) 97–104.

[16] P. Kalaitzidakis, T. Stengos, T.P. Mamuneas, Rankings of academic journals and institutions in economics, J. Eur. Econ. Assoc. 1 (6) (2003) 1346–1366.
[17] J. Horgan, From complexity to perplexity, Sci. Amer. 272 (6) (1995) 104.
[18] J. Kleijnen, Experimental design for sensitivity analysis, optimization, and validation of simulation models, in: J. Banks (Ed.), Handbook of Simulation, Wiley, New York, 1998, pp. 173–223 (Chapter 6).
[19] C. Gourieroux, A. Monfort, Simulation Based Econometric Methods, Oup/Core Lecture Series, Oxford University Press, Oxford, 1997.
[20] T. Liu, Underidentification, structural estimation, and forecasting, Econ. (28) (1960) 855–865.