Agent-based financial markets: A review of the methodology and domain

Andrew Todd*, Peter Beling*, William Scherer* and Steve Y. Yang†
*Department of Systems and Information Engineering
University of Virginia, Charlottesville, VA 22904-4747, USA
Email: {aet6cd, pb3a, wts}@virginia.edu
†Financial Engineering Division, Stevens Institute of Technology
1 Castle Point on Hudson, Hoboken, NJ 07030.
Email: syang14@stevens.edu

***Abstract*—An agent-based model is a computer simulation driven by the individual decisions of programmed agents. Such models provide a promising alternative to traditional economic modeling in that they can fully capture the diversity of agents and the institutional detail of the underlying an economic system. In this paper, we provide a brief methodological review of the agent-based approach to modeling financial markets. We review the research strategy, which is organized into a discussion of formulation, implementation, verification and validation. We conclude the paper with a review of the domain focusing on modeling market participants and market institutions.**

I. Introduction

An agent-based model (ABM) is a computer simulation driven by the individual decisions of programmed agents. Such models provide a promising alternative to traditional economic modeling in that they can fully capture the diversity of agents and the institutional detail of the underlying an economic system. For the purposes of this paper, we adopt the following definition of economic simulation as given by Lehtinen and Kuorikoski (2007):

> Simulations in economics aim at imitating an economically relevant real or possible system by creating societies of artificial agents and an institutional structure in such a way that the epistemically important properties of the computer model depend on this imitation relation [1].

Agents often enjoy the limelight when it comes to the topic of agent-based simulation, but as the authors of this definition have so aptly pointed out, institutions, or in the lingo of agent-based modeling, the topology, should be considered equally as important. In agent-based modeling, the topology is "a set of agent relationships and methods of interaction" that "defines how and with whom agents interact" [2]. In an agent-based financial model, the topology is defined by the market structure and market institutions, as well as the actual mechanisms that allow for trade. Institutions provide a convenient way to constrain the allowable actions of agents. For example, we could institute an auction mechanism and require agents to interact through placing bids and offers. Finally, there is the environment, which consists of news and other signals that serve as inputs to the agent decisions. Our paper begins with a general methodological discussion and then turns its focus to these two primary ingredients: agents and institutions.

II. Agent-based Methodology

A. Formulation

According to Hedstrom et al. (2010) and agent-based modeling research strategy proceeds as follows. First, start with a "clearly delineated social fact that is to be explained" [3]. In the case of financial markets, agent-based models often seek to account for the so-called stylized facts of financial markets. Stylized facts are simply empirical regularities that appear to be stable across markets and over time. Such facts range from statistical concepts, such as the distribution of returns, to more abstract notions such as stock market bubbles and crashes.

Pagan (1996) establishes a number of stylized facts of financial market time-series, and offers a closing comment that "statistical approaches to the modeling of financial series have possibly reached the limits of their usefulness" [4]. This comment gives a hint of motivation for generative approaches and reflects a general unhappiness with the complexity of statistical models required to capture the conditional distribution of returns. Cont (2001) also surveys a number of stylized facts of financial markets [5]. Chen, Chang and Du (2012) identify the stylized facts that are reproducible by different classes of agent-based models [6]. Overall reproducing the stylized facts of financial markets has been a general and popular approach to modeling, which is perhaps less than ideal when one considers that these models often offer little in the way of experimental design. The hope, however, is that simulation results are very robust to reasonable changes in the underlying behavior of the agents, which highlights the importance of sensitivity analysis [7].

The second step is to formulate hypotheses about the relevant micro-level mechanisms. For us, this could mean hypothesizing about agents or institutions. Designing agents can be quite involved. How do the agents make decisions? What is the composition of agents in a given market? How does the composition change over time? Do the agents learn or adapt? Institutions are also difficult. There is the market structure generally, and the specific aspects of individual exchanges, such the market mechanism and the fee schedule.

What details do we need to capture the dynamics of the market? Which are extraneous?

B. Implementation

Next, the hypotheses must be translated into a computational model. This step involves a number of practical problems to be solved. For example, how are the agents activated? That is, how do they take turns? A subtle consideration, but one that may significantly affect outcomes [8]. Consider, for example, the two activation regimes studied by Axtel (2000) [9]. In the first regime, an agent is activated once each period. In the second, an agent is activated a random number of times in a given period with a mean of 1. Simulation output should be stables across reasonable activation schemes. (See Radax and Rengs (2010) for a taxonomy of activation regimes [10].)

The actual coding and implementation of agent-based models can be quite involved. Ideally, the model, on the conceptual level, will have a clean mathematical definition, and that definition need only be implemented using the standard algorithms for stochastic simulation. However, many models stray from purely mathematical descriptions, which not only makes their actual implementation more complicated, it also makes the models more difficult to communicate to the research community. A good model should have a relatively concise description, which can be used to develop independent implementations. Results should be reproducible without the need to include code. A good practice is to follow established standards, such as the ODD protocol [11].

Researchers also need to choose from the plethora of platforms available to build agent-based models. An important consideration is performance. Tools like Netlogo offer extremely convenient and powerful platform for developing agent-based models, but they come at a performance cost [12], [13]. Some models may need to be implemented in lower level languages in order to achieve the performance required for replications and good statistical analysis. Overall, increasingly sophisticated tools are becoming available to builders of agent-based models [14], [15].

C. Verification and Validation

Once a computational representation of the model is available it needs to be verified and validated. You might ask: what is the difference? Verification is "the process of determining that a model or simulation implementation and its associated data accurately represent the developer's conceptual description and specifications," whereas validation is "the process of determining the degree to which a model or simulation and its associated data are an accurate representation of the real world from the perspective of the intended uses of the model" [16]. Verification of the model is the easy part. From a purely practical perspective, one might consider writing software tests, such as unit tests, to verify that the actual coded implementation of a model operates as expected.

Validation is more challenging. The definition of validation that we have provided is perhaps more relevant to discrete-event simulations that might be used to make decisions about an industrial process. How do we apply that definition of validation to an agent-based model that has an "intended use" of providing insight about a hypothetical economic system. Leigh Tesfastion suggests four distinct types of validation: input validation, process validation, descriptive output validation, and predictive output validation [17].

Input validation involves an examination of the exogenous inputs to the model. Such inputs may be derived from empirical observations. For example, an agent-based model might assume that there is a particular distribution of agents with in the population. For example, we might assume that half of the agents base their decision on fundamental analysis of a security and the other half base their decision on technical analysis. Input validation would involve a critical analysis of those particular assumptions.

Next there is process validation. Do the process and mechanics of the model adequately reflect the real world? For example, if we are modeling a market that employs a continuous limit order book, and our model implements the market mechanism using an equation that updates the price a discrete epochs, do we have a valid process? This is one of the most difficult areas of modeling as it involves identifying the correct level of abstraction for the model.

Finally, we have two types of output validation: descriptive output validation and predictive output validation. This distinction comes down to in-sample and out-of-sample validation. To understand it, we revisit a model that do both admirably. Mike and Farmer (2008) introduce a model of financial markets that is based on empirical estimates of order flow parameters. The model is validated by examining the statistical properties of the resulting prices process. In the paper, data from a single stock is used to design the model of order flow. That model is essentially chosen based on its in-sample performance. The model is then evaluated across a larger cross-section of stocks, which is an out-of-sample validation, or predictive output validation.

III. Agents

An artificial agent in the most general sense is an entity that perceives its environment through sensors and acts upon the observations through actuators [18]. In the context of financial markets, agents observe the market process and fundamental information about the securities being traded and initiate trades based on their observations. According to LeBaron (2001), an agent's directive is to "digest the large amounts of time series information generated during a market simulation, and convert this into portfolio decisions" [19]. However, as LeBaron notes, there are many ways to process that data, and thus many ways to define agents. Furthermore, not every agent shares identical motivation.

According to a taxonomy of traders provided by Harris (2001), there are utilitarian traders, speculators, dealers, and even futile traders [20]. Utilitarian traders include, asset exchangers, hedgers, gamblers and fledglings (among others). Hedgers trade to manage risk. Asset exchangers convert from one type of holding to a holding of more immediate value.

Gamblers trade for the thrill of it. Fledglings trade to determine if trading may be profitable. Speculators process information to predict future prices. Speculators may trade on fundamental information about the value of an asset or on technical information related to the trading process, or even a combination of both. Dealers exist to intermediate markets. Futile traders are irrational or victimized market participants that do not profit from their activity.

The variety of of financial market participants presents a tough modeling challenge. Traditionally, economics has solved this problem by working with a perfectly rational representative agent. So, how do designers of artificial markets approach the design of agents? Let's assume the goal is to produce artificial agents that reflect the heterogeneity of participants in real markets and respect the cognitive and informational constraints that those participants face. That is a tall order. Holland and Miller (1991) put the challenge this way: "Usually, there is only one way to be fully rational, but there are many ways to be less rational" [21]. While the latter is most surely true, in the context of constantly evolving financial markets, even the rational course of action is not always completely clear.

From the discussion, it is clear that agents may vary widely with respect to their motivations and how they observe and learn from their environments. Consider the extreme case of an agent that does not observe, does not learn and selects from its space of actions randomly. Here, you have what is known as a zero-intelligence agent. On the other end of the spectrum you have an agent that learns from its observations and evolves its behavior accordingly, i.e., an agent that relies on artificial intelligence. While the models on the two extremes of this spectrum are, broadly speaking, agent-based models, they represented distinct methodologies [22]. See Chen (2012) for a survey four origins of agent-based financial models including the markets origin, the cellular-automata origin, the tournaments origin, and the experiments origin [23]. We have only brushed the surface when it comes to the variety of approaches to modeling agents. It is the topology, i.e., the institutions, however, that are often neglected in the agent-based modeling literatures, and it is on that topic that we continue our discussion.

IV. Topology

Agents do not operate in a vacuum. In the case of financial markets, they are constrained by rules and regulations, market structure, and most importantly, the market mechanisms themselves, as implemented by the exchanges. As we have already pointed out, one of the most difficult tasks in agent-based modeling is reducing the universe from which to choose your agents. Carefully specifying market structure and market institutions can go a long way to reduce the complexity of designing agents. Returning to our auction example, it is easier to conceive an agent that is solely responsible for bidding in an auction than an agent that might need to make economic decisions in a much more general environment. We organize our discussion of topology for agent-based financial markets into two sections: market structure and market mechanisms.

A. Market Structure

A natural place to beging our discussion of market structure is with regulations. Regulations shape markets. In the United States, equity market structure is largely a result of Reg NMS, the National Market System. Equity markets in Europe, and in many other areas of the world, are much different. According to a list of MIC (Market Identifier Codes) there are over 1000 "exchanges, trading platforms, regulated or non-regulated markets and trade reporting facilities" globally [24]. More than 20 countries have 10 or more trading venues. The top 10 are given in Table I.

Despite the large number of venues, stocks markets are dominated by well-known exchanges such as NYSE and NASDAQ. From an agent-based modeling perspective, if we are modeling equity markets, how should we account for regulations and market structure? At a minimum, we need to account for market fragmentation. Figure 1 gives a recent measure of market share in U.S. equities. So the question is, how much fragmentation in a model implementation is enough? Could we recreate the basic dynamics with just two markets?

Countries	**Number of MIC**
United States of America	241
United Kingdom	188
Germany	71
Japan	43
Australia	31
Italy	31
Canada	23
Spain	22
Switzerland	21
Hong Kong	19

TABLE I
TOP 10 COUNTRIES BY NUMBER OF MIC CODES.

Some models have explicitly included mechanisms that reflect the reality of the fragmented market for U.S. equities. [25]. Other models have included alternative trading platforms, such as dark pools [26], [27]. However, market structure, for the most part, is largely neglected in agent-based models of financial markets. Most studies only consider a single market, and do not consider issues such as market fragmentation, market segmentation, and the heterogeneity of trading platforms. Market fragmentation is a fundamental, physical property of financial markets and cannot be neglected in modeling process.

Another crucial aspect of market structure is that of information flow. News is largely available in machine readable form, and we have already seen flash crashes resulting from posts to Twitter [28]. Agent-based models should consider a wholistic approach that not only accounts for the market structure with respect to multi-market trading and fragmented liquidity, but also account for the new means with which information flows. Yang et al (2014) study the dynamics of the Twitter financial community in an agent-based model, but these ideas have not been wholly integrated in multi-market

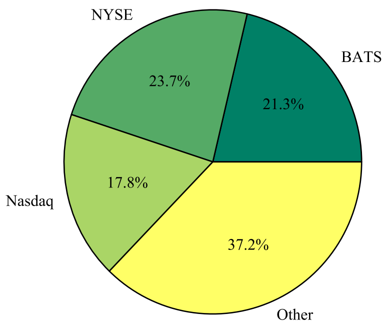

Fig. 1. 5-day moving average of market share as of November 20, 2015

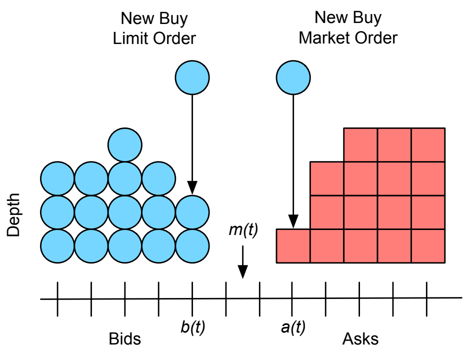

Fig. 2. Snapshot of a limit order book.

agent-based financial models [29]. Accounting for every aspect of the very complicated structure which market participants face in the global financial system is unrealistic. However, we should strive to find a balance between practicality and realism in defining the institutions and structure in which our agents interact.

B. Market Mechanisms

Not only is the broader market structure important, but so also are the details of the market mechanisms themselves. Most financial products trade in a continuous limit order book, which is a continuously evolving record of outstanding orders to buy or sell a particular security. When a limit order is submitted to an exchange it is either matched with another outstanding order, queued in the limit order book, or rejected. Limit orders have a price, quantity and side, which are the primary attributes that govern how they are processed.

The price and quantity of a limit order are restricted to multiples of the tick size and lot size, respectively. The tick size is the minimum price increment. The lot size is the minimum quantity that may be traded in a single transaction. A limit order may be rejected if it does not adhere to those restrictions. A limit order only executes if there exists an appropriately priced order on the opposite side of the market. Limit orders are queued in the book if they cannot be executed. Queued orders may be modified or canceled. Figure 2 displays a snapshot of a limit order book, which highlights important quantities, including the bid, ask and mid-quote, $b(t)$, $a(t)$ and $m(t)$, respectively.

The queued orders establish the best bid and ask prices. Market participants seeking to trade immediately will buy at the best ask price and sell at the best bid price. The queued orders are executed according to rules of precedence. The primary rule of precedence is price. The secondary rule of precedence is time. The limit order book is essentially a FIFO queue in which orders are executed based on price and origination time. However, orders may also be matched on a pro-rata basis, and matching algorithms, in general, may vary across exchanges and products.

Matching Engines	**URL**
Argo SimEx	http://www.argocons.com
Cinnober TRADExpress	http://www.cinnober.com
CME Globex	http://www.cmegroup.com/globex
Connamara	https://www.connamara.com
Gatelab exchangepath-100µs	http://www.gatelab.com
LSE Millenium Exchange	http://www.millenniumit.com
Orbixa ThymeX	http://www.orbixa.com
NASDAQ INET	http://www.nasdaqomx.com
NYSE Pillar	https://www.nyse.com/pillar
Thesys Technologies	http://www.thesystech.com

TABLE II
SAMPING OF COMMERICAL MATCHING ENGINES AND COMPANIES OFFERING MATCHING ENGINES SERVICES.

In general, exchanges offer limit order book functionality beyond the basics required to operate a continuous double auction. Limit orders may have additional attributes that govern specific aspects of their display and execution. For example, some exchanges offer the ability to place hidden orders. Other attributes might control details of execution (e.g. immediate-or-cancel, fill-or-kill or all-or-none). In the United States, exchanges have introduced order types that are tied to specifics of the market structure (e.g. orders pegged to the NBBO (national best bid and offer)). A matching engine is the software implementation of a limit order book. Table II provides an example of some exchange matching engines and other companies offering matching engines commercially.

For modeling purposes, it is sensible to implement only the core functionality required for the operation of a continuous double auction. However, it is important to understand, that not only do we need to consider the heterogeneity of the population of market participants, but also the heterogeneity of the market mechanisms themselves. Agent-based models have been used to analyze a number of specific proposals related to the operation of market mechanisms. Mizuta et al (2014 study the dynamics of markets under an up-tick rule [30]. Hayes et al (2012) a study the minimum quote life rule [31]. Price variation limits are also among the rules that have been analyzed in agent-based models [32].

V. Conclusion

In conclusion, while there are many aspects to planning and executing an agent-based research strategy for the study of financial markets, there is a large literature available to assist in some of the often neglected stages of the process including verification, validation, and even communication of models to the broader research community. Agent-based modeling provides a powerful approach to modeling economic systems and is uniquely capable of accounting for the diversity of market participants as well as complicated market structures. Agent-based models are also uniquely positioned to answer questions about specific rules and regulations related to market structure and market mechanisms.

References

[1] J. K. Aki Lehtinen, "Computing the perfect model: Why do economists shun simulation?," *Philosophy of Science*, vol. 74, no. 3, pp. 304–329, 2007.

[2] C. M. Macal and M. J. North, "Tutorial on agent-based modeling and simulation," in *Proceedings of the 37th Conference on Winter Simulation*, WSC '05, pp. 2–15, Winter Simulation Conference, 2005.

[3] P. Hedstrm and P. Ylikoski, "Causal mechanisms in the social sciences," *Annual Review of Sociology*, vol. 36, p. 49.

[4] A. Pagan, "The econometrics of financial markets," *Journal of Empirical Finance*, vol. 3, no. 1, pp. 15 – 102, 1996.

[5] R. Cont, "Empirical properties of asset returns: stylized facts and statistical issues," *Quantitative Finance*, vol. 1, no. 2, pp. 223–236, 2001.

[6] S.-H. Chen, C.-L. Chang, and Y.-R. Du, "Agent-based economic models and econometrics," *The Knowledge Engineering Review*, vol. 27, pp. 187–219, 6 2012.

[7] G. ten Broeke, G. van Voorn, and A. Ligtenberg, "Sensitivity analysis for agent-based models: a low complexity test-case," in *Proceedings of the Social Simulation conference (SSC'14)*, 2014.

[8] K. Comer, *Who Goes First? An Examination of the Impact of Activation on Outcome Behavior in Agent-based Models*. PhD thesis, George Mason University, 2014. http://gradworks.umi.com/36/23/3623940.html.

[9] R. Axtell, "Effets of Interaction Topology and Activation Regime in several Multi-Agent System," tech. rep., 2000.

[10] W. Radax and B. Rengs, "Timing matters: Lessons from the ca literature on updating," 2010. 3rd World Congress on Social Simulation.

[11] V. Grimm, U. Berger, F. Bastiansen, S. Eliassen, V. Ginot, J. Giske, J. Goss-Custard, T. Grand, S. K. Heinz, G. Huse, A. Huth, J. U. Jepsen, C. Jrgensen, W. M. Mooij, B. Mller, G. Peer, C. Piou, S. F. Railsback, A. M. Robbins, M. M. Robbins, E. Rossmanith, N. Rger, E. Strand, S. Souissi, R. A. Stillman, R. Vab, U. Visser, and D. L. DeAngelis, "A standard protocol for describing individual-based and agent-based models," *Ecological Modelling*, vol. 198, no. 12, pp. 115 – 126, 2006.

[12] E. Sklar, "Netlogo, a multi-agent simulation environment," *Artificial life*, vol. 13, no. 3, pp. 303–311, 2007.

[13] C. Nikolai and G. Madey, "Tools of the trade: A survey of various agent based modeling platforms," *Journal of Artificial Societies and Social Simulation*, vol. 12, no. 2, p. 2, 2009.

[14] A. Drogoul, E. Amouroux, P. Caillou, B. Gaudou, A. Grignard, N. Marilleau, P. Taillandier, M. Vavasseur, D.-A. Vo, and J.-D. Zucker, "Gama: multi-level and complex environment for agent-based models and simulations," in *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pp. 1361–1362, International Foundation for Autonomous Agents and Multiagent Systems, 2013.

[15] W. Dubitzky, K. Kurowski, and B. Schott, *Large-scale computing techniques for complex system simulations*, vol. 80. John Wiley & Sons, 2012.

[16] "DoD Modeling and Simulation, Verification, Validation, and Accreditation," Tech. Rep. 5000.61, Department of Defense, December 2009. http://www.dtic.mil/whs/directives/corres/pdf/500061p.pdf.

[17] L. Tesfastion, "Empirical Validation: Agent-Based Computational Economics," 2016. http://www2.econ.iastate.edu.

[18] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*. Pearson, 2009.

[19] B. LeBaron, "A builder's guide to agent-based financial markets," *Quantitative Finance*, vol. 1, no. 2, pp. 254–261, 2001.

[20] L. Harris, *Trading and exchanges: Market microstructure for practitioners*. Oxford University Press, USA, 2003.

[21] J. H. Holland and J. H. Miller, "Artificial adaptive agents in economic theory," *The American Economic Review*, vol. 81, no. 2, pp. 365–370, 1991.

[22] C. Schinckus, "Methodological comment on econophysics review i and ii: statistical econophysics and agent-based econophysics," *Quantitative Finance*, vol. 12, no. 8, pp. 1189–1192, 2012.

[23] S.-H. Chen, "Varieties of agents in agent-based computational economics: A historical and an interdisciplinary perspective," *Journal of Economic Dynamics and Control*, vol. 36, no. 1, pp. 1–25, 2012.

[24] M. LLC, "ISO 10383 - Market Identifier Codes," 2016.

[25] E. Wah and M. P. Wellman, "Latency arbitrage, market fragmentation, and efficiency: a two-market model," in *Proceedings of the fourteenth ACM conference on Electronic commerce*, pp. 855–872, ACM, 2013.

[26] S. Y. K. Mo, M. Paddrik, and S. Y. Yang, "A study of dark pool trading using an agent-based model," in *2013 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pp. 19–26, IEEE, 2013.

[27] T. Mizuta, W. Matsumoto, S. Kosugi, K. Izumi, T. Kusumoto, and S. Yoshimura, "Do dark pools stabilize markets and reduce market impacts? investigations using multi-agent simulations," in *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pp. 71–76, IEEE, 2014.

[28] T. Karppi and K. Crawford, "Social media, financial algorithms and the hack crash," *Theory, Culture & Society*, vol. 33, no. 1, pp. 73–92, 2016.

[29] S. Y. Yang, A. Liu, and S. Y. K. Mo, "Twitter financial community modeling using agent based simulation," in *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pp. 63–70, IEEE, 2014.

[30] T. Mizuta, K. Izumi, I. Yagi, and S. Yoshimura, "Regulations' effectiveness for market turbulence by large erroneous orders using multi agent simulation," in *2014 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pp. 138–143, IEEE, 2014.

[31] R. Hayes, M. Paddrik, A. Todd, S. Yang, P. Beling, and W. Scherer, "Agent based model of the e-mini future: application for policy making," in *Proceedings of the Winter Simulation Conference*, p. 111, Winter Simulation Conference, 2012.

[32] T. Mizuta, K. Izumi, and S. Yoshimura, "Price variation limits and financial market bubbles: Artificial market simulations with agents' learning process," in *2013 IEEE Conference on Computational Intelligence for Financial Engineering & Economics (CIFEr)*, pp. 1–7, IEEE, 2013.