An Agent-based Model of the Nasdaq Stock Market: Historic Validation and Future Directions.

Alexander V. Outkin

Sandia National Laboratories

Abstract. This paper presents an agent-based model of a dealer-mediated market, similar to Nasdaq [1]. We outline the overall model and representation of the market rules and order handling infrastructure, as well as representation of individual market-makers decision-making processes and strategies. The original model was created primarily to understand the effects of the then forthcoming decimalization. We discuss our original predictions on effects of decimalization and the subsequent market evolution, and the fact that predatory strategies have become more wide spread since the decimalization. We further outline the applicability of the model to represent the high frequency algorithms and effects of their interactions on the market dynamics.

1 Introduction

The global financial system is one of the most complex systems created by mankind. A key aspect of such systems is that they are composed of multiple interacting autonomous parts, complex rules of engagement, and agents who act for their own benefit and sometimes to the detriment of others. Each agent has goals, rules for action and interaction, and decision-making strategies. This interaction produces global system dynamics and generates equilibrium or disequilibrium prices of various assets.

Despite the nature of the global financial system, the majority of existing tools for the financial markets are based on the foundations of the general equilibrium theory, which originated in the 1950s, but is not considered capable of capturing the complexity and variety of agents and strategies observed in real world markets. Additionally, those tools operate under a range of rather unrealistic assumptions, including market clearing, market convergence to equilibrium prices, perfect information and rationality, profit maximizing behavior, and the absence of significant market power. Interactions are normally assumed to be non-local and non-evolutionary, which implies that the system arrives at the

[1] Much of the research outlined in this paper has been done in collaboration with Vince Darley, Tony Plate, and others at the BiosGroup. I would like to thank Walt Beyeler, Pat Finley, Brian Jones, Ed MacKerrow, and Vanessa Vargas for valuable suggestions and comments. The usual disclaimer applies.

equilibrium through a price adjustment process that is imposed on the system rather than derived from the interactions of players within the system.

This paper presents an agent-based model of the Nasdaq stock market, developed at the BiosGroup, Inc. in late 1990s - early 2000s with the purpose of understanding the effects of then forthcoming decimalization. It outlines the features of the model, the predictions made before the decimalization was implemented by Nasdaq in 2001, and market evolution after the decimalization. This model and its results has been extensively documented in Darley and Outkin (2007). Additionally, the paper outlines the framework for using only partially calibrated agent-based models for prediction.

2 Model and History

The approach here employs agent-based modeling to represent the behaviors of individual traders, market makers, system rules, and to derive the overall system dynamics. More information on agent-based modeling can be found in Schelling (1978) and Epstein and Axtell (1996), or, in application to financial markets,in Rust et al. (1993), Lux and Marchesi (1999), Darley and Outkin (2007) Darley et al. (2000, 2001), , Axtell (2003, 2005), Shubik and Smith (2004).. A part of this approach, where actions of agents and the system-level processing of those actions are modeled explicitly, is related to the area of casual models and is outlined in King (2002). Shafer and Vovk (2001) describe an exciting new approach to finance, where game theory, rather than the measure theory serves as a foundation for probability theory. This approach may apply to agent-based modeling as well, for agent-based models can be considered an extension of game-theoretic models. They share the concepts of agents, strategies, and interaction structures but differ in analytical tractability, modeling goals, and other important characteristics.

2.1 Model Outline

The model represents explicitly the interactions and the interaction structures that generate the price dynamics [2]. This includes the interaction between market makers, as well as between the market makers and investors. It includes representation for how the investor order is processed by the exchange, as well as strategies and responses of market makers and investors.

The model represents two types of agents: market makers (dealers) and investors. The market contains a single security whose fundamental value is exogenously specified: the underlying value of the security is assumed to fluctuate according to a random, externally specified process. Investors receive a noisy version of that fundamental value and act on the basis of that information, depending on the specific investor strategy: the fundamental investor types take

[2] This section is based on the model description that appeared in Darley et al. (2000).

into account the difference between the market price and their perceived fundamental value, the momentum investors mainly pay attention to the price dynamics, etc. Here, he fundamental value can be interpreted as an aggregate source of quasi-information.

Market makers can also extract information from the investor aggregate order flow and from other market observations; as a result, they generally possess information that is superior to that of the investors.

Market makers and investors are represented in the simulation as autonomous agents, which behave according to their individual strategies: these may be built in, or can arise as a result of learning or evolutionary selection. In the learning domain, we use neural networks and reinforcement learning to generate strategies for agents. The evolutionary selection determines the types of new strategies that will be introduced into the simulation when a market maker or investor goes bankrupt. The learning element is important because it allows us to investigate the possibilities resulting from strategies that have not yet been discovered by players in the real-world market.

2.2 History

This model has been created in late 1990s - early 2000s to understand the effects of decimalization and other rule changes on the Nasdaq stock market[3].

The main focus of our efforts was in the following four areas:

1. Investigating, mainly in a qualitative fashion, the consequences of regulatory and structural changes to the market (the most important being the question of minimum tick size).
2. Investigating whether our model, at least in a stylized fashion, is able to replicate some of the observed features of real-world markets.
3. Validating the model (this encompasses and includes the previous two points).
4. Designing learning agents, and investigating the behaviors they learn and their ability to perform profitably in the market.

When this effort starte, we did not know whether and when will we be able to compare our predictions with the real world outcomes, because the future of decimalization, and specifically the timing of Nasdaq decimalization, were uncertain.

Nasdaq has implemented decimal pricing over a six week period starting on March 12, 2001, starting with a limited number of stocks and then expanding it to the entire Nasdaq universe.

We have compared our prediction with the actual outcomes based on the decimalization study conducted by the Nasdaq Economic Research (2001). We found that the majority of our predictions were strongly supported by the available data.

[3] Most of this section is based on Darley and Outkin (2007) and earlier reports and papers on the Nasdaq modeling project

To generate the prediction, we calibrated the model to the audit trail data provided by Nasdaq. In particular this data represents the market makers associated with specific trades and therefore allows partially extracting and representing the market maker strategies, based on the specific transactions executed by specific market makers. In particular, we calibrated the model in two steps to:

1. Ensure that the simulated distribution of trade sizes, volumes, prices and other statistical parameters is similar to that observed in the real world.
2. Represent the real-world behaviors of market makers and investors using data sets of historical quotes and trades. In some cases, much of the quote change behavior (and thus of the market maker behavior) can be explained by very simple strategies. For example, for certain relatively simple market maker strategies, precision of calibration constituted 60-70

The following information summarizes our most important predictions:

1. Decimalization (tick size reduction) will negatively impact the price discovery process.
2. Volume will increase, potentially ranging from 15% to 600%.
3. Ambiguous investor wealth effects may be observed. (Investors average wealth may actually decrease in the simulation, but the effect is not statistically significant).
4. Phase transitions will occur in the space of market-maker strategies. (This was observed in the transactional data for the previous tick size reduction from $1/8th to $1/16th that occurred in 1997.)
5. Spread clustering may be more frequent with tick size reductions.
6. Parasitic strategies may become more effective as a result of tick size reductions.

Overall, our predictions are supported by available studies of decimalization results:

- There is strong, albeit indirect, evidence that reducing the tick size negatively impacts the price discovery process.
- There are strong indications that not all investor types benefited from decimalization. (There are ambiguous wealth effects, like those ob- served in our simulation.)
- There is unambiguous support for our prediction of phase transition in market maker strategies as a result of decimalization.
- Nasdaq Economic Research (2001) reports that there is a significant natural quote clustering in the decimalized market.
- That same source reports that there is evidence of more stepping in ahead of the customers orders (in particular on electronic communication networks, or ECNs) and more time when ECNs are on the inside market, which we interpret as indications of increased activity of parasitic strategies.

– Partially, as a result of decimalization, the hight frequency trading strategies, often associated with predatory behaviors, have become very popular.

A more detailed description of the comparisons of predicted to actual results can be found in Darley and Outkin (2007).

3 Possible extensions

This model has been calibrated using exhaustive and generally unavailable audit trail data. In particular, such data contains information on the identities of the market makers involved in a trade. This in turn allows reconstructing the market makers strategies with significant degree of confidence. However, generally such data is not available.

We propose instead use only partially calibrated agent-based models for prediction of real-world systems, where the agent-based models serve as an input into machine learning algorithms. Such learning framework can partially replace the individual market maker level data by using a learning loop operating on the level of individual models.

We illustrate the early success of this approach on the example of using the original Nasdaq model to replicate the much more recent data from the London Stock Exchange, as illustrated in the Figure 1. below.

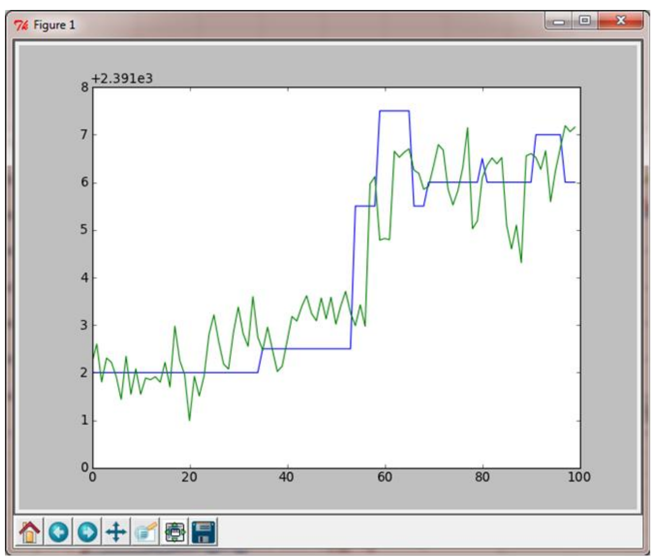

Fig. 1. Replication of the London Stock Exchange (green line) data using a set of models generated by the Nasdaq market simulation (blue line).

This approach has been described in Outkin (2012).

4 Conclusions

This paper demonstrates the success of an earlier agent-based model of the stock market and outlines a path forward for creating agent-based models when exhaustive data is not available by replacing calibration of individual models with learning that uses multiple models.

5 References

Axtell, R. (2003). Economics as Distributed Computation. In H. Deguchi et al. (Eds.), Meeting the Challenge of Social Problems via Agent-Based Simulation. Springer-Verlag.

Axtell, R. (2005). The complexity of exchange. Economic Journal, 115(504). F193F210.

Darley, V. and Outkin, A. (2007). A NASDAQ market simulation: insights on a major market from the science of complex adaptive systems. Singapore: World Scientific Publishing.

Darley, V, Outkin A., Plate T., and Gao, F. (2000). Sixteenths or pennies? Observations from a simulation of the Nasdaq stock market. Proceeding of IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering (CIFEr).

Darley, V., Outkin A., Plate T., and Gao, F. (2001). Learning, evolution and tick size effects in a simulation of the NASDAQ stock market. Proceedings of the 5th World Multi-Conference on Systemics, Cybernetics and Informatics. Orlando, Florida.

Epstein, J. M. and Axtell, R. (1996). Growing artificial societies: social science from the bottom up. The Brookings Institution.

Harris, L. (2003). Trading and exchanges: Market microstructure for practitioners, Oxford University Press.

King, J. (2002). Operational Risk: Measurement and Modelling. Wiley.

Lux, T. and Marchesi, M. (1999). Scaling and criticality in a stochastic multiagent model of a financial market. Nature, (397).

Outkin, Alexander V. (2012): Interaction-driven Learning Approaches to Complex Systems Modeling. Sandia National Laboratories Memo.

Rust, J., Miller, J., and Palmer, R. (1993). Behavior of trading automata in a computerized double auction market. The Double Auction Market: Institutions, Theories, and Evidence. Westview Press.

Schelling, T. (1978). Micromotives and Macrobehavior. W. W. Norton.

Shafer, G. and Vovk, V. (2001). Probability and Finance: Its Only a Game! Wiley-Interscience.

Shubik, M. and Smith, E. (2004). The physics of time and dimension in the economics of financial control. Physica A: Statistical Mechanics and its Applications, Volume 340, Issue 4, pp. 656-667.

6 ODD Addendum

Nasdaq simulation is a very large model with a large number of subcomponents, their possible interactions, rules, etc. Below is a brief outline of the model. More information can be found in Darley and Outkin (2007).

6.1 Purpose

The goal of the model is to represent the market operation in a single stock on the level of individual market makers and to allow investigation of possible effects of rule changes, such as decimalization, on market behavior.

6.2 Entities, state variables, and scales

Market makers, investors. Different agents may operate on different time scales.

There is a special category of learning agents. In particular, Q-learning has been used to represent learning on the level of individual market makers.

6.3 Process Overview and Scheduling

In addition to agent's "awake" state being scheduled, the agents can be "waken up" if their state changes.

6.4 Design Concepts

The model is implemented in Java. Multiple design patterns have been used in the development of the model, such as the strategy pattern and the visitor pattern.

6.5 Initialization

The model can be initialized using "stylized" data and parameters or it can be calibrated to the audit trail data for individual stocks.

6.6 Input Data

Audit trail data.

6.7 Submodels

Market makers and investors are represented as independent decision-making entities. The fundamental value of security is represented as a stochastic process.