AN AGENT-BASED FINANCIAL SIMULATION FOR USE BY RESEARCHERS

Roy Hayes
Andrew Todd
Nachapon Chaidarun
Scott Tepsuporn
Peter Beling
William Scherer

Department of Systems and Information Engineering
University of Virginia
151 Engineer’s Way
Charlottesville, VA 22904, USA

ABSTRACT
Regulators and policy makers, facing a complicated, fast-paced and quickly evolving marketplace, require new tools and decision aides to inform policy. Agent-based models, which are capable of capturing the organization of exchanges, intricacies of market mechanisms, and the heterogeneity of market participants, offer a powerful method for understanding the financial marketplace. To this end, we have worked to develop a flexible and adaptable agent-based model of financial markets that can be extended and applied to interesting policy questions. This paper presents the implementation of this model. In addition, it provides a small case study that demonstrates the possible uses of the model. The source code of the simulation has also been released and is available for use.

1 INTRODUCTION
An agent-based model is simulation of individual decision-makers. The decision-makers, referred to as agents, interact based upon their individual rules of behavior and the mechanisms or physics of the system (Macal and North 2010). For introductions to agent-based modeling see, e.g., Epstein and Axelrod (1996), Grimm and Railsback (2005), Miller and Page (2007). Agent-based models have origins in a number of fields including computer science, experimental economics, operations research and physics (Chen 2012). For a survey of early work in agent-based financial markets see LeBaron (2000). A recent survey is Chen, Chang, and Du (2012). This brief introduction will focus on origins and applications in finance and economics.

Schelling’s model of segregation is the classic example of agent-based models in economics (Schelling 1971). Schelling models neighborhoods as a cellular automaton. Two types of homeowners begin randomly distributed on a two dimensional grid. Homeowners relocate to empty locations based on a preference for similar neighbors. The model provides computational evidence that weak preferences for living with similar neighbors can have strong consequences for segregation and the composition of neighborhoods (Clark and Fossett 2008). The model demonstrates a fundamental notion in agent-based modeling: small changes in micro-level behavior of individuals may have large affects macro-level outcomes.

Experimental economics is also an important precursor to agent-based modeling. An important example is the prisoner’s dilemma tournament held by Axelrod, which provides experimental evidence for the evolution of cooperation (Axelrod 1984). An example in the context of financial markets is the tournament organized by Rust at the Santa Fe Institute in 1990. Participants submitted programs to compete in a double auction
for prize money. Analysis found that a simple rule-of-thumb outperformed advanced algorithms based on statistical predictions of future prices (Rust, Miller, and Palmer 1994). In fact, the properties of the double auction are studied in one of the earliest examples of agent-based financial markets.

Gode and Sunder coined the term “zero-intelligence” in a study in which they evaluate the efficiency of a double auction with computerized agents (Sunder and Gode 1993). The experiment replaces humans with agents programmed to behave randomly with respect to budget constraints. The experiment demonstrates the inherent efficiency of the auction mechanism. The zero-intelligence approach avoids the difficult task of modeling human behavior and is an important early influence for the models discussed in this paper (Ladley 2012).

This paper describes a novel agent-based simulation of the equity futures market and presents possible applications. The principal contribution of this work is the addition of intelligent market makers. A small case study examining low-latency trading is presented. This case study is meant to demonstrate the possible applications of the model in developing insight about market behavior. The source code for the model can be found at https://github.com/uva-financial-engineering/JinSup.

The paper is organized as follows. Section 2 gives a brief overview of previous financial agent-based models. Section 3 provides a description of the new agent-based model and the techniques used to validate it. Section 4 presents a brief case study demonstrating the possible uses of the ABM. Finally, Section 5 offers concluding thoughts and an outline of future research objectives.

2 PREVIOUS FINANCIAL AGENT-BASED SIMULATIONS

An early numerical simulation of an order-driven market is Maslov (2000). The model attempts to provide a computational explanation for statistical properties of “short-time fluctuations in prices.” The model is able to capture a remarkable number of interesting properties with a minimal number of assumptions. Although the model is a simple numerical simulation, it can be understood and discussed in the context of agent-based modeling. A new agent arrives to trade at each epoch of the simulation. The agent places an order to buy or an order to sell with equal probability. The agent then chooses to trade using a limit order with probability $\theta$ or a market order with probability $1 - \theta$. The price of the limit orders are chosen uniformly at random from a discrete set relative to the last traded price. A new bid at time $t$ takes a price in the discrete set $\{l(t) - d\tau, \ldots, l(t) - \tau\}$, where $\tau$ is the tick size. Similarly, a new offer takes a price in the set $\{l(t) + \tau, \ldots, l(t) + d\tau\}$. The parameter $\delta$ governs the cardinality of the set of prices.

As mentioned, the Maslov model is notable for its ability to capture several interesting statistical properties of price series, most notably, volatility clustering and power-law distributed price increments. The Maslov model is representative of the parsimonious statistical models common in “econophysics” (Chakraborti et al. 2011). The order flow is a statistical process and not the result of strategic behavior.

The model, however, is not suitable for general computational experiments. Although the price series exhibits interesting statistical properties, the underlying order book is unrealistic. In many markets, there is significant depth in the market throughout the day, with the exception of time periods leading up to major news announcements. The order book in the Maslov model, however, is very sparse and at times does not even have an established spread.

The assumption that traders place limit orders at uniform distance away from the last trade price is an incorrect assumption in today’s market. Paddrik et al, illustrated high frequency traders, which make up 70% of the trades in the equity market and 35% of trades in the E-mini S&P 500, place 60% of all their orders 1 tick or closer to the last trade price (Paddrik et al. 2012). Additionally, cancellation of orders was a large contributor to the flash crash (Kirilenko et al. 2011).

Challet and Stinchcombe examine Island ECN order book data from a physicist’s point of view (Challet and Stinchcombe 2001). The model they propose is a particle system model; however, it can be interpreted as a simple agent-based model. The mass of the particle is the size of the order and the price is the spatial position on a one-dimensional lattice. Placing an order is a deposition, an order cancellation is an evaporation, and a crossing limit order is an annihilation. At each time step an order is placed with
probability p and a price is drawn from a normal distribution around the best quote. Depending on the parameters the model is able to produce empirically observed characteristics of price returns and volatility clustering.

Several researchers have implemented evolutionary agent-based financial models (LeBaron, Arthur, and Palmer 1999; Arifovic and Gencay 2000; LeBaron 2001). Traders that are not profitable will either evolve or withdraw from the market. The decision rules of each agent are augmented based on the agent’s past profitability. The augmentation is controlled by a genetic algorithm (Holland 1992). These markets are capable of returning real world market characteristics, such as volatility clustering and excess returns. However, do to their pseudo-random nature it is difficult to interpret cause and effect (Cont 2007). Interpretability is a key requirement for decision makers and practitioners.

Recently Paddrik et al, developed a zero-intelligence model where the empirical distribution of the orders were derived from actual market participants (Paddrik et al. 2012). This model was the first attempt to empirically match real trader actions seen in the equity futures market. Paddrik et al, used the model to examine the flash crash. They found the events of the flash crash specified in Kirilenko et al, report could be replicated using zero-intelligence agent-based simulation (Paddrik et al. 2012). In Hayes et al, the aforementioned simulation was applied to the rule making process. The minimum quote life rule, recently adopted in Europe, is examined using the simulation. The authors found that the minimum quote life rule trades off time between executions and volatility. This was a novel finding because it was previously thought the rule would trade off bid-ask spread size and volatility (Hayes et al. 2012).

Paddrik et al, and Hayes et al, illustrated that a zero-intelligence simulations can be used by regulators to examine historical market events and study the effects of potential regulations. However, their simulation suffers from computational limitation due to the computing framework in which it is implemented. Several assumptions were made to allow for replications to be run in a timely manner. The main issue addressed in this paper is the previous simulation did not allow a traders to maintain a portfolio of multiple orders. By lifting this assumption more realistic market making strategies can be applied to the book. Secondly, the original simulation was $\frac{1}{32}$ scale which added inconsistencies due to the lower level of liquidity. The new simulation presented in this paper is a quarter scale model, which allows closer replication of real world phenomena.

3 MODEL STRUCTURE

3.1 Agent-Based Financial Market

The goal of the financial agent-based model is to replicate real world order-flow. To accomplish this the agents are designed to mimic actual practitioners in the S&P 500 E-Mini futures. With this goal in mind, agents are split into six different categories, which are originally defined in Kirilenko et al. (2011). Paddrik et al., and Hayes et al., demonstrate that heterogeneous agents are necessary to model the complex interactions of market participants. Below is a description of each class of agent.

**Small traders** are low frequency traders. They do not trade a lot or take large positions. Additionally, with equal probability they will accrue a long or short position over the duration of the simulation.

**Fundamental Buyers** are traders that seek to accrue a large long position over the duration of the simulation.

**Fundamental Sellers** are traders that seek to accrue a large short positions over the duration of the simulations. Fundamental buyers and sellers can be thought of as large institutional investors.

**Opportunistic Traders** seek to make profit off private information. This information can take the form of statistical arbitrage, sentiment analysis, etc. A single stochastic signal is given to all opportunistic traders, representing this private information. This signal changes the probability that opportunistic traders’ place a buy versus sell order. Below is the signal equation,

$$P_{buy_t} = MAX(30\%, MIN(70\%, P_{buy_{t-1}} + U[-20\%, 20\%])).$$
Unlike the aforementioned traders, opportunistic traders have position limits. In other words, opportunistic traders will not take an absolute position larger than a defined threshold. These thresholds are found empirically (Paddrik et al. 2012). Additionally, they seek to end the day with little to no inventory.

**Market Makers** place orders on both sides of the market and attempt to profit off the bid-ask spread. They have position limits and will trade aggressively (cross the bid-ask spread) if they reach their position limits. Additionally, market makers will attempt to end the day with no inventory.

**High Frequency Traders** are modeled as faster, more aggressive market makers. Their probability of placing a buy or sell order is altered based on the difference between best bid and ask. The equation below gives the probability that the next order will be a buy order,

\[ P_{buy} = \frac{Q_B}{Q_B + Q_S} \]

Where \( Q_B \) is the quantity at the best bid price and \( Q_S \) is the quantity at the best ask. This behavior was observed in data provided by the Commodity and Future Trade Commission and reported by Brogaard (Brogaard 2010). HFTs also have strict position limits and end the trading day with little or no inventory.

The agents are modeled as zero-intelligence, meaning their actions are controlled by empirically derived probability distributions. All probability distributions were originally estimated for previous simulations and are separately estimated for each class of traders. For further explanation on how these distributions were estimated please refer to Paddrik et al. (2012). The random distributions control order size and placement, as well as new order arrival and order cancellation rates. Figure 1 illustrates how these distributions are combined to create an agent class.

The simulation is a discrete event model, with a resolution of 1 millisecond. At each time step a set of agents, which are eligible to place and/or cancel an order, is returned. Agents in this set are selected in random sequence and place and/or cancel an order according to their underlying random distributions. These distributions control the size of the order, the price placement of the order, whether the order is a buy or a sell, as well as the cancellation rates. A trade automatically occurs if the order crosses the bid-ask spread. The next section will report validations performed against the S&P 500 E-Mini futures.

### 3.1.1 Financial Agent-Based Model Validation

As a method for validating the simulation, the statistical properties of the simulated markets are compared to the S&P 500 E-Mini market data. Comparing a vector of statistics of the simulated data to a vector of statistics from the real market is just one way of validating an agent-based model of financial markets. Validating agent-based financial markets is still an active area of research. The statistical properties checked in our validation are the non-normal distribution of returns, lack of autocorrelation in returns, volatility clustering and aggregational normality.
Mandelbrot observed the fat tails of the distribution of prices changes and that sample second moment typically varies in an erratic fashion (Mandelbrot 1963). This has caused various suggestions regarding the form of the distribution, ranging from the Student-t, hyperbolic, normal inverse Gaussian, and others, but no general consensus exists for the form of the tails for all markets. That is to say, the returns observed in financial markets are non-normal at the tails of the distribution. Figure 2 below compares the simulation returns to the real market. Notice that both distributions are non-normal at the tails of the distribution.

![Normality Test of Emini S&P 500 Returns](image1)

**Figure 2: Normality comparison.**

A criterion for an accurate representation of a market is that returns are not autocorrelated. To test this, we run a test for autocorrelation of returns. Figure 3, compares the simulations return autocorrelation to that of the real market. It should be noted that both the real and simulated market have statistically high autocorrelation for several minutes, before dying out. These autocorrelations are generally due to microstructure effects such as the bid-ask bounce. However, over a longer time horizon returns are not autocorrelated.

![Autocorrelation of E-Mini S&P 500 Minute Return](image2)

**Figure 3: Autocorrelation of returns comparison.**

Volatility clustering may be measured by the autocorrelation of absolute returns. This was first noted by Mandelbrot and was finally translated into agent-based models by Kirman and Teyssi`ere when they discovered a model would exhibit autocorrelation patterns in the absolute returns if a variable was used to herd positive or negative opinion (Mandelbrot 1963, Kirman and Teyssi`ere 2001).

\[ \text{Herd}_t = \text{Herd}_{t-1} + \text{unif}(-N,N) \]
The simulation implements the same herding mechanism to influence the decision of opportunistic traders in the model. This creates a similar the autocorrelation pattern in the absolute returns that is seen in the S&P 500 E-Mini futures minute price returns. Figure 4, compares real world and simulated autocorrelation of absolute price returns. It should be noted that the autocorrelation pattern does not always occur. This indicates that the model is capable of weak volatility clustering.

![Autocorrelation of E-Mini S&P 500 Absolute Minute Returns](image1)

![Autocorrelation of Simulated E-Mini S&P 500 Absolute Minute Returns](image2)

Figure 4: Autocorrelation of absolute return comparison.

Lastly, the simulation is tested for aggregational normality. Aggregational normality states that as one increases the time scale over which the returns are calculated, the distribution approaches the Gaussian form. This cross-over phenomenon was noted by Kullmann et al. (1999) where the evolution of the Pareto exponent of the distribution with the time scale is studied. Figure 5, only shows this phenomenon in the simulated data. Since market data with time resolution of a minute was used as the comparison it was not possible to demonstrate aggregational normality on the real world data. However, since the primary goal is to validate the simulation we felt it was sufficient demonstrate aggregational normality in the simulation.

![Aggregational Normality Comparison](image3)

Figure 5: Aggregational normality comparison.

The next section describes the intelligent high frequency market maker strategy examined in this paper. This demonstrates the capability of this simulation to incorporate intelligent traders.

### 3.2 Intelligent High Frequency Market Maker

A market maker is a broker-dealer that accepts the risk of holding shares or contracts in specific financial markets, which facilitates trading in these markets. Market makers place buy and sell limit orders. The
goal of the market maker is to offset each completed buy order with a corresponding sell order in the process capturing the difference between the best bid and ask. Market makers place strict rules on inventory management because they are susceptible to large price movements. Therefore, market makers try to avoid trading with informed traders. Informed traders are defined as traders that trade in the direction of a persistent price change. This is also known as adverse price selection.

Due to the competitive nature of electronic trading, specific strategies are not generally known. However, using work done by Cont, Kukanov, and Stoikov (2013) and Brogaard (2010) a HFT market making strategy was generated. Brogaard found that there exist a class of HFTs that passively buy and sell on both sides of the market (Jones 2013). These HFTs resemble classic market makers and generally make profits by capturing the spread. Cont et al, found that up to 76% of price change can be attributed to the difference between the best bid and ask.

Using this information a passive HFT market making strategy was generated. The strategy is completely passive (issuing no market orders or aggressive limit orders), and attempts to avoid adverse selection using a published method of detection.

1. The market maker places a single buy limit order for one contract on the top ten bid prices and does the corresponding for the sell orders.
2. The best ask (bid) order will be canceled if the bid ask volume difference exceeds (drops below) a specified threshold.
   \[-X < (Q_B - Q_S) < X\]
3. If the long (short) inventory limit is reached the market maker cancels all buy (sell) orders until inventory drops below (goes above) half the inventory limit.
4. Replace executed order unless difference threshold or inventory limits are violated.

The next section describes the experimental procedure, as well as the results.

4 CASE STUDY
4.1 Experimental Procedure

Over the last decade market participants and exchanges have increased in speed. A number of market participants now compete on speed to capture profits related to intermediating retail order flow. These market participants are generally known as high frequency traders (HFTs). Researchers have primarily focused on the effect of high-frequency trading on market quality, especially volatility. There is still no broad consensus regarding HFT and market quality (Kirilenko et al. 2011; Brogaard 2010; Zhang 2010; Haldane 2011).

Although there has been a surge in research, several questions still remain about HFTs. One puzzling question is why are there so few HFT firms? This study will seek to answer this question, by adding increasing number of the previously defined intelligent HFTs into a simulated market. The HFTs’ gross profit was examined as intelligent HFT parameters are changed. Gross profit is defined as profit made before transaction cost and operating expenses. HFT firms’ transaction cost are not public information, therefore transaction cost was not simulated. Several researchers have indicated that many HFT strategies resemble traditional market making (Brogaard 2010, Haldane 2011). Therefore, the authors’ contend that the previously defined intelligent HFT is an adequate representation of real life HFT firms.

This case study is small in scope. It is designed to demonstrate experiments that can be run using the agent-based simulation. The next section examines the results of the case study.

4.2 Results

A main effects plot was generated using the gross profit for each trial run, shown in Figure 6 below. It should be first noted that the main effect plot illustrates the HFT strategy is profitable. However, the number
of HFT market makers has a non-linear impact on the mean gross profit of firms. This indicates that there is a limited amount of money that can be made using this type of algorithm and the market opportunity is arbitraged away as the number of firms is increased. This finding provides an explanation for declining HFT profits. Given HFT strategies are known to be similar (Brogaard 2010, Jones 2013) and the possible profits of HFTs are limited, as shown in the simulation, there exist a maximum carrying capacity of HFT firms. The results indicate that this maximum carrying capacity will be a relatively low number of firms, compared to other types of trading firms.

Figure 6 illustrates the aforementioned results. Decreasing the action delay (increasing the HFT speed) has very little impact on overall profits. This is because they are already the fastest agents in the simulation. However, increasing the number of intelligent HFTs in the market, greatly impacts the average HFT profit. This indicates that the market has a carrying capacity.

![Main Effects Plot for Profit](image)

5 CONCLUSION

Policy makers are in need of a tool that allows them to study proposed regulations over a variety of market conditions. The authors’ contend that agent-based models provide regulators and policy makers with a controlled laboratory environment, in which to study proposed regulations. To this end we have developed a financial agent-based simulation.

This paper presents a financial agent-based model that is capable of incorporating intelligent agents. The source code has been released at https://github.com/uva-financial-engineering/JinSup. Researchers are encouraged to utilize and improve the model. A small case study was presented, which demonstrates how the model can be used to examine relevant research questions.

The simulation struggles to consistently produce volatility clustering. This may indicate that the herding mechanism may not be the cause of volatility clustering. Therefore, future iterations of the simulation will seek to examine how to use intelligent agents to generate the volatility clustering.

REFERENCES


Hayes, Todd, Chaidarun, Tepsuporn, Beling, and Scherer


**AUTHOR BIOGRAPHIES**

**ROY L. HAYES** is a PhD Candidate in the Department of Systems and Information Engineering at the University of Virginia. His previous research leveraged agent-based modeling and machine learning techniques to model high frequency trading. His current research examines reinforcement learning’s applicability to real time strategy games. His email address is rlh8t@virginia.edu.

**ANDREW TODD** is a PhD candidate in the Department of Systems and Information Engineering at the University of Virginia. His research interests include financial engineering, agent-based modeling of financial markets and market microstructure. His email address is aet6cd@virginia.edu.

**NACHAPON CHAIDARUN** is a recent graduate from the University of Virginia with a BA in Computer Science and Mathematics. He currently works as a software developer at Duolingo, an online language education startup. His email address is nc5rk@virginia.edu.

**SCOTT TEPSPORRN** is a BS Computer Engineering and Computer Science undergraduate at the University of Virginia graduating May 2015. He has helped plan and build multiple time domain simulations and has experience with high performance computing. His current research involves improving throughput variability of large data transfers. His email address is spt9np@virginia.edu.

**PETER A. BELING** is an Associate Professor in the Department of Systems and Information Engineering at the University of Virginia. Dr. Beling’s research interests are in the area of decision making in complex systems, with emphasis on Bayesian scoring models, machine learning, and mathematical optimization. His research has found application in a variety of domains, including lender and consumer credit decision-making, analysis of high frequency trading strategies, and man-machine decision systems. His email address is pb3a@virginia.edu.

**WILLIAM T. SCHERER** is a full professor, who has served on the University of Virginia Department of Systems and Information Engineering faculty since 1986. He has authored and co-authored numerous publications on intelligent decision support systems, combinatorial optimization and stochastic control. His current research focuses on systems engineering methodology, financial engineering and intelligent transportation systems. His email address is wts@virginia.edu.