OFFICE OF FINANCIAL RESEARCH

WORKING PAPER

14-09 | November 25, 2014

Effects of Limit Order Book Information Level on Market Stability Metrics

Mark Paddrik
Office of Financial Research
Mark.Paddrik@treasury.gov

Roy Hayes
University of Virginia
rlh8t@virginia.edu

William Scherer
University of Virginia
wts@virginia.edu

Peter Beling
University of Virginia
pb3a@virginia.edu

Effects of Limit Order Book Information Level on Market Stability Metrics

Mark Paddrik, Roy Hayes, William Scherer, Peter Beling

Abstract

Using an agent-based model of the limit order book, we explore how the levels of information available to participants, exchanges, and regulators can be used to improve our understanding of the stability and resiliency of a market. Ultimately, we want to know if electronic market data contains previously undetected information that could allow us to better assess market stability.

Using data produced in the controlled environment of an agent-based model's limit order book, we examine various resiliency indicators to determine their predictive capabilities. Most of the types of data created have traditionally been available either publicly or on a restricted basis to regulators and exchanges, but other types have never been collected. We confirmed our findings using actual order flow data with user identifications included from the CME (Chicago Mercantile Exchange) and New York Mercantile Exchange (NYMEX). Our findings strongly suggest that high-fidelity microstructure data in combination with price data can be used to define stability indicators capable of reliably signaling a high likelihood for an imminent flash crash event about one minute before it occurs.

Acknowledgments: The authors are very grateful to Philip Maymin, Michael Wellman, Nathan Palmar, Greg Feldberg and Wendy Wagner-Smith for their help and guidance with this paper. Special acknowledgement is due Andrew Todd for his help in the development of the agent-based simulation used in this paper.

1. Introduction

On May 6, 2010, the U.S stock market experienced one of the most severe price drops in its history. Known as the May 6th Flash Crash, the Dow Jones Industrial Average fell 5 percent in less than five minutes. The bulk of losses were recovered in the next 15 minutes of trading. The crash raised concerns about the stability of capital markets and resulted in a joint investigation by the U.S. Securities and Exchange Commission (SEC) and the U.S. Commodity Futures Trading Commission (CFTC) (CFTC and SEC, 2010).

The May 6th Flash Crash seems to be singular in the sense that no following events have rivaled its depth, breadth, and speed of price movement. Flash crashes on a smaller scale, however, occur frequently. Between 2006 and 2012, there were 18,520 incidents that we term *mini flash crashes*, in which a single security experienced abrupt and severe price change over a short period of time (Johnson et al., 2012). Although the underlying causes of mini flash crashes vary (Golub, 2012), they share some uniform markers of stress that, if recognized in advance, could be the basis for market intervention aimed at increasing price stability.

The metrics TR-VPIN (Easley et al., 2011), and BV-VPIN (Easley et al., 2012) were both developed to try to predict flash crashes. These metrics have considerable limitations because they were constructed using in-sample metric calibration with no out-of-sample testing, have high false positive rates, and are general ineffective at predicting short term volatility once accounting for volume dynamics (Andersen and Bondarenko, 2014). Moreover, the VPIN metrics use only transactional information and do not take into account even basic information on about the limit order book, which is the predominant trading mechanism in today's electronic equities and futures markets. However, it has been observed that the limit order book offers a structure that makes it possible to examine strains on the state of price formation and overall market stability (Duong, 2014).

In this paper, we examine the hypothesis that the quality of market stability metrics can be improved by making use of order flow information. A central point of investigation is to assess the increases in metric quality associated with moving from publically available data (such as Level 2 quotes) to increasingly more detailed data such as that held by regulators, which may contain trader identification or other restricted information. This hypothesis is difficult to test empirically, because much of the historical data about the inner workings of the limit order book, often called microstructure data, is available only to regulators and exchanges and, in many cases, is too complex and time consuming to examine. But regulatory attention to this process is needed, because price discovery is viewed as the most important product that security or futures markets offer and must be treated as a public good, essential to financial stability (Hasbrouck, 1995).

A number of stability metrics are considered in this paper, including several designed specifically to make use of fine-grained information such as Level 6 data. Stability metrics are assessed on their ability to predict impending destabilization in the form of flash crashes. We tested market data produced by an agent-based simulation model that incorporates a limit order book mechanism and associated microstructure data at all levels. The simulation model has zero-intelligence agents, and so is limited in fidelity but is tuned to have styled facts similar to those of the historical market. Our key conclusion is that metrics using more detailed market microstructure data had the greatest power to predict flash crashes in the simulated market. To validate the conclusions from the simulation, we used a small set of actual market data to show that the metrics most effective on the simulated data were also most effective in predicting the May 6, 2010, Flash Crash of the E-Mini S&P 500 futures market and the September 17, 2012, mini flash crash of the Light Sweet Crude Oil futures market one minute in advance.

The paper is organized as follows:

- Section 2 gives general background on the electronic order book and what occurs when it becomes unstable in advance of and during flash crashes. The concepts introduced in this section are foundational to the stability metrics proposed in Section 4.
- Section 3 describes the agent-based simulation model used to produce the principal datasets used to test the stability metrics.
- Section 4 covers how different data levels may be used to determine possible measures of market stability.
- Section 5 presents results for testing of the stability metrics on the simulated data simulated.
- Section 6 describes the ability of each stability metric to predict two historical flash crashes.
- Section 7 summarizes our results and conclusions, including a discussion of the opportunities that this method may offer to operational risk management.

2. Background

The rapid and wide adoption of computer and communications technology during the early 1990s drove many exchanges interested in increasing market participation and decreasing costs to build electronic platforms for trade. The majority of financial trading activity moved off the physical trading floors and on to electronic exchanges. Trading automatically using sophisticated computers and algorithms enabled the exchanges to offer a wider set of financial instruments, which was once restricted by physical floor space.

This transition to electronic exchanges required a new way to discover the price of an asset in market format that would allow participants to interact in a manner similar to that of the physical trading floor. The result was the electronic order book system of trade, which now dominates the market. The change has been widely accepted and is generally seen as effective.

Along with this transformation, however, new concerns for regulators arose about market participant behavior and events such as flash crashes. These concerns established the need for regulators to develop the capability to leverage the vast amount of electronic data to evaluate participants' impacts and prevent similar flash crash events.

Market stability is heavily dependent on the mechanism of trade and price formulation and examining the data produced from the electronic order book mechanism may provide the information needed to prevent future price destabilizations. The first step is to understand how the electronic order book functions.

The Electronic Limit Order Book Market

The electronic order book, the main mechanism for price discovery and trade on exchanges, is a publicly visible mechanism providing bids and offers. The order book contains a price and a quantity of limit orders (Glosten, 1994). A trade occurs when a new order is entered or an existing order's price is modified so the order crosses the spread of best bid and best ask offers. The new or modified order transacts against the best opposing orders using a price-time priority, removing them from the order book until the new or modified order is either fully or partially filled, leaving the rest of the quantity of the order to become the new best bid or offer. Other order types are not publicly visible, such as stop-loss and max-show orders that follow additional rules. However, all end up becoming limit orders before execution.

Price discovery results from the deliberation between the buyers and sellers who place, modify, and cancel orders, leading to trades when an agreeable price can be found. The individual decisions and interactions between market participants generate the underlying microstructure of the limit order book, which then results in the pricing of assets. Figure 1 gives an example of how the market typically responds to reflect this price discovery process. Although the electronic order book and its rules are relatively simple, the behavior of a market is complex, because it includes a large number of participants who interact with each other stochastically, making analysis difficult (Anderson, 1988).

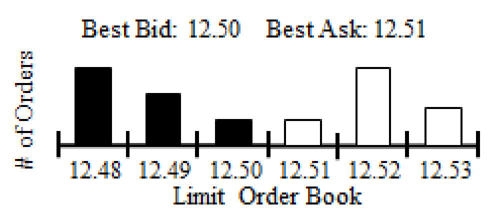
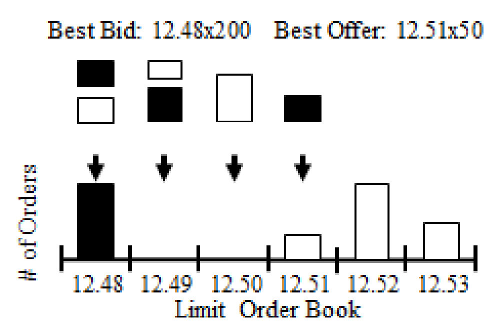
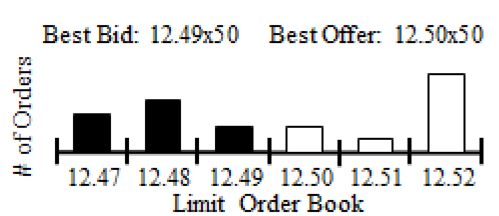
<table>
<tr><th>Step</th><th>Event</th><th>Order Book Plot</th><th>Last Trade Price</th></tr>
<tr><td>1</td><td>A relatively large sell order arrives that crosses the bid-ask spread of $12.50/$12.51, consuming limit buy orders from $12.50 to $12.49</td><td>
</td><td>$12.50</td></tr>
<tr><td>2</td><td>Prices adjust sharply to the executed trade, leaving a $0.03 bid-ask spread between $12.48/$12.51</td><td rowspan="3">
</td><td rowspan="3">$12.48-$12.51</td></tr>
<tr><td>3</td><td>The widened spread represents a greater uncertainty of the price of the asset</td></tr>
<tr><td>4</td><td>The wide spread is followed by a surge in new orders at various price levels, representing the price discovery with increased certainty</td></tr>
<tr><td>5</td><td>As new orders are introduced, the bid-ask spread narrows and a new equilibrium price is reached</td><td>
</td><td>$12.49</td></tr>
</table>

Figure 1: The Dynamics of the Limit Order Book

Market Design Problems and Controversy

The price discovery process of the limit order book depicted in Figure 1 does not always settle in an orderly manner. The market can suddenly lose apparent resiliency, resulting in prices soaring or plunging for short periods of time. These events, mini flash crashes, are characterized by the price of an asset drastically changing for short periods of time until the market can recover.

The recovery process can happen in two ways. The first, and most common, allows market forces to reach an equilibrium price naturally through the continual placing of new orders into the pricing mechanism. The second is an intervention, known as a circuit breaker, where the exchange pauses trade for a predetermined amount of time, stopping the price mechanism. This gives market participants more time to observe market conditions and generate new orders to bring equilibrium back to the pricing discovery process (Subrahmanyam, 1994).

Although such pricing problems occurred some 9,000 times between 2007 and 2010 (Nanex, 2011), they were not seen as a concern because they could be stopped through the circuit breaker process and their impact was minimal. However, in 2010, one of these destabilizing events made headlines when most major equities, futures, and commodities markets plunged in near unison, causing more than a trillion dollars of asset value to be lost in less than five minutes before recovering.

The CFTC-SEC staff report on the events of May 6th Flash Crash identified an automated execution algorithm that sold a large number of contracts as one catalyst for the crash. The algorithm, which was run on the E-mini S&P 500 futures market, kept pace with the market selling about 9 percent of the previous minute's volume. This process, though used previously with no known negative impact, triggered a cascade of selling by firms using high-speed automated trading (AT), which led to "hot potato trading" where these firms rapidly acquired and then liquidated positions among themselves, resulting in steadily declining prices (Kirilenko, 2011).

These events unfolded as described in Figure 2.

Event	Model Price Points	Effect
Large sell execution algorithm initialized		Large aggressive sell orders enter the market
Automated trading systems begin to retreat from the market as prices fall	24 ticks below the moving average	Market depth begins to disappear
Fundamental traders withdraw from the market as stop-loss orders are triggered	70 ticks from price at start of day	Market depth disappears and more sell orders are executed
A market pause is initiated by the exchange	Price falls at least or more than 1.3 percent per second per second	Gives time for slower market participants to enter trades into the order book

Figure 2. Table of events that occurred during the Flash Crash of May 6, 2010.

The effect of this event led to heavy speculation and controversy about how automated trading participants trade in markets and whether they may threaten market stability. Though automated systems had been credited with increasing trading volumes and the narrowing of price spreads on markets, which have traditionally been seen as indicators of good market quality (Menkveld, 2010; Hendershot, 2011), others have questioned the value of such systems. Zhang (2010) showed there is a negative correlation in the relationship between a market's ability to incorporate information about firm fundamentals into asset prices and the presence of automated trading.

Zhang's work suggests automated trading orders may have a negative influence on the market's ability to price assets by obscuring information in the order book. Since the majority of high-frequency automated trading market participation orders are placed within the first few ticks of the best bid-ask (Clark-Joseph, 2013), automated trading orders can narrow price spreads while giving the appearance of liquidity. However, when actually transacted, prices may prove to be fleeting and less liquid than indicated.

Electronic Market Data

Market participants, exchanges, and regulators understand the state of electronic markets and pricing of assets by examining market data. Financial exchanges give different degrees of data on markets with varying levels of privacy and informational content available. Traditionally, the degree of information given or sold by exchanges can be broken down into levels of information associated with the state of the electronic order book (O'Hara, 2005; Bethal, 2011). We have broken down this information into six levels in Figure 3.

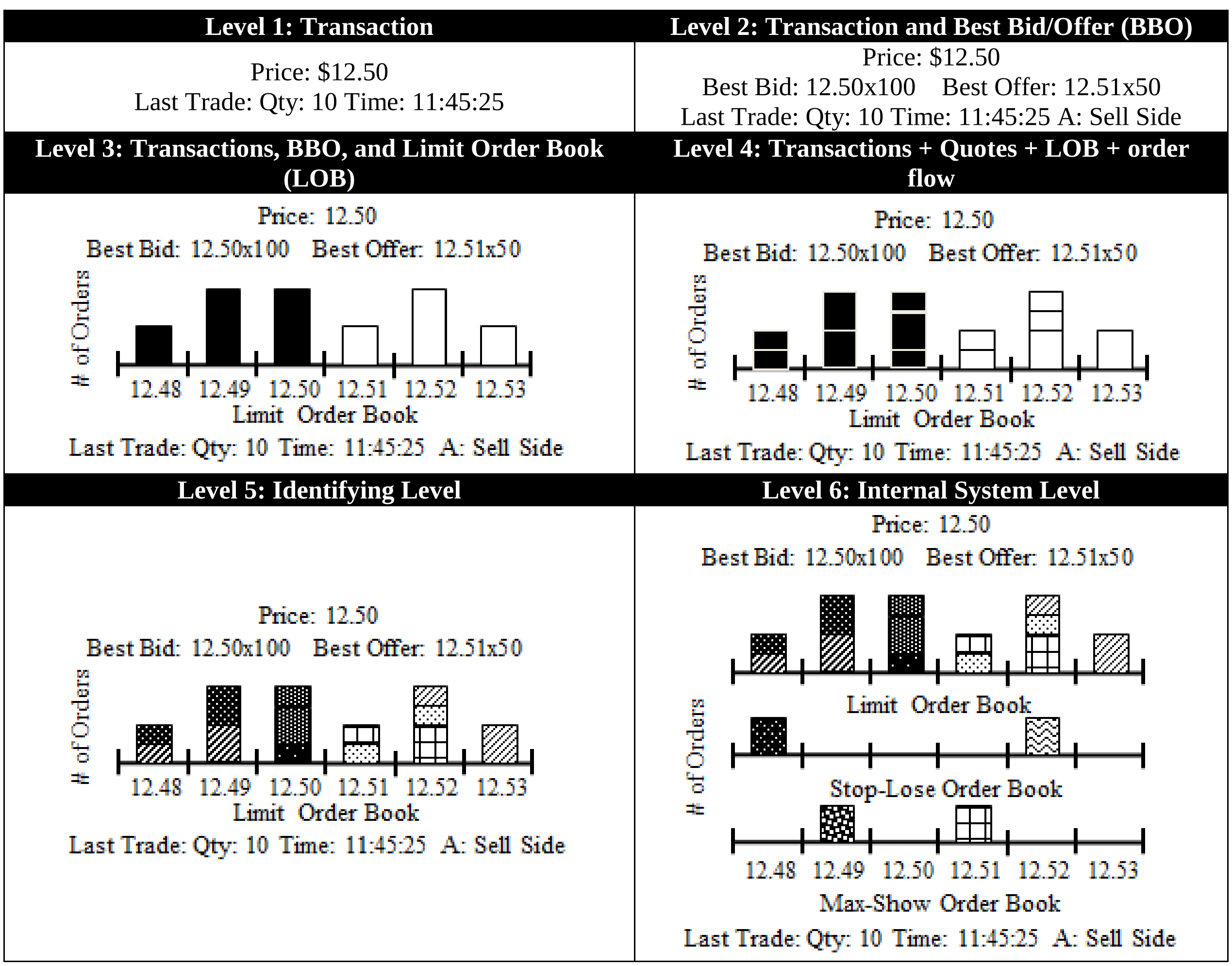


Figure 3: The Levels of Financial Market Data

Levels 1 and 2 typically are available on public market feeds that allow participants to examine trade prices and quoted best bid and offer prices with the demand at that particular price, which are the result of aggregated order flow communications between an exchange and individual participants of a market. The distinction between Level 1 and Level 2 is that the second level gives the participant some underlying knowledge about the state of the market by giving the best bid and ask offers instead of using the last reference point of trade price.

Bid Quantity	Bid Price	Ask Price	Ask Quantity	
50	$10.56	$10.57	140	
36	$10.55	$10.58	67	
142	$10.54	$10.59	89	
32	$10.53	$10.60	52	
49	$10.52	$10.61	103	
100	$10.51	$10.62	40	
110	$10.50	$10.63	205	
65	$10.49	$10.64	178	
258	$10.48	$10.65	245	
178	$10.47	$10.66	90	
Last Trade:	**Price**	**Quantity**	**Time**	**Aggressor**
	$ 10.57	10	12:56:24.047	Buy Side

Figure 4: Example of Level 4 Data

Levels 3 and 4 are publically available data sets, but are traditionally distributed at a cost to users. These data levels provide a snap shot view of the limit order book out to some depth (typically 10 units on both sides of the best bid and offer price). Figure 4 shows an example of what a market participant receiving Level 3 data quotes would see. Level 4 separates the aggregated snap shot view of the limit order book depths data of Level 3 by providing the raw order flow data that is used fill the limit order book mechanism.. The order flow data contains new, modified, and canceled limit orders place into the mechanism.

Level 5 and Level 6 information is commonly kept private and is available only to exchanges and regulators to verify that market participants are complying with rules and processes. Level 5 data provides identification information along with order flow data, so participants know who placed a specific order, canceled or modified an order, or who traded a security. (In Level 4, this order flow data is anonymized.)Level 6 includes all order flow and order book data, regardless of order type (marketable, limit, stop-loss, max-show, etc.), which makes it possible to track and account for all events that occur in a market. Such detail allows users to create a picture of the state of the market at any given time.

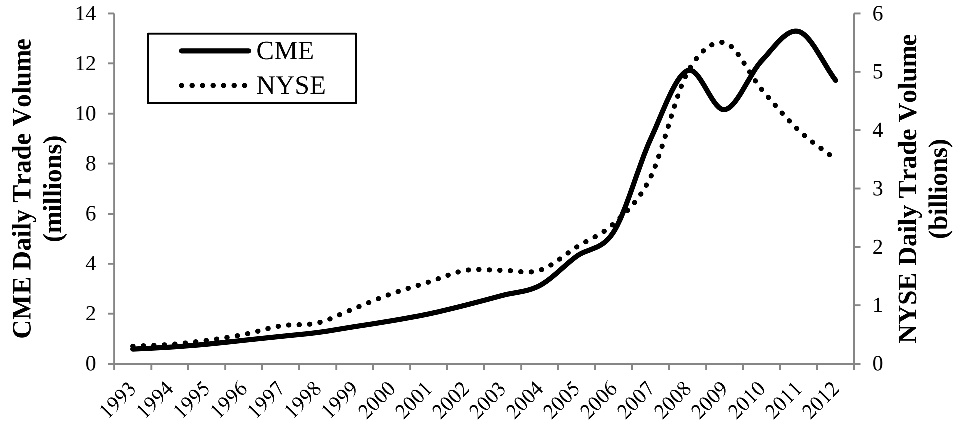


Figure 5: Average Daily Trade Volume on the New York Stock Exchange and the Chicago Mercantile Exchange

These data levels give various depths of information about the market, but they can be complicated to monitor because of the enormous volume of U.S. financial trades per day, as seen in Figure 5. Trade data itself represents only about 1 percent of total order flow data, which contains new orders, cancel orders, modify orders, and trade data. The message traffic of order flow data is about 100 times larger than trade data alone. (CME Group, 2013). Because of the large quantities of data, users — market participants, exchanges, and regulators — typically will only use a subset of the total data available to them. For example regulators generally use a combination of Level 2 and Level 3 data, with the addition of some private account information when investigating market events.

In addition, much of this type data is not stored because higher level data is extensive and complex, requiring tremendous amounts of computational processing capacity, and such data collection and storage is costly. Consequently, large quantities of useful information that might prove helpful to regulators' study of unexpected or destabilizing market events, such as flash crashes, are not available.

Simply requiring regulators to collect market data at higher levels (such as levels 4, 5, or 6), when the cost-benefit has not been proven, would not be feasible. Building and analyzing market designs through simulation modeling offers a lower cost and technically effective experimental platform for evaluating the informational value of higher fidelity of financial market data.

3. The Agent-based Model of the Limit Order Book

The transition to electronic order books and automated trading has been an impetus for looking at market microstructure with new lenses (O'Hara, 2014). Market simulations offers an experimental environment for examining features and characteristics of interests to academics and regulators by providing a controlled way to create representative market data useful for analysis (Hayes et al., 2014). Because little level 4, 5, or 6 electronic order book data is readily available, particularly to the academic community; simulation is an especially necessary tool.

Agent-based modeling offers the capable of capturing the organization of exchanges, the intricacies of the trading process, and the heterogeneity of market participants, and is thus a powerful method for analyzing financial market (Bookstaber, 2012). Agent based models (ABM) simplify complex systems by including a set of individual agents, a topology and an environment (Farmer, 2010).

In a typical ABM of a financial market, the market participants are agents, the market mechanism is the topology and the exogenous flow of information that is relevant to the market is the environment. As a system's complexity increases, the ability to directly correlate its macro-level behavior with the changes in the underlying micro-level behavior and parameters diminishes. This requires that the micro-level behavior and parameters be more fully examined and defined to understand their relationship to the macro outcomes. ABM can provide regulators with an experimental environment that can comprehend complex system outcomes.

For this reason, designers of ABM's face a trade-off between complexity and explanatory power. In the context of financial markets, for example, agents are endowed with varying degrees of sophistication with regard to how they adapt, predict and optimize within their environment. To create an ABM of a financial market, it is important to determine what features are needed for creating a proper representative market without overly complicating the framework. Because a market is a system constructed of the participants and the rules governing their interaction (e.g. trade and pricing), it is necessary to focus on these two parts and set aside, for simplification purposes, the uncertain environment aspect which has indefinite set of conditions. Focusing on the agents and the topology is important to note, because it presumes that the market participants in this simulation will focus on their individual conditions and the single market in which they operate and not on environmental variables like other markets or economic condition that could affect their decisions. We will later implement an agent to help create the environmental conditions of the crash.

3.1 The Agents

Of the two components our model tries to capture, the agents are more difficult to specify as their design needs to mimic that of true market participants. The difficulty of balancing specificity of traders' characteristics while still making agents that are representative of the general population is difficult. Current literature suggests that the markets are divided into subcategories of participants and the combinations of trading styles that are responsible for emergent market events. These combinatorial aspects led to the design of multiple categories of trading agents in the simulation. From work done by Kirilenko et al. (2011), it was possible to characterize market participants in to one of six classes of agent types, though in real life a trader may play multiple roles:

i. ***Fundamental Buyers**:* take long positions on the asset and trade with a low frequency.

ii. ***Fundamental Sellers**:* take short positions on the asset and trade with a low frequency.

iii. ***Small:*** randomly take small long or short positions (50%/50%) on the asset and trade with a very low frequency.

iv. ***Market Makers:*** take the position of straddling both sides of the market by taking long and short positions on an asset. These intermediaries' are generally seen as liquidity providers to markets through their placement of orders on both sides of the order book.

v. ***Opportunistic Traders:*** take long or short positions on the asset during the duration of the market day like a fundamental participant. However, they implement trading strategies that make them resemble Intermediaries because they do not take large positions.

vi. ***High Frequency Traders (HFT):*** take long or short positions on an asset for short periods and trading with high frequency near the best-ask and best-bid sides of the book. HFTs are generally described as simple momentum strategy (Clark-Joseph, 2013). As the bid/ask depth becomes imbalanced, HFTs will tend to trade in the same direction as the imbalance. In other words, if there are excess bids in the order book, HFTs have a higher probability of placing a buy order; such a strategy was indicated in real market data provided by the CFTC. In general, HFT's will allow themselves to take large positions for short periods of time, but will try to be neutral by the end of day.

Traditional economic approaches to build agents focuses on the role of intelligent (rational) agents that seek to maximize a utility function. However, this assumes that humans have perfect information and adapt instantly. People are often less than rational, which is a topic of recent interest of behavioral economics. Furthermore, rational economic theory relies on the abstraction of equilibrium methods to solve these problems, which has been demonstrated to be non-existent (Farmer and Foley 2009).

However as early as the 1990's zero-intelligence agent descriptions have been used to represent aggregate market participant behavior by using random sampling of empirically generated distributions to characterize the agent decisions. First introduced by Gode and Sunder (1993) in a double auction market, they found that the allocation efficiency of the market derived larger from its structure and was indecent of trader's motivation, intelligence, or learning. This methodology has come to dominate the ABM limit order book literature with several expansion of the models (Maslov, 2000; Challet and Stinchombe, 2001; Farmer et al., 2005; Preis et al., 2006) and found to explain the majority of general dynamics like spread variance and price diffusion rates.

Our model will consider three such zero intelligence characteristics, based on empirical data of the S&P 500 E-mini futures market, three days prior and including the May 6th 2010 crash[1], to describe the agents in this model:

1. Order Arrival Rate: A new orders is placed by an agent by an arrival process that follows a Poisson distribution with a mean observed from the empirical order submission rate of a class.[2]

2. Order Size: An order's size is randomly selected based on the empirical distribution of order sizes of a class.

---

[1] The data and the agent distributions used in the construction of the limit order book ABM cannot be made public out of concerns of providing detailed trading strategy secrets on order size and placement.

[2] Agents only manage a single order at a time. If an agent has an old order that is still in the order book at the time it is scheduled to place a new order, it cancels the old order before adding the new order to the market.

3. Order Placement: An order's price is randomly selected based on the empirical distribution of orders placed based of the distance from the current best bid and best ask of a class.

Additionally we must know two other characteristics to fully characterize the agents in the definitions stated above which requires using features from the near-zero intelligence literature (Paddrik et al., 2012; Wah and Wellman, 2013). These characteristics include which side of the book to place an order (i.e. place a buy or sell order) and whether the agents have limits on their inventory. These characteristics are partially connected because the probability of a buy or sell order depends on how close an agent is to a position limit.

1. Position Limit: The decision to give an agent a limit is based on the averages derived from Kirilenko et al. (2011).[3] From these observations *Market Makers, Opportunistic Traders, and High Frequency Traders* all appeared to have some form of risk control built into their behavior that limited the positions that they take throughout the day. A governing algorithm was placed on top of a random uniform distribution for the decision to buy or sell, as seen below. The algorithm provides position limits control to all these agents but allows the process to be applied to all the agent

$$P(Buy) = \begin{cases} 1 - \dfrac{CurrentPoisition}{MaxPosition} & CurrentPoisition > \dfrac{MaxPosition}{2} \\ \dfrac{CurrentPoisition}{|MinPosition|} & CurrentPoisition < \dfrac{MinPosition}{2} \\ 0.5 & otherwise \end{cases}$$

$$P(Sell) = 1 - P(Buy)$$

2. HFT Buy Sell Decision: Additionally this class of trader would use the order book depth of the best bid and best ask to predict price, a trend that follows with work done by Cont et al. (2013). This replace the random uniform distribution decision to buy or sell when $CurrentPoisition$is between $\frac{1}{2}MaxPosition$ and $\frac{1}{2}MinPosition$ with:

$$P(Buy) = \frac{BestBidDepth}{BestBidDepth + BestAskDepth} \qquad P(Sell) = \frac{BestAskDepth}{BestBidDepth + BestAskDepth}$$

3.2 The Market Topology

The design of the simulated market exchange systems follows a traditional "price-then-time" order book market using a set of agents to insert transactional messages to the order book. By proving this, topography asset price creation on the part of the market participants can take place through their individual actions (order, cancel), and the market matching engine connecting them (execute).

[3] See Appendix A for exact numbers.

This format is consistent with a double auction market traditionally used in financial and commodities markets (Friedman and Rust, 1993).

3.3 Model Validation

With the agents and the correct market topology in place,[4] the model was validated in Paddrik et al. (2012) to show it resembles the manner in which that market typically functions by assessing the "stylized facts" (Kaldor, 1961) of the pricing series associated with that market. Stylized facts are a set of statistical characteristics found in the price-time series data which represents the macro dynamic of the underlying demand and supply of individual participants in the market. The features traditionally tested included the distribution of price returns, volatility clustering, absence of autocorrelation of price returns, and aggregation of price returns (Maslov, 2000; Challet and Stinchombe, 2001).

Additionally, a comparison was done between the simulation's output and empirical data of the S&P 500 E-mini futures market for the period under investigation. As Table 1 shows, the model appears to have the same characteristics in trade volume and cancelation rates.

Table 1: Real versus Simulated E-Mini: Trade Volume and Cancellation Rate

Trader Type	Simulated Volume	Actual Volume	Simulated Cancelation Rate	Actual Cancelation Rate
Small	1%	1%	40%	20-40%
Fundamental Buyers	10%	9%	44%	20-40%
Fundamental Sellers	10%	9%	44%	20-40%
Market Makers	10%	10%	35%	20-40%
Opportunistic	31%	33%	50%	40-60%
High Frequency	38%	38%	77%	70-80%

3.4 Model Output

Using the simulation environment, we were able to produce limit order book data for various market conditions which allowed us to observe stable and destabilized markets pricing conditions. The destabilizing condition in our simulation was introduction of an algorithmic trader, similar to the one that initiated the May 6th Flash Crash, which was the catalyst to other mini flash crashes. Although all destabilizing events do not occur exactly this way, our method produced conditions in a controlled manner for producing experimental data for testing.

[4] The code for the ABM will be made available upon request by emailing Mark.Paddrik@treasury.gov

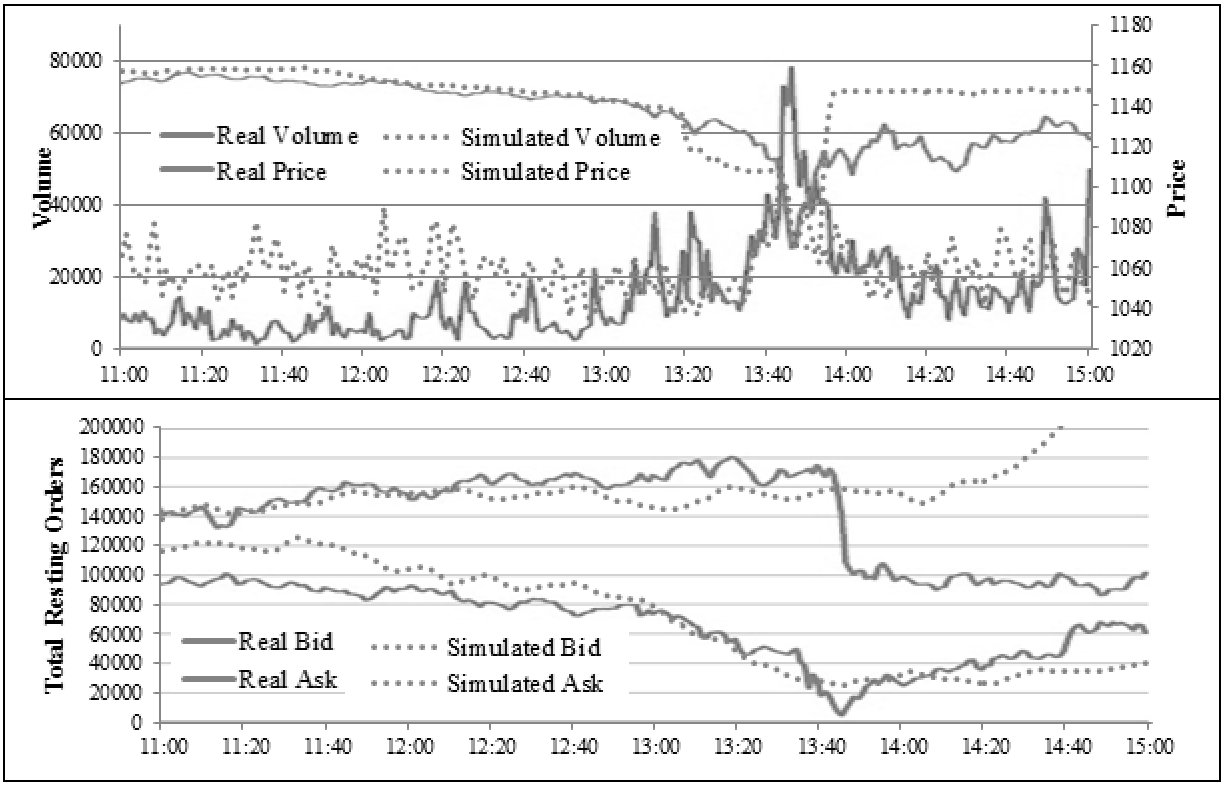


Figure 6: Real versus Simulated E-Mini: Volume, Moving Average Price, and Order book Depth

Figure 6 illustrates the results of the simulated price and the moving average volume of a particular instance of the simulation. The graph of the actual E-mini S&P 500 flash crash is shown for comparison. Note the widening separation in the order books bid-ask resting order depth quantities and then the sudden collapse.

4. Measures of Limit Order Book Market Stability

Traditional stability, or the resiliency of a market, is represented by the term liquidity, which refers to the ability to transform one asset into another in a short period without losing value. The ease of this transformation, in terms of time required and price impact is seen as a measure of a market's state or health (Benic, 2008). Unfortunately, liquidity is a multidimensional phenomenon, making it difficult to encompass in a single measure. Harris (2003) defines liquidity using four dimensions:

1. **Trading time.** Defined as the ability to execute the transaction immediately at the current price. The waiting time between trades is the measure for trading time.
2. **Tightness.** The ability to buy and sell an asset at about the same price at the same time, generally referenced as the instantaneous spread between the best bid and best offer.
3. **Depth.** The ability to buy or sell a certain amount of an asset without influencing the quoted price. A sign of illiquidity would be an adverse market effect on price when trading occurs.

4. **Resiliency**. The ability to buy or sell a certain amount of an asset with little influence on the quoted price. Market depth regards only the volume of best bid and best ask prices, but resiliency takes the elasticity of supply and demand into account.

Resiliency, unlike the other three, has no metric associated with its definition, which is often seen as an opportunity to observe several features of the electronic order book system when trying to characterize resiliency of the market's state.

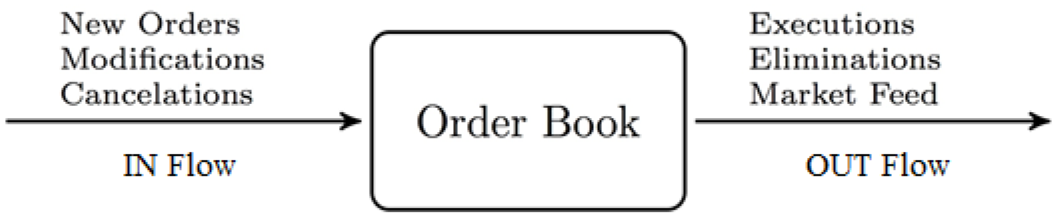


Figure 7: Market State Diagram

Figure 7 depicts the state of a market's order book with respect to order flow —in and out and current state of orders are depicted. The "IN" flows are characterized by new orders and orders modified to increase quantity, which increase the inventory of resting orders held in the order book. Order "OUT" flows are orders leaving the order book that have been executed or canceled and orders modified to decrease quantity, which also decreases the inventory of resting orders held in the order book. The order book's inventory of resting orders indicates the current demand and supply of participant interest for that market.[5]

To date, most studies done examining market resiliency have been focused on the impact order out flows have on price (Brandt, 2004; Love, 2008; Evans, 2002; Easley, 2010), but little has been done examining the effect of order in-flows or the state of resting orders in the order book. In this examination, we looked at five different indicators that measure price, order flow, and order book state to different degrees using the different levels of market data to search for which metric might reliably signal the risk of destabilizing events.

VPIN

Volume-Synchronized Probability of Informed Trading (VPIN) is an existing indicator developed by Easley, López de Prado and O'Hara (2010). VPIN is the ratio of average unbalanced volume to total volume in a period and requires Level 2 data. Heuristically, the VPIN metric "measures the fraction of volume-weighted trade that arises from informed traders as the informed tend to trade on one-side of the market and so their activity leads to unbalanced volume." (Easley et al., 2010) In other words, there might be an increasing toxicity of orders entering the market. This metric requires

[5] It is important to note that some of these metrics contain parameters that require tuning. In our work, we will select similar time scales for all the parameters so they produce new measurements every 30 seconds on average. We do not spend time calibrating these parameters to better fit the data, which has been an area of concern in previous research (Andersen and Bondarenko, 2014). .

being parameterized for the volume bucket size (V) and the number of buckets needed to get an accurate measure (n). In our simulation, $V = 50$, and $n = 50$[6].

$$VPIN = \frac{\sum_{i=1}^{n} \left| V_i^S - V_i^B \right|}{nV}$$

V: *the volume in every bucket*
V_i^B: *volume from buyer-initiated trades*
V_i^S: *volume from seller-initiated trades*
n: *the number of buckets used to approximate the expected trade imbalance*

Price Impact

Price Impact in limit order book markets is an existing indicator based on a measure of the effect a trade has on an asset's value. For this study, we selected the adverse selection component found by decomposing the price spread along the lines of Glosten (1987) and used by Hendershott et al. (2011) in examining the price impact of algorithmic trading on financial markets. Empirical evidence shows order book liquidity and adverse selection are inversely related (Frey, 2006).

The Price Impact indicator is the cumulative response of the quoted price to a one-time unit shock in the order flow as a measure of adverse selection, which accounts for persistence in order flow that we might translate as indication of sharp order flow imbalance (Hasbrouck, 1991). This metric requires using aggression (the side of the order book that initiated the trade) as found in Level 2 data. It is then parameterized for a time in the future $(t + x)$, and the current trade price (at time, *t*) is compared. (Note: In this study, x was 30 seconds.)

$$Price\ Impact = \frac{Q_t(M_{t+x} - M_t)}{m_t}$$

Q_t: *an indicator variable (+1: for buyer-initiated trades, -1: for seller-initiated trades)*
M_t: *midpoint price is the average of the best bid / ask at the time of the trade (t)*
x: *time in the future*

Window Spread Indicator

The window spread, a new indicator we developed as part of this work, is a measure of the price movement over a period that demonstrates the impact the current price is facing due to market order flow without regard to quantity traded, so it only requires Level 1 data. This metric is parameterized to the length of time that the window carries minimum and maximum values. In the simulation, we set this to 30 seconds.

$$Window\ Spread = \max_{t=0,\ldots,i}(P_t) - \min_{t=0,\ldots,i}(P_t)$$

P_t: *prevailing midpoint price at the time of the trade (t)*

---

[6] The parameters were selected based on the average volume bucket, V, would be equal to 30 seconds during a normal period so the model would produce new values every 30 seconds. The number of buckets was selected during based on the (Easley et al., 2010) value.

OBWA Spread

The Order Book Weighted Average Spread (OBWA) is new indicator we developed as part of this work. OBWA is a weighted average bid price and ask price spread of the order book using the first 10 best bid ticks and best ask ticks. The spread gives an aggregated measure of where the majority of resting orders are in the order book. The smaller the value indicates, the closer in agreement exists on price between the majority of participants. OBWA requires level 4 data.

$$OBWA\ Spread = \frac{\sum_{i=1}^{n} P_i S_i}{\sum_{i=1}^{n} S_i} - \frac{\sum_{j=1}^{m} P_j S_j}{\sum_{j=1}^{m} S_j}$$

S: *order size*
P: *trade price*
i: *ask order* , **j**: *bid order*
n: *number of bid limit orders,* $\boldsymbol{m}$: *number of ask limit orders*

Manual Spread

The Manual Spread (MS) is another new indicator we developed as part of this work. The indicator is a weighted average bid price and ask price spread of the order book using the first 10 best bid ticks and best ask ticks prices. Orders made by automated trading systems are not included in the average to remove any false indications of market resiliency that might be introduced because automated trading orders may not provide a true indication of price spread compared to manual orders. Without automated trading orders, we expect a better indication of the strength of the market to take “acidic” order flow (orders that consume the depth of the book). The caveat to this indicator is traders must be categorized traders and their orders classified, requiring Level 5 data. We identified automated trading accounts by the quantity of trading they did solely — 2,000 or more actions (new, canceled, or modified orders) in a single market — which is a quantity we determined beyond a human’s ability to manage.

$$Manual\ Spread = \frac{\sum_{i=1}^{n} P_i S_i}{\sum_{i=1}^{n} S_i} - \frac{\sum_{j=1}^{m} P_j S_j}{\sum_{j=1}^{m} S_j}$$

S: *order size*
P: *trade price*
i: *ask order* , **j**: *bid order*
n: *number of bid limit orders,* $\boldsymbol{m}$: *number of ask limit orders*

5. Measure Performance in the Agent-Based Model Environment

To test how well our indicators would respond to market conditions that arise from random and complex market transactions, as well as the possible effect of automated trading, we ran the simulation 2,000 times on a single i5 processor, taking 23 hours to run. Each run held the initial conditions with respect number of participants, their behavior, and market rules constant, varying only the selling behavior of the large selling algorithm used on May 6, 2010.

$$Large\ Sell\ Algorithm: C_t = V_{t-1} * PTV$$

$\mathbf{C_t}$: *number of contracts to sell at time t*
$\mathbf{V_{t-1}}$: *volume traded in previous minute (t-1)*
PTV: *percent of trade volume algorithm desires to be of entire market's volume*

The algorithm determines how many contracts it wants to sell, C_t, in the following minute based on a preset percentage, PTV, of the previous minutes volume traded, V_{t-1}. PTV is varied from 0 percent to 9 percent, which resulted in a rapidly increasing probability of a mini flash crash above about 4 percent, as seen in Figure 8.

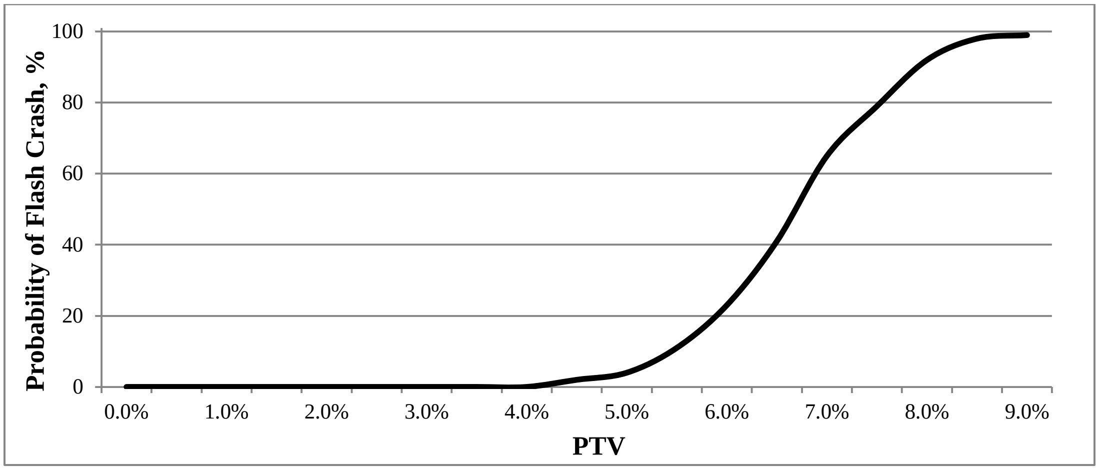


Figure 8: Probability Density Function Graph of a Flash Crash in Simulated E-Mini Market

The indicator results we observed in each of these simulations runs yielded the following set of distributions for the indicators under the two conditions of a normal price decline (circuit breaker triggering market pause) and flash crash resulting runs as seen in the probability density function (PDF) and cumulative density function (CDF) curves in Figure 9. The results showed the center of all indicator distributions, most notably the OBWA distribution, shifted to the right, demonstrating that, on average, all the indicators read higher values before the flash crash run versus the normal run.

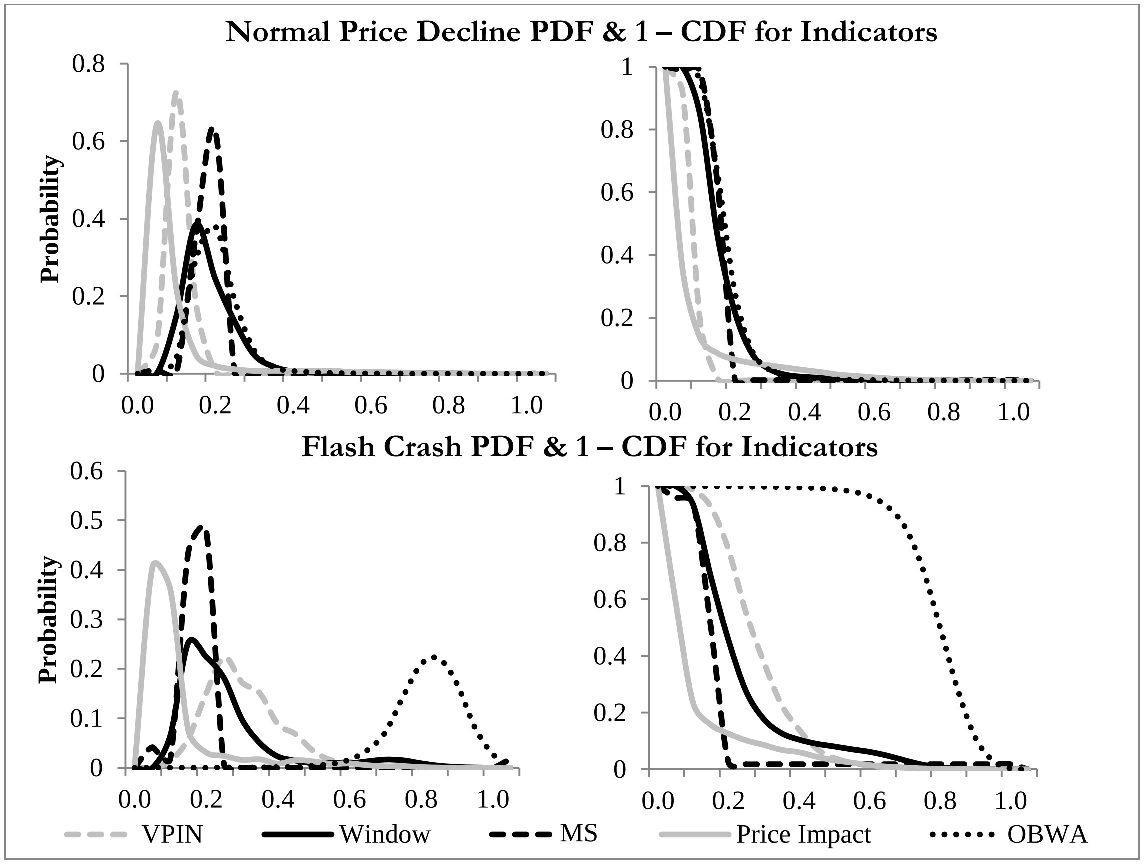


Figure 9: Indicator PDF and 1 – CDF in Simulated E-Mini Market

When we further examined the flash crash runs, we were interested in how much advanced warning we might expect given the readings of the indicators before the market pause. In Figure 10, the PDF and CDF of the maximum value observed before the market pause showed the highest values were within 30 seconds of the event for all indicators. The Window Spread and Price Impact indicators seemed to be slightly better as early warning predictors.

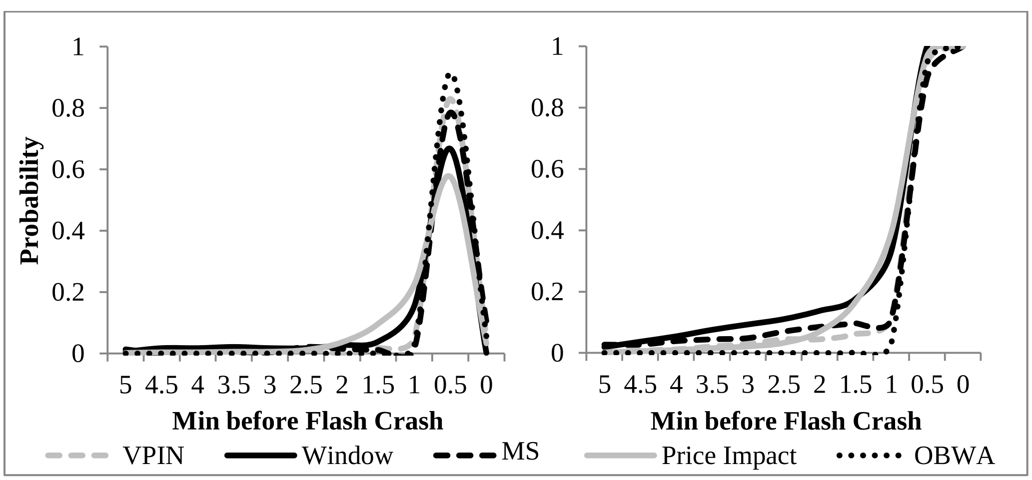


Figure 10: Indicator Advance Warning PDF and CDF

Finally, we compared the reliability of the predictive capabilities of the five indicators using receiver operating characteristic (ROC) curves. Figure 11 illustrates the performance of the indicators in relation to the true-positive state (a flash crash was predicted and occurred) and false-positive state (a flash crash was predicted but did not occur) as discrimination thresholds were varied. Measures of the maximum value were observed for the indicators at 30-second intervals from 5 minutes to 30 seconds before the event of a market pause. Figure 11 shows the strongest predictive power at 30 seconds, because it had the most predictive power, as seen in figure 9.

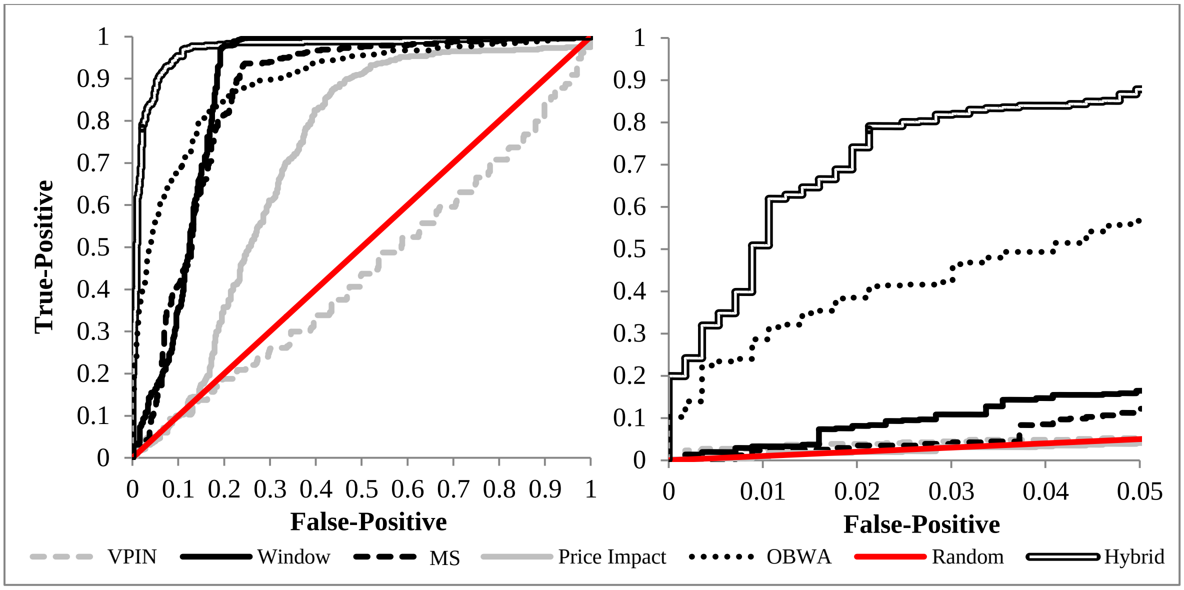


Figure 11: ROC Curves of Indicators: 30-Second Interval (Left) and Scale Modified of 30 –Second Interval (Right)

As the results show, no indicator of a market's resiliency and susceptibility to a mini flash crash is simple or perfect. The indicators focusing on measuring the state of the order book were able to predict the mini flash crash with the highest degree of accuracy. However, none of these individual indicators could encompass the dynamics of order flow and the state of the order book.

Combining the best of each of the measures into a *hybrid* indictor provided the strongest predictive value and an earlier warning signal. Using logistic regression, we found the predictive power of the weighting the Window (62 percent), Manual (19 percent) and OBWA (19 percent) Spread indicators seem to be significant in combination, resulting in a more trustworthy early-warning indicator that functions both when a market is near equilibrium and when high quantities of marketable orders are being placed.

It is important to note that the hybrid indicator gives a high true-positive-to-false-positive rate, as seen in the expanded portion of the ROC curve in Figure 11. This rate is an important consideration when determining if an acceptable indicator exists for implementing a new market policy in advance of a potential price destabilization, such as stopping a live trading system. The high true-positive-to-false-positive rate of the hybrid indicator is a good sign that an indicator market participants and exchanges could trust may exist for financial markets.

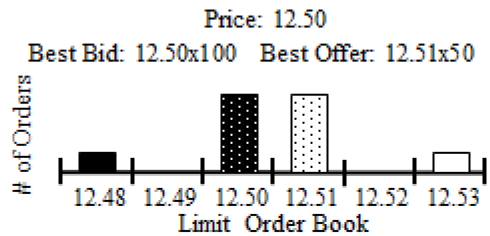


Figure 12: Book during a Lower Resiliency Event. AT = dotted, Manual = solid

A second interesting result of our simulation is the difference in predictive power observed in the OBWA and Manual spread indicators. The predictive power of the OBWA indicator is typically stronger showing that automated trading behavior is not overly skewing market liquidity during the most extreme points, assuming that behavior is constant. Yet, during tipping points, resting orders placed by automated trader can hide the real spread in demand and supply otherwise evident from the manually placed bid-ask resting orders, which would be demonstrated by a growing gap of ticks between the best bid-ask. Figure 12 demonstrates this point, showing an example of an order book where automated trading orders make the spread appear tighter than the best manual placed orders could support. Consequently, demand and supply appear to be closely matched, which might serve as a catalyst for needless trading to occur.

6. Vetting Measure Performance Using Real Data

As a form of validation of the results on simulated data; we used two real-world datasets to examine the performance of indicators. These tests required obtaining Level 6 datasets so we could accurately reconstruct the order book, tracking each limit order as it was entered, modified, and removed from the limit order book, as well as the execution of other order types to accurately reconstruct events. In addition, we gave each order an identifier that classified it as belonging to an automated trader or manual. The final difficulty was to find events that occurred in single order book markets and which, more importantly, experienced a mini flash crash like event.

We used two unique data sets to investigate the predictive power of the indicators:

i. September 19, 2012. Drop in the front month WTI Crude Oil Futures market when the price of dropped $4 over four minutes (much of which occurred during 30 seconds of that period) and reached its low at 12:55PM CST.
ii. The Flash Crash of May 6, 2010. The crash originated in the front month E-Mini S&P 500 Futures market [CFTC & SEC. 2010] where the price dropped 3 percent over four minutes. At 1:45PM CST, it reached its low when the Chicago Mercantile Exchange circuit breaker was triggered before the price was able to recover in the contract.

Using four days of data, three full days selected between 2011 and 2012 and the full day from the day of the price events previously mentioned, we examined the indicators' behavior (see Appendix A for market statistic to compare to simulation). By comparing values on normal days to those seen during price-event days, we were able to construct a "likelihood ratio" for each indicator to represent how it changed from a market acting normally without a flash event to one experiencing a flash event.

$$Flash\ Crash\ Likelihood\ Ratio = \frac{1 - CDF(F)}{1 - CDF(N)}$$

N: *normal day data*
F: *flash day* data

The results for real market data are shown in Figure 13 and were similar to those in the simulated data sets. We observed the Window Spread indictor as the most confident indicator of the group before the lowest price observed in the set. However, the MS and OBWA indicators were able to show earlier warning signs of the markets becoming abnormal. Using the three in combination as a hybrid indicator, we were able to get the best of the both with observing a ratio above 100 signaled the likelihood of a flash crash more than a minute in advance of the worst phase of the two price events.

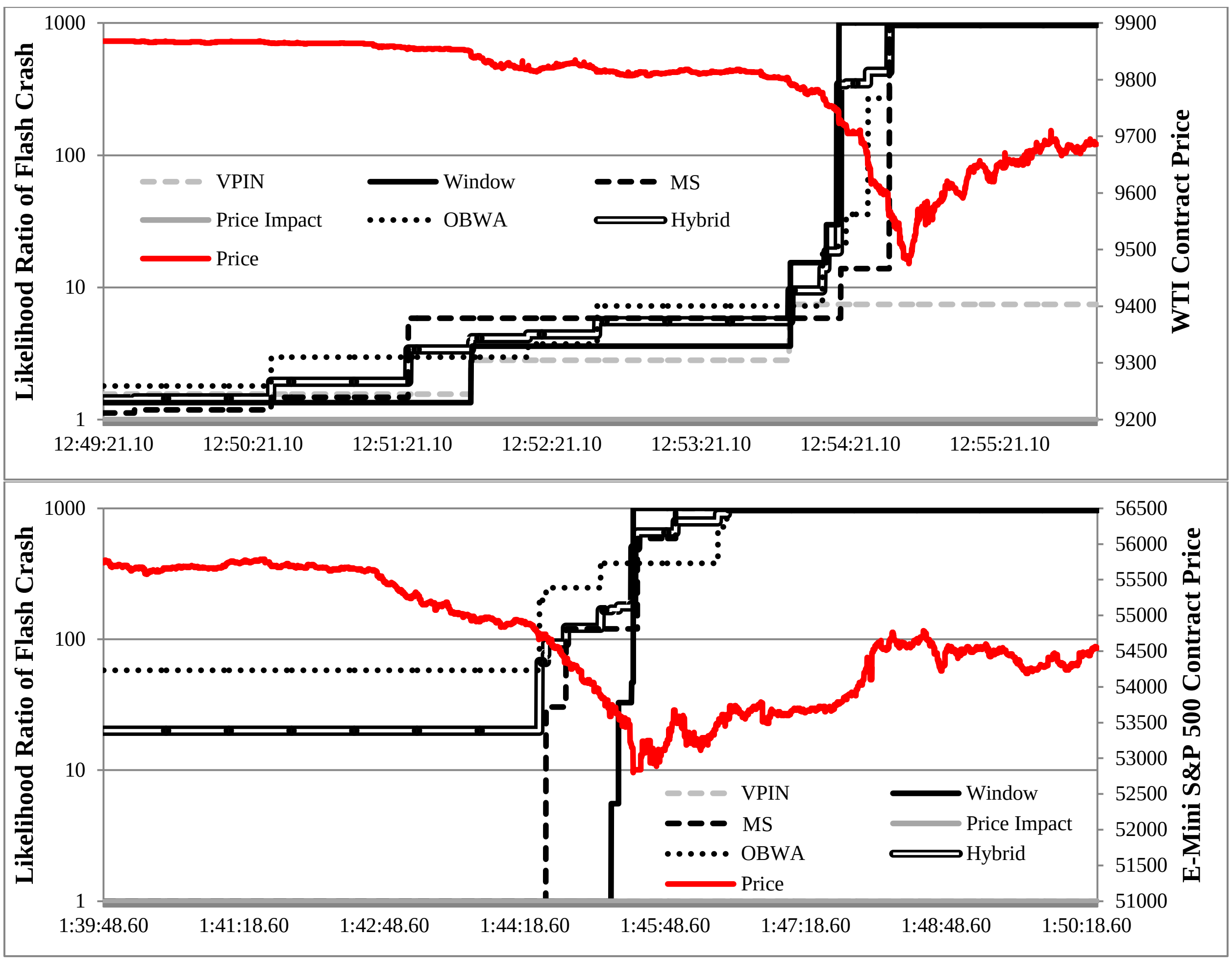


Figure 13: WTI Crude Oil (top) & E-Mini S&P 500 (bottom) Futures Price Events

7. Conclusions

Although access to the full scope actual market data is limited, data describing almost everything that happens in markets does exist and could be made available, allowing for a complete picture of the markets. The challenge is to determine how to best make use of this overwhelming amount data to keep order in the markets. This paper discusses the available data, how it relates to the price discovery process, and the challenges and possible solutions of using it to understand market stability.

Effective analysis of the electronic order book's microstructural data has been limited by both the data's accessibility, because of size and complicated structure, and the lack of a controlled environment to examine it. Through the development and application of an agent-based model faithfully replicating the limit order book structure, we were able to produce normally sensitive and private data in a controlled environment.

We then examined the different levels of market data produced from the model and relationships to the price discovery process. We have in this paper identified a possible set of data driven indicators that could provide a level of predictive capability to potentially avoid destabilizing pricing events. Applying these indicators developed comparing market simulation data to actual market data, we found advance warning signals could be detected a full one minute before two well-documented market flash crash events occurred.

The results of our work demonstrates that with available data, underlying market microstructure state of the order book can be a more accurate predictive indicator of market's resiliency than order flow, even though order flow is a significant component. These findings not only help to demonstrate the important role higher level data can play in regulators' assessment of financial markets but also how agent-based models can be used as a testing platform for our understanding of markets and financial stability.

References

Anderson, P., Arrow, K., & Pines, D.(1988). The Economy as an Evolving Complex System. Redwood City, Addison-Wesley Co.

Andersen, T., Bollerslev, T., Diebold, F. & Vega, C. (2007). Real-time Price Discovery in Global Stock, Bond and Foreign Exchange Markets. Journal of International Economics 73, 2, 251-277.

Andersen, T. & Bondarenko, O. (2014). VPIN and the flash crash. Journal of Financial Markets, 17, 1-46.

Bethel, E., Leinweber, D., Rübel, O., & Wu, K. (2011). Federal market information technology in the post flash crash era: roles for supercomputing. In Proceedings of the fourth workshop on High performance computational finance, 23-30.

Benic, V. & Franic. I. (2008). Stock Market Liquidity: Comparative Analysis of Croatian and Regional Markets. Financial Theory and Practice 32, 4, 477-498.

Berger, D., Chaboud, A., Chernenko, S., Howorka, E., & Wright, J. (2008). Order flow and exchange rate dynamics in electronic brokerage system data. Journal of International Economics 75, 93-109.

Brandt, M. & Kavajecz, K. (2004). Price discovery in the U.S. Treasury market: The impact of order flow and liquidity on the yield curve. Journal of Finance 59, 2623-2654.

CFTC & SEC. (2010). Findings regarding the market events of May 6, 2010. September 30, 2010.

Challet, D. & Stinchombe, R. (2001). Analyzing and modeling 1 + 1d. Physica A: Statistical Mechanics and its Applications 300, 285-599.

Clark-Joseph, A. D. (2013). Exploratory trading. Unpublished job market paper. Harvard University, Cambridge, MA

CME Group. (2013). CME MDP Message Statistics. Web. http://beta.cmegroup.com/market-data/distributor/market-data-platform.html

Cont, R., Kukanov, A., & Stoikov, S. (2013). The price impact of order book events. Journal of Financial Econometrics 12, 1, 47-88.

Duong, H. N. & Kalev, P. S. (2014). Anonymity and the Information Content of the Limit Order Book. Journal of International Financial Markets, Institutions and Money.

Easley, D., López de Prado, M., O'Hara, M. (2010). The microstructure of the 'flash crash': Flow toxicity, liquidity crashes and the probability of informed trading. Working Paper.

Easley, D., López de Prado, M., & O'Hara, M. (2011). The microstructure of the flash crash: Flow toxicity, liquidity crashes and the probability of informed trading. Journal of Portfolio Management 37, 2, 118-128.

Easley, D., López de Prado, M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. Review of Financial Studies 25, 1457-1493.

Evans, M. & Lyons, R. (2002). Order flow and exchange rate dynamics. Journal of Political Economy 110, 170-180.

Farmer, J. D., Patelli, P., & Zovko, I. (2005). The predictive power of zero intelligence in financial markets. *Proceedings of the National Academy of Sciences of the United States of America* 102, 6, 2254-2259.

Farmer, J. D. & Foley, D. (2009). The economy needs agent-based modelling. Nature 460, 7256, 685-686.

Friedman, D. & Rust, J. (1993). The Double Auction Market: Institutions, Theories, and Evidence. Redwood City, Addison-Wesley Co.

Frey, S., Grammig, J. (2006). Liquidity supply and adverse selection in a pure limit order book market. Empirical Economics 30, 1007–1033.

Gode, D. K. & Sunder, S. (1993). Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. Journal of Political Economy 101, 1, 119-137.

Glosten, L. (1987). Components of the bid ask spread and the statistical properties of transaction prices. Journal of Finance ,42, 1293–1307.

Glosten, L. (1994). Is the electronic open limit order book inevitable?. Journal of Finance, 49, 1127-1162.

Golub, A., Keane, J., & Poon, S. H. (2012). High Frequency Trading and Mini Flash Crashes. *arXiv preprint arXiv:1211.6667*.

Harris, L. (2003). Trading and Exchanges: Market Microstructure for Practitioners. Oxford University Press. New York.

Hasbrouck, J. (1991). Measuring the information content of stock trades. Journal of Finance, 46, 179-207.

Hasbrouck J. (1995). One security, many markets determining the contributions to price discovery. Journal of Finance, 50, 1175–1199

Hasbrouck, J., & Saar, G. (2013). Low-latency trading. Journal of Financial Markets, 16, 4, 646-679.

Hayes, R., Todd, A., Chaidarun, N., Tepsuporn, S., Beling, P. & Scherer, W. (2014). An Agent-Based Financial Simulation for Use by Researchers. *Available at SSRN*

Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity?. The Journal of Finance, 66, 1-33.

Johnson, N., Zhao, G., Hunsader, E., Meng, J., Ravindar, A., Carran, S., & Tivnan, B. (2012). Financial black swans driven by ultrafast machine ecology. *Available at SSRN*.

Kaldor, N. (1961). Economic growth and capital accumulation. The Theory of Capital, Macmillan, London.

Kim, Y. & Yang, J. (2004). What makes circuit breakers attractive to financial markets? A survey. Financial Markets, Institutions & Instruments 13, 109–146.

Kirilenko, A., Kyle, A., Samadi, M. & Tuzun, T. (2011). The Flash Crash: The Impact of High Frequency Trading on an Electronic Market. Available at SSRN: http://ssrn.com/abstract=1686004

Krantz, M. (2011). Mini Flash Crashes Worry Traders. USA Today. 16 May 2011. Web. <http://www.usatoday.com/money/markets/2011-05-16-mini-flash-crashes-market-worry_n.htm>.

Love, R., Payne, R. (2008). Macroeconomic news, order flows, and exchange rates. Journal of Financial and Quantitative Analysis 43, 467-488.

Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modelling and simulation. Journal of Simulation 4, 3, 151-162.

Maslov, S. (2000). Simple model of a limit order-driven market. Physica A: Statistical Mechanics and its Applications, 278, 571-578.

Menkveld. A. J. (2010). High frequency trading and the new market maker. Working Paper.

Nanex. (2011). Flash Crash Analysis Continuing Developments: Flash Equity Failure for 2006, 2007, 2008, 2009, 2010, and 2011. http://www.nanex.net/FlashCrashEquities/FlashCrashAnalysis_Equities.html

O'Hara, M. (2014). High Frequency Market Microstructure. Working Paper. Cornell University.

Oliven, K. & Rietz, T. (2002). Suckers are born but markets are made: Individual rationality, arbitrage and market efficiency on an electronic futures market. Management Science.

Paddrik, M., Hayes, R., Todd, A., Yang, S., Scherer, W. & Beling, P. (2012). An Agent-based Model of the E-Mini S&P 500 and the Flash Crash. IEEE Computational Intelligence for Financial Engineering and Economics.

Preis, T., Golke, S., Paul, W. and Schneider, J. J. (2006). Multi-Agent Based Order Book Model Of Financial Markets, Europhysics Letters, 75 (3), pp. 510516.

SEC. (2010). Concept release on equity market structure. Release No. 34-61358; File No. S7-02-10.

SEC. (2012). Investor Bulletin: New Measures to Address Market Volatility. Release 23 July 2012. Web <http://www.sec.gov/investor/alerts/circuitbreakersbulletin.htm>

Subrahmanyam, A. (1994). Circuit Breakers and Market Volatility: A Theoretical Perspective. The Journal of Finance. 49: 237–254.

Wah, E. & Wellman, M. P. (2013). Latency arbitrage, market fragmentation, and efficiency: a two-market model. In Proceedings of the fourteenth ACM conference on Electronic commerce, 855-872.

Zhang, F. (2010). High-Frequency Trading, Stock Volatility, and Price Discovery. Working Paper.

Appendix A

Trader Type	#of Traders	Rate of New Orders	Position Limits	Market Volume
Small	6880	7200 sec	-∞ – ∞	1%
Fundamental Buyers	1268	60.0 sec	0 – ∞	9%
Fundamental Sellers	1276	60.0 sec	-∞ –0	9%

Market Makers	176	20.0 sec	-120 – 120	10%
Opportunistic	5808	120 sec	-120 – 120	33%
High Frequency	16	0.35 sec	-3000 – 3000	38%
Manual	15232	3308 sec	$-\infty$ – ∞	52%
AT	192	18.1 sec	-360 – 360	48%

Table 1: Simulated S&P 500 E-Mini Market Participation