Cross-Section of Mini Flash Crashes and Their Detection by a State-Space Approach

Chyng Wen Tee*† Christopher Ting*

Abstract

We define mini flash crashes as significant deviation from a normal price process, and formulate a robust and simple state-space model for efficient detection of mini flash crashes. Compared to existing methodologies, we demonstrate that our method can capture a much broader range of mini flash crashes. As one would expect, mini flash crashes spike when there are major information events in the market. Our research shows that stocks with larger market capitalization, smaller volatility, higher averaged price, and larger liquidity are consistently more robust against these extreme price movements. These market microstructure behaviors are consistent with the literature studying the impact of public information arrivals on price discovery. This suggests that mini flash crashes occur because market assimilates new information at an increasingly faster speed, complementing the view that high-speed trading increases information efficiency.

Keywords: market microstructure, mini flash crashes, state-space model, liquidity, high speed trading, information efficiency.

*Lee Kong Chian School of Business, Singapore Management University, 50 Stamford Road #05-01, Singapore 178899.

†Corresponding author, email: cwtee@smu.edu.sg. Tel: +65 6828 0819

1 Introduction

It is well known that irrespective of methods or technological advances, executions of sizeable order would always generate large price impact and slippage (Black (1971)). Price continuity is more likely maintained when trading is conducted in blocks of small, individual trades. However, trading stocks across fragmented markets is risky as mini flash crashes occur more frequently than previously thought. Mini flash crashes are defined as events in which intraday asset prices experience an abrupt drop or surge within a short time (a few seconds), and may or may not be followed by a flash recovery offsetting much of the losses or gains. Among all flash crashes occurred to date, the May 6^{th} 2010 Flash Crash is the most notable and rigorously studied (see Madhavan (2012) and Kirilenko et al. (2017)). But what constitutes a mini flash crash is highly dependent on both the definition and the selection criteria used to detect them. In our survey, we have identified five major definitions of mini flash crashes widely used by academic and practitioner alike. The first commonly used rule-of-thumb definition for mini flash crashes is as follows:

- at least 10 ticks in the same direction;
- time window does not exceed 1.5 seconds;
- price change exceeds 0.8%.

This definition is proposed by Nanex, a trading technology company and data service provider. Based on this definition, Nanex have identified a large number of crashes after scanning their trade database (see Nanex (2010)). Researchers following this definition of mini flash crashes in subsequent studies include Johnson et al. (2013) and more recently Ozenbas and Schwartz (2018).

A second frequently used definition of mini flash crashes is to count the number of trades outside the previous second's bid-ask range — in other words, transactions executed at trade price above the ask price or below the bid price. Intuitively, persistent transaction outside the quoted spread is a measure of desperate buyers or sellers rushing to trade, which can be interpreted as a mini flash crash. Researchers using this definition include Agrawal et al. (2018), who investigate the impact of social media on high-speed price momentum and mean-reversion. The list of mini flash crashes detected under this definition can be obtained from Wharton Research Database (WRDS).

A third definition for mini flash crashes is to use a market microstructure model to estimate return and label intervals with residuals that belong to the top percentile (e.g. 99.9^{th}) as extreme price movements (EPMs). Brogaard et al. (2018) use this method to study the behavior of high-frequency traders during extreme events. A fourth definition of mini flash crashes is provided by Lee and Mykland (2012), who use nonparametric empirical methods to detect jumps in the underlying price process in the presence of general dependent noise processes, taking local volatility into account. Finally, a fifth definition is given by Dugast and Foucault (2014), who define mini flash crashes as large sudden price drops or spikes followed by quick price reversals, i.e. "V-shape" or "inverted V-shape" price movements. In this definition, price movements that do not revert will not be classified as mini flash crashes.

Despite the recent progress, we have identified two areas where further research is necessary. First, while the numerous existing definitions of mini flash crashes are intuitive and applicable, albeit with varying degree of robustness, it is not ideally suited to real-time mini crash detection, given the need to closely track the number of uni-directional tick, the magnitude of the price change, the width of the time window, or to perform in-sample model calibration or regression. In addition, as we will demonstrate in this paper, the choice of the cutoff in these three crash parameters might potentially result in otherwise sizeable flash crashes to be missed. Consequently, an alternative method that is more robust, efficient, and scalable is desirable for real-time application. In this work, we formulate a robust state-space model to accurately capture abnormal price movement during flash crashes. This new method circumvents the shortcomings of existing methodologies, and is able to detect a much broader range of mini flash crashes.

Second, a fundamental understanding on the impact of these mini flash crashes, in terms of their descriptive statistics, as well as a cross-sectional breakdown of their distribution, remain inadequate. Do stocks with different market capitalization, volatility, liquidity, and price level behave differently during mini flash crashes? What impact does mini flash crashes have on market microstructure? Before dismissing them as disruptions to the market, it is important to first understand their influence to intraday market structure. The analysis should cover a broad class of stocks and include a sufficiently long time window for the results to be representative. In this work, we report in-depth statistical analyses conducted based on the mini flash crashes detected using the state-space model approach formulated in this paper.

This paper is organized as follows. We review existing literature in this area in Section 2.

In Section 3 we formulate a novel way to identify mini flash crashes, which makes use of a state-space model on the trade data based on a Kalman filter approach. This is followed by a description on the dataset used in this work, and in-depth statistical analyses are presented in Section 4, where we present a full breakdown of crash statistics across market capitalization, volatility, price level, and liquidity. Regression analysis is also performed to identify relationship between crash statistics with these parameters. Finally, conclusions of this study are drawn in Section 5.

2 Literature Review

Mini flash crashes and their implications on the financial market are increasingly the subject of intense research effort. A thorough understanding on the mechanism, along with the full implication of mini flash crashes, is vital for market regulation and the implementation of effective preemptive measures. As Kirilenko and Lo (2013) point out, modernization of regulatory oversight is crucial in the quest to bring relevant financial regulation into the digital age. Insights on the mechanisms of flash crashes will become the basis for formulating effective and prudent regulatory framework.

Johnson et al. (2013) suggest that mini flash crashes may be a result of interaction among several trading algorithms, or a positive feedback loop induced by market conditions. On the other hand, Golub and Poon (2012) suggest the aggressive use of the Intermarket Sweep Order (ISO), an exemption to the Order Protection Rule in the National Market Regulation (Reg NMS), by high frequency traders as a possible cause. This is because Reg NMS only protects the top of the order book, so that trade execution on exchanges with limited liquidity might result in trades that are far away from the national best bid-offer. More recently, Ozenbas and Schwartz (2018) and Agrawal et al. (2018) have also studied the impact of social media and other exogenous source of mini flash crashes.

High-speed algorithmic trading (AT) and market fragmentation are frequently considered as the implicit cause of flash crashes. As Schapiro (2010) points out, trading activities in U.S. equity market is split among over 13 public exchanges, more than 30 dark pools and over 200 internalizing broker-dealers, making it difficult for all participants of the capital markets to access liquidity in an organized manner. However, O'Hara and Ye (2011) discovered that market fragmentation generally reduces transactions costs and increases execution speeds. Fragmentation does increase short-term volatility, but prices are more efficient in that they are closer to being a random walk. Eventually, as trading activities becomes faster, in tandem with the fundamental changes in the market structure, market participants must adapt their trading strategies or risk falling victim to the predatory algorithms of high speed traders looking to act on information revealed by them (see O'Hara (2014)).

An important follow-up question posed by researchers is whether AT improve market liquidity. In the attempt to answer this question, Hendershott, Jones and Menkveld (2011) analyze New York Stock Exchange (NYSE) automated quote dissemination. Their findings indicate that algorithmic trading improves liquidity and enhances the information content of market quotes. In a related work, Hasbrouck and Saar (2013) also find that increased low-latency activity improves traditional market quality measures such as short-term volatility, spreads, and displayed depth in the limit order book, suggesting that increased low-latency activity does not necessarily work to the detriment of long-term investors. Indeed, Colliard (2013) shows that high speed market participants decrease the likelihood of short-lived mispricings by trading against price pressure. Dugast and Foucault (2014) specifically investigate the interplay between informational efficiency and price reversals. Their studies led them to conclude that cheaper fast trading technologies simultaneously raise information efficiency and the frequency of mini flash crashes, thus explaining two apparently incompatible observations: that the perception by market participants that mini flash crashes are increasingly frequent, and the academic findings suggesting that high frequency traders contribute to price discovery. For instance, Brogaard et al. (2018) show that during extreme price movements, high frequency traders (HFTs) provide liquidity by absorbing imbalances created by non-HFTs, unless the extreme event is occurring to multiple stocks, in which case HFTs will stop providing liquidity to manage their risk.

Criticisms of high-speed traders do exist. For instance, Wah and Wellman (2013) argued that an infinitely fast arbitrageur profits from market fragmentation by reaping the surplus when the two markets diverge due to latency in cross-market communication. Sparrow (2012) also points out that adverse selection can occur when some market participants have faster access to market events. Nevertheless, Lee, Cheng and Koh (2011) allude to the fact that that while some pundits blame HFTs, their analyses are unable to identify a direct link to high-frequency trading *per se*, but rather the domination of market activities by trading strategies that are responding to the same set of market variables in similar ways, as well as various pre-existing market microstructural safety mechanisms with unintended consequences when triggered simultaneously.

Menkveld (2014) highlights that the proliferation of HFTs and the fragmentation in financial markets are interrelated, and the two aspects of the problem ought to be considered jointly. For example, consider the Flash Crash of May 6 2010. Analyses presented in Kirilenko et al. (2017) suggest that HFTs did not trigger the Flash Crash, but their responses to the unusually large selling pressure on that day did exacerbate market volatility. A similar conclusion is reached in an independent study by Menkveld and Zhou (2018), who argued that the crash cannot be attributed to a single agent but really is the product of agent interaction.

3 Model

In this section, we present a state-space model based on a Kalman filter approach to identify mini flash crashes. We define crashes as asset returns experiencing a significant deviation from a normal process. As we will demonstrate, this simple yet intuitive definition allows us to formulate a robust methodology to capture the majority of mini flash crashes. A salient feature of the Kalman filter approach is its ability to solve the state estimation problem within a Bayesian framework. We begin with the common practice in the literature to postulate that stock price follows a lognormal process

$$dS_t = \mu_t S_t dt + \sigma_t S_t dW_t,$$

where $W_t \sim N(0, t)$ is a standard Brownian motion following normal distribution with a mean of 0 and a variance of t, μ_t is the drift coefficient, and σ_t is the volatility coefficient. An application of Itô's formula yields the log-price process from timestamp t to the next at $t+\Delta t$. In general, the drift term (μ_t) is negligible over a short time horizon – this is generally an adequate assumption for high frequency data. It follows that

$$\log S_{t+\Delta t} \approx \log S_t + \sigma_t W_{\Delta t}$$

3.1 State-space model

Let $x_i = \log \tilde{S}_{t+\Delta t}$ denote the *true* log-price process which is unobservable, while $z_i = \log S_{t+\Delta t}$ denote the *traded* log-price, which is interpreted as a measurement of the true log-price process subject to noise. We propose the following state-space model to describe the log-price processes at the transaction level:

$$\begin{cases} x_i = x_{i-1} + \epsilon_{p,i-1}, & \epsilon_{p,i-1} \sim N(0, \sigma_{p,i}^2 \cdot \Delta t_{i-1,i}) \\ z_i = x_i + \epsilon_{m,i}, & \epsilon_{m,i} \sim N(0, \sigma_{m,i}^2) \end{cases}$$

where $x_i = \log S_i$ is the log-price on the i^{th} transaction, ϵ_p is the process noise, and is assumed to be normally distributed, while ϵ_m is the measurement noise, and is also assumed to be normally distributed and independent of ϵ_p. Here, σ_p^2 and σ_m^2 are the filter parameters that need to be estimated (see Harvey (1991) for further exposition on the Kalman filter approach). In our formulation, we have made the intuitive assumptions that the variance of the true price process is proportional to the time elapsed between each transaction, so that the variance of the process noise (ϵ_p) is proportional to $\Delta t_{i-1,i}$. We note in passing that the application of Kalman filter in finance is not new. For instance, Hasbrouck (1999) uses this concept to formulate a dynamic model of discrete bid and ask quotes.

We define the *a priori* (e_i^-) and *a posteriori* (e_i) errors, respectively, as

$$e_i^- = x_i - \hat{x}_i^-$$
$$e_i = x_i - \hat{x}_i$$

where $\hat{x}_i^-$ is estimated from $\hat{x}_{i-1}$ while $\hat{x}_i$ is estimated based on measurement z_i. Intuitively, the *a priori* error is quantified based on the difference between x_i and $\hat{x}_i^-$ (estimated before measurement z_i is obtained), while the *a posteriori* error is quantified based on $\hat{x}_i$, the estimation based on measurement z_i. The *a prior* estimate error is $\mathbb{E}[(e_i^-)^2]$, and the *a posteriori* estimate error is $\mathbb{E}[e_i^2]$. In Kalman filter formulation, we say that the state estimate $\hat{x}_i$ of the unknown true log-price at i given measurement z_i is a linear combination of the *a priori* estimate $\hat{x}_i^-$ and a weighted difference between an actual measurement z_i and a measurement prediction $\hat{x}_i^-$:

$$\hat{x}_i = \hat{x}_i^- + \kappa(z_i - \hat{x}_i^-),$$

where κ is the filter gain coefficient. The rationale is that if the distance between z_i and $\hat{x}_i^-$ is small, there is little surprise, hence $\hat{x}_i$ should be close to the *a priori* estimate $\hat{x}_i^-$. However, if this distance is large, then the true state estimate should be revised up or down based on the parameter κ. The objective will then be to choose κ in order to minimize the *a posteriori* error

e_i. To this end, we compute the *a posteriori* error variance

$$\sigma_{e,i}^2 = \mathbb{E}[e_i^2] = \mathbb{E}[(x_i - \hat{x}_i)^2] = (1-\kappa)^2(\sigma_{e,i}^-)^2 - \kappa^2\sigma_{m,i}^2.$$

Next, we take the partial derivative with respect to κ to obtain the minimum variance. The κ value that minimizes the *a posteriori* error is given by

$$\kappa = \frac{(\sigma_{e,i}^-)^2}{(\sigma_{e,i}^-)^2 + \sigma_{m,i}^2}.$$

Substituting back to the equation for the *a posteriori* error covariance, we obtain

$$\sigma_{e,i}^2 = (1-\kappa)(\sigma_{e,i}^-)^2.$$

Finally, we note that the *a priori* error variance can be expressed as

$$(\sigma_{e,i}^-)^2 = \mathbb{E}[(e_i^-)^2] = \mathbb{E}[(x_i - \hat{x}_i^-)^2] = \sigma_{e,i-1}^2 + \sigma_p^2.$$

In most Kalman filter application, the process noise (σ_p^2) and measurement noise (σ_m^2) are assumed to be constant (Harvey (1991)). However, for high frequency intraday financial data, there has been numerous research showing that volatility exhibits time-varying properties. For instance, a well-known study on intraday volatility is presented in Andersen and Bollerslev (1997). We note that the flexibility of the state space model makes it possible to allow such generalized autoregressive conditional heteroskedasticity (GARCH) features. Following Engle and Sokalska (2012), we divide each trading day (labeled by n) into 5-minute intervals (labeled by j) and calibrate an intraday volatility forecasting model based on multiplicative component GARCH. We use a GARCH model for high-frequency intraday financial returns, which specifies the conditional variance to be a multiplicative product of daily, diurnal, and stochastic intraday volatility. Intraday equity returns are described by the following process

$$r_{n,j} = \sqrt{h_n \cdot s_j \cdot q_{n,j}} \cdot e_{n,j}, \quad e_{n,j} \sim N(0,1).$$

where h_n is the daily variance component, s_j is the diurnal variance pattern, $q_{n,j}$ is the intraday variance component with $\mathbb{E}[q_{n,j}] = 1$, and $e_{n,j}$ is an error term. The intraday returns are assumed to be serially uncorrelated. For the daily variance component (h_n), we follow Andersen and Bollerslev (1997) by modeling and forecasting it using a daily GARCH model.

3.2 Computational Procedure & Crash Detection

In this section, we lay out the computational procedure for our state-space model for efficient mini flash crashes detection. We slice each day (labeled by n) into 5-minute interval (labeled by i). The total variance for each day is calculated as

$$\mathrm{RV} = \sum_{j=1}^{N} (x_j - x_{j-1})^2 = h_n.$$

A daily GARCH(1,1) model is used to forecast the daily variance (labeled by n):

$$h_n^2 = \omega + \alpha \varepsilon_n^2 + \beta h_{n-1}^2$$

For each 5-minute interval intraday (labeled by j), the intraday GARCH model is given by

$$\begin{aligned}
r_{n,j} &= x_{n,j} - x_{n,j-1}, \\
r_{n,j} &= \sqrt{h_n \cdot s_j \cdot q_{n,j}} \cdot e_{n,j}, \quad e_{n,j} \sim N(0,1) \\
q_{n,j} &= \omega + \alpha z_{n,j-1}^2 + \beta q_{n,j-1}, \quad z_{n,j} = \frac{r_{n,j}}{\sqrt{h_n \cdot s_j}} = \sqrt{q_{n,j}} \cdot e_{n,j}
\end{aligned}$$

The diurnal component, s_j, is estimated as

$$\hat{s}_j = \frac{1}{N} \sum_{n=1}^{N} \frac{(x_{n,j} - x_{n,j-1})^2}{h_n}.$$

For illustration purpose, Figure 1 plots the volatility information for Apple (AAPL) stock over the full year of 2010. The top subplot is the diurnal volatility component, while the bottom subplot is the composite intraday volatility. The spike in volatility corresponds to the May 6 Flash Crash. With the intraday volatility in place, we then set the process and measurement variances ($\sigma_{p,i}^2$ and $\sigma_{m,i}^2$) in the state-space model and proceed to estimate the states of the log-price in a standard Kalman filter recursive two-step procedure:

1. Time update: *a priori* prediction step

$$\hat{x}_i^- = \hat{x}_{i-1}^-$$
$$(\sigma_{e,i}^-)^2 = \sigma_{e,i-1}^2 + \sigma_{p,i}^2$$

2. Measurement update: *a posteriori* correction step

$$\kappa_i = \frac{(\sigma_{e,i}^-)^2}{(\sigma_{e,i}^-)^2 + \sigma_{m,i}^2}$$
$$\hat{x}_i = \hat{x}_i^- + \kappa_i(z_i - \hat{x}_i^-)$$
$$\sigma_{e,i}^2 = (1 - \kappa_i)(\sigma_{e,i}^-)^2$$

Under the Kalman filter formulation, time and measurement updates are performed on the log-price, giving us the latest *a posteriori* state and error estimates. Using these estimates, we can determine the distance between the measurement price (actual traded log-price) and the true price, which is unobservable and estimated by the state-space model. Our model formulation allows us to provide a robust and intuitive definition of mini flash crashes – whenever the measurement price deviates from the true price beyond a predefined threshold amount. We can calculate this distance in units of standard deviation based on the state error estimates, and explicitly specify this threshold as a z-score.

To summarize, there are three parts to the procedure:

1. Intraday and daily GARCH for volatility estimation
2. State-space model update
3. Identify transactions executed far away from true price as crashes

The mechanism behind mini flash crash detection using our state-space model is developed further in Section 4.

4 Analysis and Results

4.1 Dataset

The data used in this investigation is obtained from NYSE's Daily Trade and Quote (TAQ) database. The daily TAQ data set used here are definitive transcripts of the consolidated activity, where all trades are time-stamped to the millisecond. In addition to the timestamp and trade price, each trade event also contains information pertaining to the trade volume, trade venue and trade condition. In this work, our focus is on large market capitalization stocks – the selection criterion is therefore based on annual S&P 500 index constituent, including recent changes made over the period studied. Apart from TAQ, we also use Compustat Database to obtain information on market capitalization, volatility, and other stock variables. A total of four calendar years (2010-2013) is included in our analysis. This dataset is representative in that it includes several known cases of major flash crashes (refer to Section 4.2 for more information). Table I summarizes the number of trades and volume transacted each year. Our investigation covers over 11 billion trades with an overall volume of approximately 2.96 trillion, executed over 17 trading venues.

We use the volume Herfindahl index to provide a measure on market fragmentation. This metric is commonly used in the industrial organization literature to measure fragmentation or concentration. The volume Herfindahl index for a given stock on day t is defined as

$$H_t^v = \sum_{k=1}^{K} \left(s_t^k\right)^2,$$

where s_t^k is the volume share of venue k on day t, and K is the number of trading venues. The Herfindahl index ranges from 0 to 1, with higher figures indicating less fragmentation in that particular stock. Our results reveal that over the periods studied, the Herfindahl index reduces progressively, indicating that the market is becoming more fragmented over time.

Table II shows the trade and volume breakdown by reporting exchanges of the data we have included in this investigation. Overall, a total of 572 unique tickers are included in the analysis, with an average of 518 tickers over each year. NASDAQ Stock Exchange has the largest amount of trades and volume among all reporting exchanges, followed by NYSE ARCA (SM).

4.2 Identification of Flash Crashes

4.2.1 Standard Approach

Mini flash crashes are characterized by a sudden burst in price volatility, where the price moves in a uni-directional fashion for multiple ticks within a short time span, often (but not necessarily) followed by an equally quick price reversal process. Both up and down crashes are possible. To provide a basis for comparison, we begin by applying the widely used definition:

- uni-directional tick for more than a predefined threshold tickcount (10)
- time window is within a predefined short interval (1.5 seconds)
- price change exceed a predefined threshold (0.8%)

Figure 2 provides two samples of actual mini flash crashes. The left figure shows a downward crash of International Business Machines Corporation (IBM) stock at NYSE on August 3, 2010, while the right figure shows an upward surge of Salesforce Inc. (CRM) stock at NYSE on October 22, 2010.

Figure 3 plots the actual ΔP, i^c and Δt with daily breakdown (under the same mini flash crash threshold). We note again the large number of crashes on May 6, 2010. On this day, price changes on a number of the down crashes are in excess of 60%, which is expected, given that there are a handful of stocks crashing to a penny before bounding back. These "penny" stocks include ACN (Accenture Plc.), CNP (CenterPoint Energy Inc.), EXC (Exelon Corp.), and PG (Procter & Gamble). We note also that while on average the time window of each crash last less than $1.5s$, on May 6 the crash time window occurs quite uniformly up to $1.5s$.

Figure 4 plots the histogram of the actual ΔP, i^c and Δt for each crash. Here, we see that most flash crashes are characterized by small uni-directional price changes, small tick counts and occurring over a short time window. From the histogram it is clear that crash frequency decays with increasing ΔP and i^c, implying that large price change or large tick count in a flash crash occurs less frequent, which is expected. However, for time window, while the majority of the flash crashes occurred within a short time span, we note that apart from the first bucket, crashes are close to uniformly distributed across all the buckets, highlighting the lack of strong time-dependence in mini flash crashes.

To gain a better intuition behind the relationship between thresholds for price change ($|\Delta P_{th}|$), tick count (i^c_{th}) and time window width (Δt_{th}) in flash crash identification, in Figure 5 we vary the threshold for each of these parameter while holding the other two constant. We then proceed to count the number of mini flash crashes identified and plot the corresponding histogram. Here, we observe that the number of mini flash crashes is a decreasing function of both price change threshold and tick count threshold. In other words, as price change and tick count threshold are increased, the number of flash crashes detected decreases, which conforms to our intuition. On the other hand, we observe that the number of mini flash crashes is a increasing function of the time window threshold. In other words, as the threshold of the time window width is increased, the number of flash crashes increases.

However, price change, tick count and time window are all measures of price variation and volatility clustering. Given the implicit relationship between these parameters, defining threshold values for all three of them is liable to misspecification, leading to a fraction of mini flash crashes being missed. For instance, a mini flash crash with a large price change occurring over only a handful of tick counts might be missed by this identification method which set a threshold value for tick count. Similarly, it is also possible for sizeable price change to occur in a non-uni-directional manner across several seconds and a large number of ticks, and consequently misses the time window threshold to be identified as a mini flash crash. Figure 6 provides examples of such instances discussed above. The left figure shows the stock MEE (Massey Energy Corporation) at Direct Edge X Stock Exchange on November 5, 2010, experiencing a price change of 9.25% over $6.5s$. However, the price movement is not strictly uni-directional. On the right figure, the upward surge of the stock CHK (Chesapeake Energy Corporation) at NASDAQ Stock Exchange on April 16, 2010 is significant, but with less than 10 downward ticks, it fails to satisfy the definition of mini flash crash according to the definition above. These examples illustrate how a broader class of mini flash crashes will be missed if we scan for mini flash crashes by specifying hard threshold values for price change, tick count, and time window.

4.2.2 State-space Model Approach

As we have alluded to in earlier sections, we define mini flash crashes as abnormal price process. Under the state space model formulated in the previous section, transaction prices can be interpreted as noisy measurement of the unobservable true price process. For every observed transaction price, the state space model will update the latest state estimates of the true price and the state errors. Under this framework, mini flash crashes can be readily identified whenever transaction prices occur beyond a threshold distance away from the true price process. This distance can be conveniently measured in units of the state error estimates, which lends itself to intuitive interpretation.

One of the two main contribution of this paper is the formulation of a state-space model based on the Kalman filter approach as a more robust framework to identify mini flash crashes. Using the state estimates of the true price and measurement (trade) price processes, along with the error estimates, we can readily track the distance between the true price and the measurement price. We specify a z-score that measures the distance in units of the state estimate error, and identify deviations in the two price processes greater than this threshold as mini flash crashes. In other words, mini flash crashes are intuitively defined as significant deviation away from a normal process in returns. The effect of volatility clustering, and the relationship between price change and time window are implicitly accounted for within the framework.

Figure 7 provides an illustration of how Kalman filter method can capture sudden movement in prices. The left column shows three randomly selected crashes during the May 6, 2010 Flash Crash, and the right column shows three randomly selected crashes during the April 23, 2013 hoax tweet crash (where a false news tweet caused market to plunge, only to recover within seconds). The cross (blue) denotes the traded price, while the dot (green) denotes the true price estimate based on our state-space model. The dashed (red) lines denote the error estimate around the true price. These figures clearly illustrates the intuition behind the mechanism of the state-space model in mini flash crash detection: plotting the *a posteriori* true state with $\pm z \cdot \sigma$ overlay (where z is the z-score and σ is the standard error of the price process), we see that in general the true state is following the measurement process very closely. However, in the event of a mini flash crash, due to the rapid and extreme price movements, the measurement process deviates away from the true price process very quickly, as the Kalman filter formulation dictates that the true price process cannot deviate significantly away from its previous value, as its variance is proportional to the time elapsed from one transaction to the next. On the other hand, whenever a flurry of trade activities (measurement) takes place within a short time window at comparable prices, the width of the $\pm z \cdot \sigma$ overlay will reduce, as the state-space model becomes increasingly certain that the true price is not far away from the transaction price. In the event of a flash crash where the trade price moves away in a drastic fashion, the measurement process will cross over this $\pm z \cdot \sigma$ overlay. We will then identify this extreme price movement as a flash crash.

Using the state-space model formulated, we proceed to scan our entire dataset for flash crashes over the 4 calendar year included in this investigation (2010–2013). Table III shows the number of flash crashes detected as the threshold z-score is varied from 2 to 12. As one would expect, for smaller z-score values, the number of crashes detected is high. On the other hand, choosing too high a threshold will result in only a small number of crashes captured. For this work, we have chosen the threshold value of $z = 6$, and the results reported in subsequent sections of this paper are based on this threshold and the 7,492 crashes detected. Nevertheless, we have performed robustness check to verify that the conclusions and statistical properties reported in our paper remain consistent when other z-score thresholds are selected instead.

Figure 8 shows the number of mini flash crashes detected on each day over the period included in this analysis (Jan-2010 through to Dec-2013). Major economic or market events are labeled with red dashed lines for days with large number of mini flash crashes. These events are also tabulated in Table IV. From Figure 8 and Table IV, we clearly observe that the number of mini flash crashes spikes up on days when there are major market events or important information releases. Going through the number of crashes per day over the period included in our study, the top 10 event days include the Flash Crash (May 6, 2010), the Hoax tweet (April 23, 2013), the Knight Capital Glitch (August 1, 2012), Black Monday (August 8, 2011), and a collection of important Federal Reserve announcements on quantitative easing and bond buying programs, resulting in jumps in the US stock market. This indicates that mini flash crashes could be interpreted as processes where market assimilates new information or events at very high speed. This in turn lead to price crashes or surges which may or may not be followed by a quick reversal, depending on the validity of the new information.

4.3 Statistical Analyses of Mini Flash Crashes

Market microstructure theory postulates that price movements occur primarily due to the arrival of new information, and this information is incorporated efficiently into market price. Other variables such as trading volume and market liquidity are observed to be related to the volatility of the returns.

In this section, we perform in-depth statistical analyses on the 7,492 mini flash crashes detected based on our state-space model approach. The objective is to study important crash statistics like overall price change, number of transactions, and overall time taken of each mini flash crash, and to assess their sensitivities to factors like market capitalization, liquidity, volatility, average price level, and the trading venue of the stock. To this end, we group mini flash crashes into quintile to compare their statistics. This also allows us to understand the robustness of different stocks during mini flash crashes.

4.3.1 Market Capitalization

We start by investigating the sensitivity of crash statistics to the market capitalization of a stock experiencing mini flash crash. Stocks with larger market capitalization are expected to be more robust against extreme price movements due to their larger size and generally higher liquidity. This intuition is backed by academic research. For example, Brogaard, Hendershott and Riordan (2014) mention that larger market capitalization stocks are more liquid, have smaller spreads, with more market depth, and return volatility is also the lowest. Other robustness measure of larger market capitalization stocks have also been documented in Conrad, Wahal and Xiang (2015).

Table V provides a full breakdown of crash statistics by average market capitalization, split into quintile. A total of 7,492 crashes are detected across the 5 market capitalization quintiles. Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are also tabulated. Our results show that as market capitalization of the stock increases, the mean price change of the crashes decreases, while tick count and time window increases. This observation conforms to the intuition that stocks with larger market capitalization in general tend to be more liquid and have more depth. Consequently, their price movement are smaller during crashes, and it takes a larger number of ticks and a longer time period compared to smaller market capitalization stocks which are less liquid. Anderson-Darling test, Kruskal test, and Fligner test are performed across the crashes categorized under the 5 quintiles to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the 5 market capitalization quintiles. The test results conclusively reject the null hypothesis, supporting our observation that different market capitalization result in different crash statistics, with larger market capitalization exhibiting more robust characteristics during mini flash crashes.

4.3.2 Volatility

Volatility is a widely used gauge in the financial market to convey information about the extent of fluctuation in the asset return. According to Aït-Sahalia and Saglam (2013), liquidity provision declines when market volatility increases, which can lead to episodes of market fragility. In this context, volatility can also be viewed as a market risk indicator. Consequently, stocks with smaller annualized volatility should be expected to be more robust to extreme price movements.

There are two main ways to define and quantify volatility. The first is to define it as the square root of intraday realized variance, which can be calculated as the sum of log price changes squared. Alternatively, one could also use the option volatility, defined as the implied volatility to calculate the fair value of stock options using a Black-Scholes model. In the table below, we use the option implied volatility as the measure of volatility, though we have also performed the same analysis using intraday realized volatility and obtained similar conclusion.

Table VI provides a full breakdown of crash statistics by annual volatility, with the 7,492 crashes split across the 5 annual volatility quintiles. Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are also tabulated. Note that as the annual volatility of the stock increases, the mean price change of the crashes also increases, implying that more volatile stocks experience larger price movements during crashes. On the other hand, tick count and time window both decreases as the volatility of the stocks increases. This is intuitive, as crashes of more volatile stocks are generally expected to occur in shorter time spans. Anderson-Darling test, Kruskal test, and Fligner test are performed across the crashes categorized under the 5 quintiles to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the 5 annual volatility quintiles. The test results conclusively reject the null hypothesis, demonstrating that the volatility of a stock will affect its crash statistics, with smaller volatility stocks exhibiting more robust behavior during mini flash crashes.

4.3.3 Liquidity

The size of transactions is informative in reflecting the forces of supply and demand during price discovery beyond the intrinsic information. In fact, high liquidity facilitates price discovery, as demonstrated by Barclay and Hendershott (2003). Stocks with higher liquidity should be expected to be more robust during extreme price movements. Numerous empirical studies (Karpoff (1987) and Barclay, Litzenberger and Warner (1990)) have identified a positive relationship between trading volume and volatility. Weber and Rosenow (2006) argue that limited liquidity will manifest in stock price fluctuations, and also that limited liquidity is a necessary prerequisite for the occurrence of extreme price fluctuations, manifesting as high volatility. While one would not expect extreme price changes to be caused by limited trading volume alone, investigation carried out by Manganelli (2005) does support the hypothesis that the combined effect of trading volume and liquidity can account for price changes quantitatively.

Table VII provides a full breakdown of crash statistics by average volume, split into quintile. Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are also tabulated. As the average trading volume of the stock increases, the mean price change (in %) of the crashes decreases. This shows that more liquid stocks are more robust and less sensitive to extreme price movements. Anderson-Darling test, Kruskal test, and Fligner test are performed across the crashes categorized under the 5 quintiles to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the 5 average price quintiles. The test results conclusively reject the null hypothesis, showing that trading volume or liquidity information does impact crash statistics, with more liquid stocks being more robust during mini flash crashes.

4.3.4 Average Price

Apart from standard features like size, volatility, and liquidity, another interesting stock characteristic to investigate is its average price. As Birru and Wang (2016) point out, the level of a company's stock price is arbitrary as it can be adjusted by the company itself by altering the number of shares outstanding. However, there is clear evidence that nominal prices do influence investors' behavior. Low-priced stocks tend to be less liquid and have smaller market capitalizations. Investors exhibit a psychological bias in the manner in which they perceive nominal prices. Gompers and Metrick (2001) show that institutional ownership is increasing in stock price, while Kumar and Lee (2006) find that individuals tend to hold low-price stocks. Furthermore, Green and Hwang (2009) suggest nominal prices are relevant to investors when constructing and rebalancing their portfolios. Given this, one should expect higher dollar value stocks to be more robust during extreme price movements.

Table VIII provides a full breakdown of crash statistics by average price, split into quintile. Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are also tabulated. As the average price of the stock increases, the mean price change (in %) of the crashes decreases. This shows that stocks with lower dollar value experience larger price movements (in percentages) during crashes, while stocks with higher dollar-value are comparatively more robust. Similarly, high dollar value stocks experience larger amount of tick counts and longer time windows during crashes. Anderson-Darling test, Kruskal test, and Fligner test are performed across the crashes categorized under the 5 quintiles to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the 5 average price quintiles. The test results conclusively reject the null hypothesis, confirming that average price level does impact crash statistics, with higher priced stocks exhibiting more robust characteristics during mini flash crashes.

4.3.5 Exchanges

Schapiro (2010) pointed out that trading activities in U.S. equity market is fragmentation over numerous trading venues with varying degree of liquidity and efficiency. This section set out to determine whether mini flash crashes occur with comparable frequency and characteristics over these trading venues. Table IX provides crash statistics grouped by exchanges. We split the 7,492 crashes detected across the top 5 reporting exchanges by number of crashes, and group the rest of the exchanges into "others". Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are tabulated. The largest amount of crashes detected takes place at NYSE, followed by NASDAQ. However, price change during crashes at NYSE and NASDAQ are among the smallest, while tick count during crashes are among the largest. Notably, other reporting exchanges experience comparatively larger price change and smaller tick count during crashes. Anderson-Darling test, Kruskal test, and Fligner test are performed across the crashes to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the reporting exchange groups. The test results conclusively reject the null hypothesis.

4.3.6 Regression Analysis

In Table X, we perform ordinary least square (OLS) regression of crash statistics (price change, tick count, and time window) on the underlying stock parameters:

$$C_i = \beta_0 + \beta_1 \mathrm{MCap}_i + \beta_2 \bar{\sigma}_i + \beta_3 \bar{S}_i + \beta_4 \bar{V}_i + \beta_5 \mathbb{1}_{\mathrm{NQEX}} + \beta_6 \mathbb{1}_{\mathrm{EDGE}} + \beta_7 \mathbb{1}_{\mathrm{NYSE}} + \beta_8 \mathbb{1}_{\mathrm{PACF}} + \beta_9 \mathbb{1}_{\mathrm{EDGX}} + \epsilon_i,$$

where C_i is either ΔP, i_c, or Δt of stock i, MCap_i is the market capitalization of the stock, σ_i is the volatility of stock, $\bar{S}_i$ is the average price of the stock, $\bar{V}_i$ is the trading volume of the stock, and $\mathbb{1}_{\mathrm{EDGE}}$, $\mathbb{1}_{\mathrm{NYSE}}$, $\mathbb{1}_{\mathrm{PACF}}$ and $\mathbb{1}_{\mathrm{EDGX}}$ are dummy variables which equals 1 if the crash happens at that specific exchange, and 0 otherwise. In addition to the regression coefficient β, the standard errors σ_{se} and the significance level (two tail test) are also presented. The regression model adjusted R^2 for $|\Delta P|$, i_c and Δt are given by 0.22, 0.88 and 0.43, respectively.

The β coefficients show that larger market capitalization stocks generally experience smaller price movement, and takes more ticks and longer time during crashes. Furthermore, stocks with higher volatility or lower liquidity experience large price movement in smaller amount of ticks during crashes. Apart from that, the larger the stock price, the less the extreme its price movement during crashes. The t-statistics are computed with the Newey and West (1987) adjustment for serial correlations. We apply the recommendation in Newey and West (1994) to select the number of lags with the Bartlett kernel. The regression results are consistent with the comprehensive statistical tests results reported in earlier part of this section.

5 Conclusions

The consensus that mini flash crashes are becoming increasingly frequent has prompted many researchers to investigate its nature and implication. While mini flash crashes have been the subject of a number of recent studies, most of the research efforts focus on relatively small sets of data. In this work, we have performed in-depth analyses on 572 US large cap stocks (S&P500 constituent stocks, including its most recent revisions) over 4 calendar years (2010-2013), comprising of more than 11 billion trades and a trade volume tally of approximately 2.9 trillion.

We begin by pointing out how existing methodologies could potentially lead to sizeable mini flash crashes being missed. To address this, we formulate a state-space model, and define mini flash crashes intuitively as price movement experiencing a significant deviation from a normal process. In order to account for intraday volatility into our state-space model, we incorporate the intraday GARCH model proposed by Engle and Sokalska (2012) in our formulation. This allows us to perform state estimates in our model by taking diurnal variation of price volatility into consideration. We apply this methodology to our entire dataset, and found a large number of mini flash crashes.

Plotting the number of crashes per day across the 4 calendar years included in our study, we observe that the number of crashes spike when there are major information releases or market events. We proceed to conduct comprehensive statistical analysis on the 7,492 mini flash crashes detected. Our results indicate that stocks with larger market capitalization, lower volatility, higher liquidity, and higher price are more robust during mini flash crashes — they generally experience smaller price changes (in %), and take a larger amount of tick count (transaction) and a longer time window to crash. Regression analysis is also performed, to support the conclusion. We note that these observations are compatible with market microstructure literature on the role of information on trading.

Our work suggests that mini flash crashes can be interpreted as market assimilating new information at an ever increasing speed. This complements the view in recent research that high speed algorithmic trading improves liquidity by speeding up price discovery.

References

Agrawal, S., Azar, P. D., Lo, A. W. and Singh, T. (2018), 'Momentum, mean-reversion, and social media: evidence from StockTwits and Twitter', *Journal of Portfolio Management* **44** (7), pp. 85–95.

Aït-Sahalia, Y. and Saglam, M. (2013), 'High-frequency traders: taking advantage of speed', *National Bureau of Economic Research, Cambridge, MA*, [Online] Available at: https://www.nber.org/papers/w19531.

Andersen, T. G. and Bollerslev, T. (1997), 'Intraday periodicity and volatility persistence in financial markets', *Journal of Empirical Finance* **4** (2-3), pp. 115–158.

Barclay, M. J., Litzenberger, R. H. and Warner, J. B. (1990), 'Private information, trading volume, and stock-return variances', *Review of Financial Studies* **3** (2), pp. 233–254.

Barclay, M. and Hendershott, T. (2003), 'Price discovery and trading after hours', *Review of Financial Studies* **16** (4), pp. 1041–1073.

Birru, J. and Wang, B. (2016), 'Nominal price illusion', *Journal of Financial Economics* **119** (3), pp. 578–598.

Black, F. (1971), 'Toward a Fully Automated Stock Exchange', *Financial Analysts Journal* **27** (4,6).

Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A. and Sokolov, K. (2018), 'High frequency trading and extreme price movements', *Journal of Financial Economics* **128** (2), pp. 253–265.

Brogaard, J., Hendershott, T. and Riordan, R. (2014), 'High frequency trading and price discovery', *Review of Financial Studies* **27** (8), pp. 2267–2306.

Chakravarty, S., Jain, P. K., Upson, J. and Wood, R. (2012), 'Clean Sweep: Informed Trading through Intermarket Sweep Orders', *Journal of Financial and Quantitative Analysis* **4**, pp. 415–435.

Colliard, J.-E. (May 2013), 'Catching Falling Knives Speculating On Market Overreaction', *European Central Bank Working Paper Series* **1545**.

Conrad, J., Wahal, S. and Xiang, J. (2015), 'High-frequency quoting, trading, and the efficiency of prices', *Journal of Financial Economics* **116** (2), pp. 271–291.

Dugast, J. and Foucault, T. (2014), 'False News, Informational Efficiency and Price Reversals', *HEC Paris Research Paper* (FIN-2014-1036), [Online] Available at SSRN: `http://ssrn.com/abstract=2398904`.

Engle, R. F. and Sokalska, M. E. (2012), 'Forecasting intraday volatility in the US equity market: multiplicative component GARCH', *Journal of Financial Econometrics* **10** (1), pp. 54–83.

Golub A. Keane, J. and Poon, S. H. (2012), 'High Frequency Trading and Mini Flash Crashes', *Manchester Business School Working Paper Series*, [Online] Available at: `http://arxiv.org/abs/1211.6667`.

Gompers, P. A. and Metrick, A. (2001), 'Institutional investors and equity prices', *Quarterly Journal of Economics* **116** (1), pp. 229–259.

Green, T. C. and Hwang, B.-H. (2009), 'Price-based return comovement', *Journal of Financial Economics* **93** (1), pp. 37–50.

Harvey, A. C. (1991), *Forecasting, structural time series models and the Kalman filter*, Cambridge University Press: Cambridge.

Hasbrouck, J. (1999), 'The dynamics of discrete bid and ask quotes', *Journal of Finance* **54** (6), pp. 2109–2142.

Hasbrouck, J. and Saar, G. (2013), 'Low-Latency Trading', *Journal of Financial Market* **16** (4), pp. 646–679.

Hendershott, T., Jones, C. M. and Menkveld, A. J. (2011), 'Does Algorithmic Trading Improve Liquidity?', *Journal of Finance* **66** (1), pp. 1–33.

Johnson, N., Zhao, G., E., H., Qi, H., Johnson, N., Meng, J. and Tivnan, B. (2013), 'Abrupt rise of new machine ecology beyond human response time', *Scientific Reports* **3** (2627).

Karpoff, J. (1987), 'The relation between price changes and trading volume: a survey.', *Journal of Financial and Quantitative Analysis* **22**, pp. 109–126.

Kirilenko, A. A., Kyle, A. S., Samadi, M. and Tuzun, T. (2017), 'The flash crash: the impact of high frequency trading on on electronic market', *Working Paper* **72** (3), pp. 967–998.

Kirilenko, A. A. and Lo, A. W. (2013), 'Moore's law versus Murphy's law: algorithmic trading and its discontents', *Journal of Economic Perspectives* **27** (2), pp. 51–72.

Kumar, A. and Lee, C. M. C. (2006), 'Retail investor sentiment and return comovements', *Journal of Finance* **61** (5), pp. 2451–2486.

Lee, S. S. and Mykland, P. A. (2012), 'Jumps in equilibrium prices and market microstructure noise', *Journal of Econometrics* **168** (2), pp. 396–406.

Lee, W. B., Cheng, S. F. and Koh, A. (2011), 'Would price limits have made any difference to the "Flash Crash" on May 6, 2010?', *Review of Futures Markets* **9**, pp. 55–93.

Madhavan, A. (2012), 'Exchange-Traded Funds, Market Structure, and the Flash Crash', *Financial Analysts Journal* **68** (4), pp. 20–35.

Manganelli, S. (2005), 'Duration, volume and volatility impact of trades', *Journal of Financial Markets* **8** (4), pp. 377–399.

Menkveld, A. J. (2014), 'High Frequency Traders and Market Structure', *The Financial Review.*

Menkveld, A. J. and Zhou, B. Y. (2018), 'The flash crash: a cautionary tale about highly fragmented markets', *Forthcoming in Management Science.*

Nanex (2010), 'Flash Equity Failures in 2006, 2007, 2008, 2009, 2010, and 2011', [Online] Available at: `http://www.nanex.net/FlashCrashEquities/FlashCrashAnalysis_Equities.html`.

O'Hara, M. and Ye, M. (2011), 'Is market fragmentation harming market quality?', *Journal of Financial Economics* **100** (3), pp. 459–474.

O'Hara, M. (2014), 'High-Frequency Trading and Its Impact on Markets', *Financial Analysts Journal* **70** (3).

Ozenbas, D. and Schwartz, R. A. (2018), 'Do high frequency trading firms provide two-sided liquidity?', *Journal of Portfolio Management* **44** (7), pp. 63–74.

Schapiro, M. L. (2010), 'Strengthening Our Equity Market Structure', [Online] Available at: `http://www.sec.gov/news/speech/2010/spch090710mls.htm`.

Scholtus, M., van Dijk, D. and Frijns, B. (2014), 'Speed, algorithmic trading, and market quality around macroeconomic news announcements', *Journal of Banking and Finance* **38**, pp. 89–105.

Sparrow, C. (2012), 'The failure of continuous markets', *Journal of Trading* **7** (2), pp. 44–47.

Wah, E. and Wellman, M. P. (2013), 'Latency Arbitrage, Market Fragmentation, and Efficiency: A Two-market Model', *Proceedings of the* 14^{th} *ACM Conference on Electronic Commence*, pp. 855–872.

Weber, P. and Rosenow, B. (2006), 'Large stock price changes: volume or liquidity?', *Quantitative Finance* **6** (6), pp. 7–14.

Welch, G. and Bishop, G. (2006), 'An Introduction to the Kalman Filter', *Technical Report TR95-041, University of North Carolina.*

Table I: Descriptive Statistics of Stocks Included in Analysis

This work aims to focus on large market capitalization stocks – the selection criterion is based on annual S&P 500 index constituent, including recent changes made over each period. A total of four calendar years (2010-2013) is included in the analysis. 572 unique tickers are included in the analysis, with an average of 518 tickers over each year. This dataset is representative in that it includes several known cases of major flash crashes. The table below summarizes the number of trades and volume transacted each year. The volume Herfindahl index, a measure of market fragmentation, is also included in the table. Over the periods studied, the Herfindahl index reduces progressively, revealing the fact that the market is becoming more fragmented over time.

	2010	2011	2012	2013	Total
Tickers	516	519	518	519	572
Trade ($\times 10^6$)	2937.78	3102.39	2561.75	2403.68	11005.6
Volume ($\times 10^9$)	911.38	805.36	652.17	592.98	2961.89
Herfindahl Index	0.2988	0.263	0.194	0.1743	0.2116

Table II: Summary Statistics of Trade & Volume across Exchanges

Trading activities in U.S. equity market is split among over numerous public exchanges. For the dataset used in our study, Nasdaq Exchange has the largest amount of trades and volume among all reporting exchanges, followed by NYSE ARCA (SM).

Exchange	Trade ($\times 10^6$)	Volume ($\times 10^9$)
NASDAQ	3836.31	1161.93
NYSE ARCA (SM)	1473.64	306.23
BATS	1348.53	247.38
New York Stock Exchange	1210.30	495.33
NASDAQ NSX TRF	719.19	250.23
Direct Edge X	654.24	139.93
NASDAQ OMX BX	454.56	86.40
Direct Edge A	436.16	85.83
FINRA TRF	347.06	75.31
BATS Y	324.09	53.57
NASDAQ OMX PSX	79.85	20.83
National Stock Exchange	50.04	11.99
CBOE	29.62	6.15
International Securities Exchange	20.60	6.18
NYSE MKT	15.62	2.78
FINRA ADF	3.42	1.07
Chicago Stock Exchange	2.36	10.74

Table III: Threshold z-score and Mini Flash Crashes Detected

The number of mini flash crashes detected as the threshold z-score in the state-space model is varied from 2 to 12. For small threshold values, the number of crashes detected is high, as one would expect. On the other hand, choosing too high a threshold will result in only a small number of crashes identified. For this work, we have chosen the threshold value of $z = 6$, and the results reported in subsequent sections of this paper are based on this threshold and the 7,492 crashes detected. Nevertheless, we have performed robustness check to verify that the conclusions and statistical properties reported in our paper remain consistent when other z-score thresholds are selected instead.

z-score	Crashes
2	182,877
4	22,121
6	7,492
8	3,691
10	2,071
12	1,308

Table IV: Top 10 Event Days with Largest Number of Crashes

Details on date, event, and the number of mini flash crashes identified on the top 10 event days with the largest number of crashes. Table is arranged in chronological order.

Date	Event	Number of Crashes
May 6, 2010	The Flash Crash	454
November 3, 2010	Fed announced purchase of $600bn of Treasuries	84
August 8, 2011	Black Monday	204
September 21, 2011	$400 billion stimulus, markets tumble	76
June 28, 2012	Euro sovereign debt crisis summit, US market affected	56
August 1, 2012	Knight Capital Glitch	114
September 13, 2012	3^{rd} round of quantitative easing, market jumps	142
April 23, 2013	Hoax tweet and flash crash	136
September 28, 2013	Dow, S&P hit record as Fed holds off tapering	392
December 18, 2013	Fed announcing tapering in bond buying program	368

Table V: Descriptive Statistics & Statistical Tests by Market Capitalization

Full breakdown of crash statistics by average market capitalization, split into quintile. A total of 7,492 crashes are split across the 5 market capitalization quintiles. Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are tabulated. Our results show that as market capitalization of the stock increases, the mean price change of the crashes decreases, while tick count and time window increases. This is intuitive, since stocks with larger market capitalization in general tend to be more liquid and therefore exhibit greater depth. Consequently, on average their price movement are less, while tick count is larger and time window is longer, compared to smaller market capitalization stocks which are comparatively less liquid. Anderson-Darling test, Kruskal test, and Fligner test are also performed across the crashes categorized under the 5 quintiles to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the 5 market capitalization quintiles. The test results conclusively reject the null hypothesis.

Panel A: Descriptive Statistics of Crashes by Market Capitalization

Avg		Price Change (%)			Tick Count			Time Window (s)		
MarketCap	Crashes	Mean	Median	Stdev	Mean	Median	Stdev	Mean	Median	Stdev
3.46	916	1.532	0.860	0.5640	5.90	5.0	1.85	0.772	0.40	0.2748
7.78	1,317	1.298	0.756	0.6168	6.79	5.0	2.95	0.794	0.40	0.2952
12.82	1,322	1.076	0.730	0.3084	7.20	6.0	3.29	0.796	0.45	0.3258
21.68	1,626	1.018	0.704	0.2484	7.71	6.0	3.73	0.840	0.45	0.2916
91.53	2,311	0.871	0.678	0.1950	8.83	7.0	4.91	0.872	0.50	0.3162

Panel B: Statistical Tests on Crash Statistics by Market Capitalization

Anderson-Darling Test	155.982***	364.846***	10.343***
Kruskal Test	457.536***	777.711***	11.523***
Fligner Test	996.315***	1433.057***	77.740***

Table VI: Breakdown & Statistical Tests by Volatility

Full breakdown of crash statistics by volatility, split into quintile. Similar to Table V earlier, 7,492 crashes are detected across the 5 volatility quintiles. Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are tabulated. Note that as the volatility of the stock increases, the mean price change of the crashes also increases, implying that more volatile stocks experience larger price movements during crashes. On the other hand, tick count and time window both decreases as the volatility of the stocks increases. This behavior is expected, as crashes of more volatile stocks are generally expected to occur in shorter time spans. Anderson-Darling test, Kruskal test, and Fligner test are also performed across the crashes categorized under the 5 quintiles to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the 5 volatility quintiles. The test results conclusively reject the null hypothesis.

Panel A: Descriptive Statistics of Crashes

		Price Change (%)			Tick Count			Time Window (s)		
Volatility	Crashes	Mean	Median	Stdev	Mean	Median	Stdev	Mean	Median	Stdev
0.17	2,220	0.870	0.680	0.2196	8.97	7.0	4.82	0.822	0.40	0.3126
0.23	1,515	1.014	0.720	0.2724	7.56	6.0	3.58	0.746	0.40	0.2748
0.29	1,589	1.064	0.732	0.2100	7.27	6.0	3.56	0.928	0.55	0.3186
0.37	1,263	1.297	0.736	0.6414	6.69	5.0	2.80	0.804	0.45	0.3078
0.60	905	1.553	0.854	0.5574	6.00	5.0	2.10	0.804	0.45	0.2904

Panel B: Statistical Tests on Crash Statistics by Volatility

Anderson-Darling Test	139.713***	362.324***	22.758***
Kruskal Test	399.370***	770.777***	38.526***
Fligner Test	898.843***	1428.727***	114.867***

Table VII: Breakdown & Statistical Tests by Liquidity

Full breakdown of crash statistics by average trading volume, split into quintile – a total of 7,492 crashes distributed across 5 average volume quintiles. Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are tabulated. As the average volume of the stock increases, the mean price change (in %) of the crashes decreases. This shows that stocks with lower volume (liquidity) experience larger price movement during crashes. Similarly, stocks with lower liquidity tend to taker a smaller amount of tick count during crashes, while stocks with higher average volume are comparatively more robust – their price changes are smaller, with larger tick counts during crashes. Anderson-Darling test, Kruskal test, and Fligner test are also performed across the crashes categorized under the 5 quintiles to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the 5 average volume quintiles. The test results conclusively reject the null hypothesis.

Panel A: Descriptive Statistics of Crashes

Avg		Price Change (%)			Tick Count			Time Window (s)		
Volume	Crashes	Mean	Median	Stdev	Mean	Median	Stdev	Mean	Median	Stdev
1.14	1847	1.594	0.746	0.3072	7.51	6.0	3.78	0.834	0.42	0.3060
2.22	1682	1.276	0.740	0.3696	7.60	6.0	3.84	0.872	0.45	0.3300
3.49	1409	1.131	0.738	0.4794	7.44	6.0	3.66	0.786	0.45	0.2910
5.74	1299	1.082	0.726	0.5310	7.18	6.0	3.49	0.800	0.48	0.2904
17.02	1253	0.090	0.718	0.4056	8.24	6.0	3.29	0.812	0.50	0.2892

Panel B: Statistical Tests on Crash Statistics by Average Volume

Anderson-Darling Test	9.375***	41.137***	9.572***
Kruskal Test	16.298***	92.438***	8.054*
Fligner Test	9.572***	8.054***	10.923**

Table VIII: Breakdown & Statistical Tests by Price Level

Full breakdown of crash statistics by price level, split into quintile – a total of 7,492 crashes distributed across the 5 average price level quintiles. Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are tabulated. As the average price of the stock increases, the mean price change (in %) of the crashes decreases. This shows that stocks with lower dollar-value experience larger price movement during crashes. Similarly, stocks with lower average prices tend to taker a smaller amount of tick count and a shorter time window during crashes, while stocks with higher average prices are comparatively more robust – their price changes are smaller, with larger tick counts and longer time windows during crashes. Anderson-Darling test, Kruskal test, and Fligner test are also performed across the crashes categorized under the 5 quintiles to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the 5 average price quintiles. The test results conclusively reject the null hypothesis.

Panel A: Descriptive Statistics of Crashes

Avg		Price Change (%)			Tick Count			Time Window (s)		
Price	Crashes	Mean	Median	Stdev	Mean	Median	Stdev	Mean	Median	Stdev
15.02	704	1.680	0.856	0.7290	5.46	5.0	0.88	0.742	0.40	0.2682
29.93	1,110	1.372	0.712	0.2952	6.73	6.0	2.11	0.788	0.40	0.2850
42.84	1,583	1.158	0.708	0.2508	7.60	6.0	3.19	0.750	0.45	0.2958
61.88	1,904	1.034	0.714	0.3804	8.16	6.0	4.19	0.870	0.45	0.3204
122.15	2,191	0.774	0.714	0.2424	8.18	6.0	5.17	0.882	0.55	0.3300

Panel B: Statistical Tests on Crash Statistics by Average Price

Anderson-Darling Test	118.110***	451.395***	31.947***
Kruskal Test	332.659***	910.720***	1790.718***
Fligner Test	31.947***	56.095***	178.800***

Table IX: Breakdown & Statistical Tests by Exchange

Crash statistics split by exchange. We split the 7,492 crashes detected across the top 5 exchanges by number of crashes, and group the rest of the exchanges into "others". Mean, median, and standard deviation on crash parameters including price change (ΔP), tick count (i_c), and time window (Δt) are tabulated. The largest amount of crashes detected takes place at NYSE, followed by Nasdaq. However, price change during crashes at NYSE and Nasdaq are among the smallest, while tick count during crashes are among the largest. Notably, other exchanges experience comparatively larger price change and smaller tick count during crashes. Anderson-Darling test, Kruskal test, and Fligner test are performed across the crashes to test the null hypothesis that the distribution, median, and variance of price change, tick count, and time window are identical across the exchange groups. The test results conclusively reject the null hypothesis.

Panel A: Descriptive Statistics of Crashes

		Price Change (%)			Tick Count			Time Window (s)		
Exchange	Crashes	Mean	Median	Stdev	Mean	Median	Stdev	Mean	Median	Stdev
New York Stock Exchange	2,508	0.730	0.652	0.0786	10.09	9.0	5.03	0.508	0.15	0.2400
NYSE ARCA (SM)	930	0.860	0.692	0.1752	7.46	6.0	3.21	0.708	0.35	0.2796
NASDAQ	924	0.834	0.674	0.1530	8.18	7.0	4.13	0.534	0.25	0.2286
BATS	746	1.030	0.742	0.2550	6.88	6.0	2.72	0.856	0.50	0.2970
Direct Edge X	503	1.156	0.824	0.2580	6.26	5.0	2.24	0.960	0.55	0.3270
Others	1,835	1.888	1.094	0.7392	4.63	5.0	1.10	1.396	0.90	0.3390

Panel B: Statistical Tests on Crash Statistics by Exchange

Anderson-Darling Test	234.926***	391.813***	209.097***
Kruskal Test	681.322***	880.485***	490.509***
Fligner Test	1247.523***	1128.641***	378.162***

Table X: Regression analysis on crash and descriptive statistics.

Ordinary least square (OLS) regression of crash statistics (price change, tick count, and time window) on the underlying stock parameters:

$$C_i = \beta_0 + \beta_1 \text{MCap}_i + \beta_2 \bar{\sigma}_i + \beta_3 \bar{S}_i + \beta_4 \bar{V}_i + \beta_5 \mathbb{1}_{\text{NQEX}} + \beta_6 \mathbb{1}_{\text{EDGE}} + \beta_7 \mathbb{1}_{\text{NYSE}} + \beta_8 \mathbb{1}_{\text{PACF}} + \beta_9 \mathbb{1}_{\text{EDGX}} + \epsilon_i,$$

where C_i is either ΔP, i_c, or Δt. In addition to the regression coefficients β's, the standard errors σ_{se} and the significance level (two tail test) are also presented. The regression model adjusted R^2 for $|\Delta P|$, i_c and Δt are given by 0.22, 0.88 and 0.43, respectively. The β coefficients show that larger market capitalization stocks generally experience smaller price movement, and takes more ticks and longer time during crashes. Although the regression parameter suggests that more volatile stocks experience larger price movement, and takes shorter amount of time during crashes, this result is not statistically significant. The larger the stock price, the less the price movement during crashes. The t-statistics are computed with the Newey and West (1987) adjustment for serial correlations. We apply the recommendation in Newey and West (1994) to select the number of lags with the Bartlett kernel.

	Price Change			Tick Count			Time Window		
	β	σ_{se}	t	β	σ_{se}	t	β	σ_{se}	t
MCap_i	-0.1328	0.016	-8.133***	1.0832	0.064	16.805***	0.0090	0.010	1.928**
$\bar{\sigma}_i$	0.41	0.021	1.614*	-0.351	0.066	-1.703*	-0.0199	0.013	-0.528
$\bar{S}_i$	-0.0006	0.000	-3.210***	0.0016	0.001	1.617*	0.0008	0.000	3.064***
$\bar{V}_i$	-0.0053	0.002	-4.851***	-0.0073	0.002	-1.817*	0.0015	0.001	0.414
$\mathbb{1}_{\text{NYSE}}$	-0.5713	0.035	-16.142***	4.4998	0.130	34.681***	-0.3616	0.017	-21.409***
$\mathbb{1}_{\text{PACF}}$	-0.5181	0.036	-14.314***	1.9601	0.096	20.410***	-0.2626	0.017	-15.435***
$\mathbb{1}_{\text{NQEX}}$	-0.5351	0.037	-14.629***	2.7135	0.133	20.462***	-0.3497	0.017	-20.657***
$\mathbb{1}_{\text{BATS}}$	-0.4309	0.035	-12.343***	1.3923	0.087	15.993***	-0.1853	0.018	-10.515***
$\mathbb{1}_{\text{EDGX}}$	-0.3660	0.035	-10.559***	0.7786	0.078	9.941***	-0.1308	0.020	-6.504***
Intercept	1.2075	0.058	20.761***	3.3888	0.171	19.847***	0.5568	0.026	21.641***

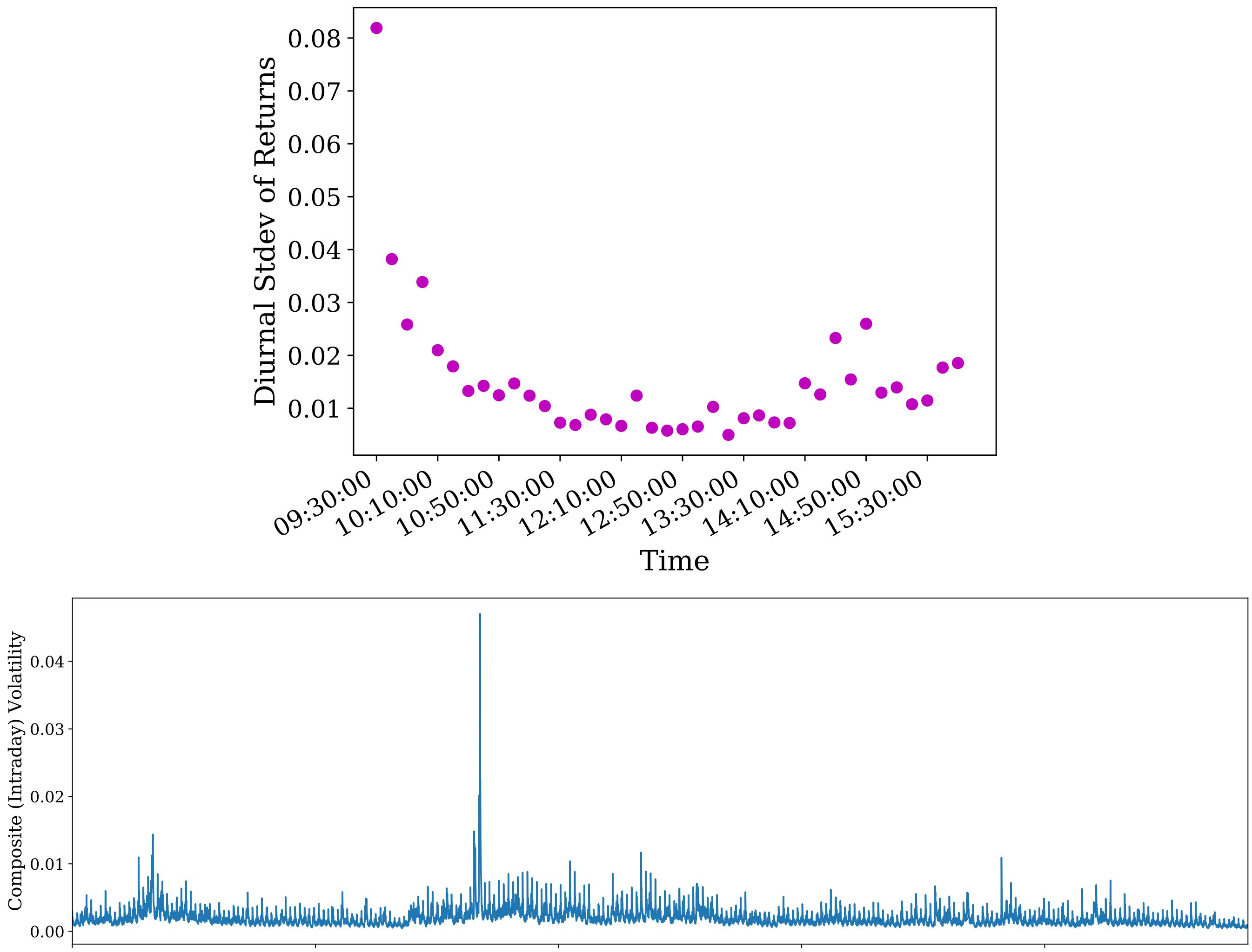

Figure 1: Sample volatility information for Apple stock (AAPL) over the entire year of 2010. The upper subplot is the diurnal volatility component. The lower subplot is the the composite intraday volatility calculated based on Engle and Sokalska (2012). The spike in volatility corresponds to the May 6 Flash Crash.

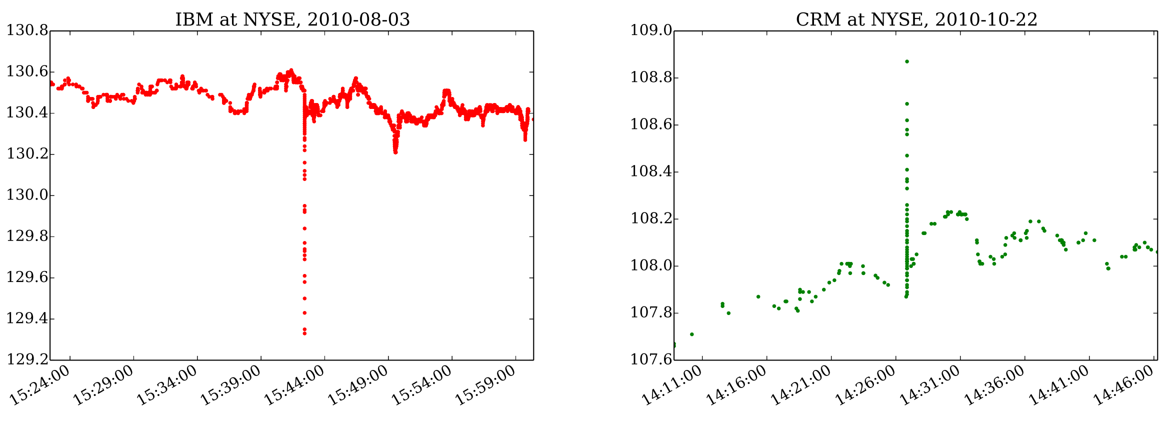

Figure 2: Two samples of mini flash crashes: left figure shows a downward crash of IBM stock at NYSE on August 3, 2010; right figure shows an upward surge of CRM stock at NYSE on October 22, 2010. In both cases, the prices swiftly revert to the prevailing level immediately after the crash/surge.

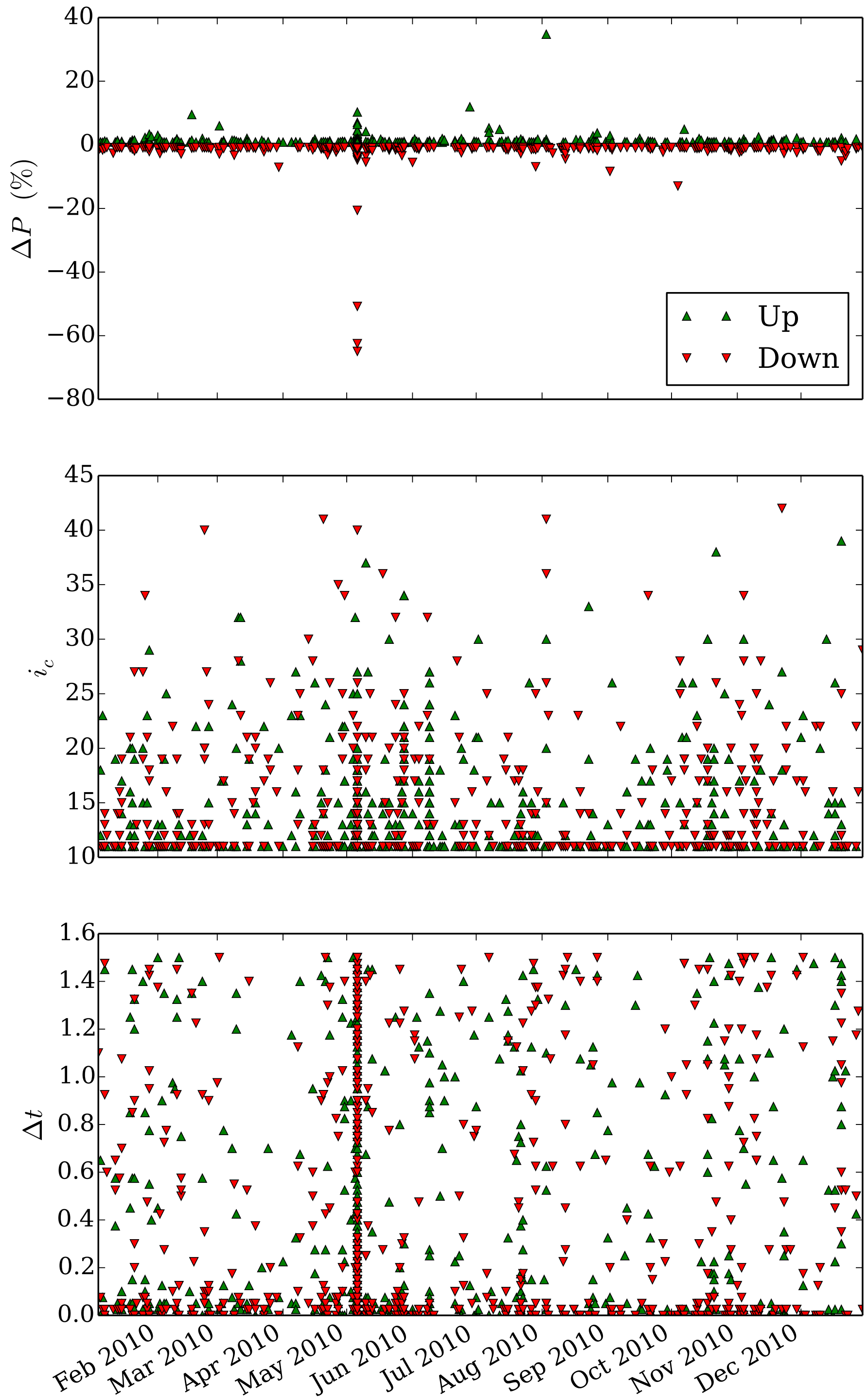

Figure 3: Daily breakdown of price change (%), tick count and crash time window.

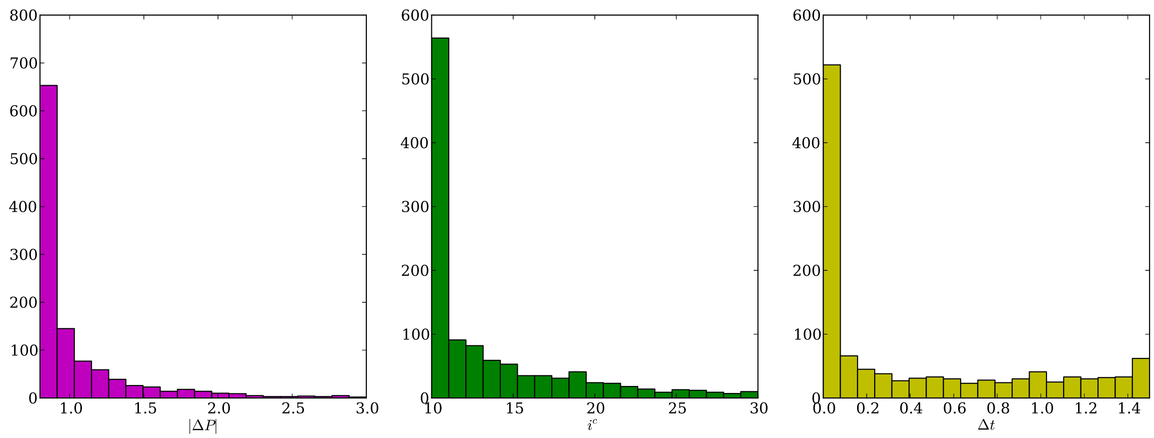

Figure 4: Histogram of Nanex crashes. $|\Delta P_{th}| = 0.8\%$, $i^c_{th} = 10$, $\Delta t_{th} = 1.5s$

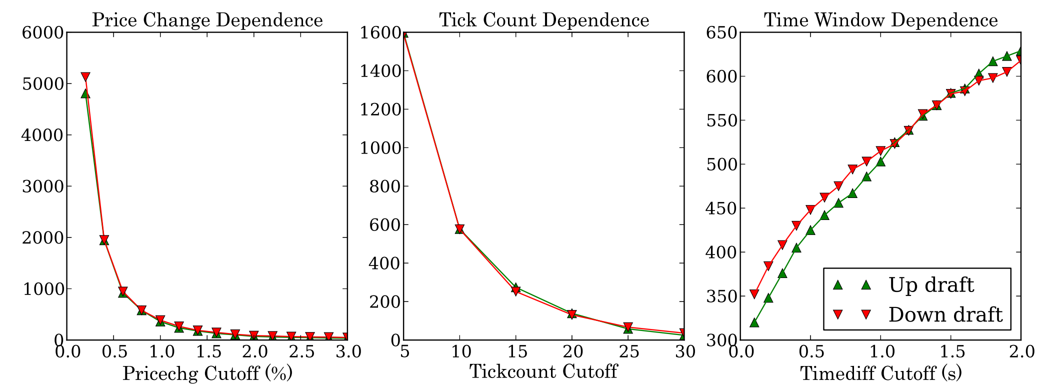

Figure 5: Registered flash crash dependence on cutoff parameters using existing methodology.

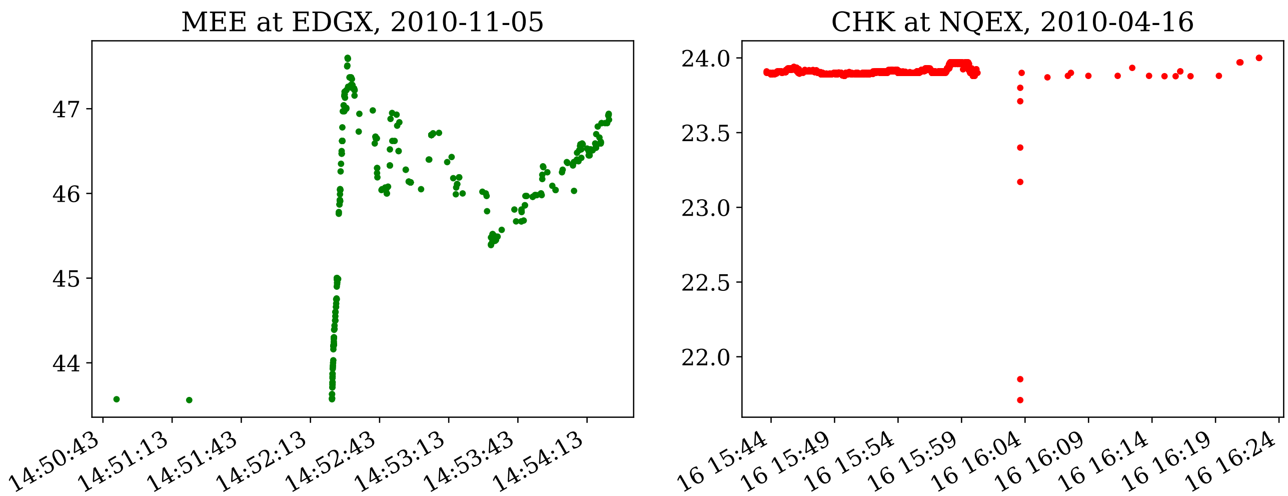

Figure 6: Two examples of crashes that will be missed by conventional crash detection. On the left figure, the upward surge occurs over a period of approximately 6 seconds and not strictly monotonically. On the right figure, the crash in price is significant, but with only 7 downward ticks. Both fail to satisfy the definition of mini flash crash in existing methodology.

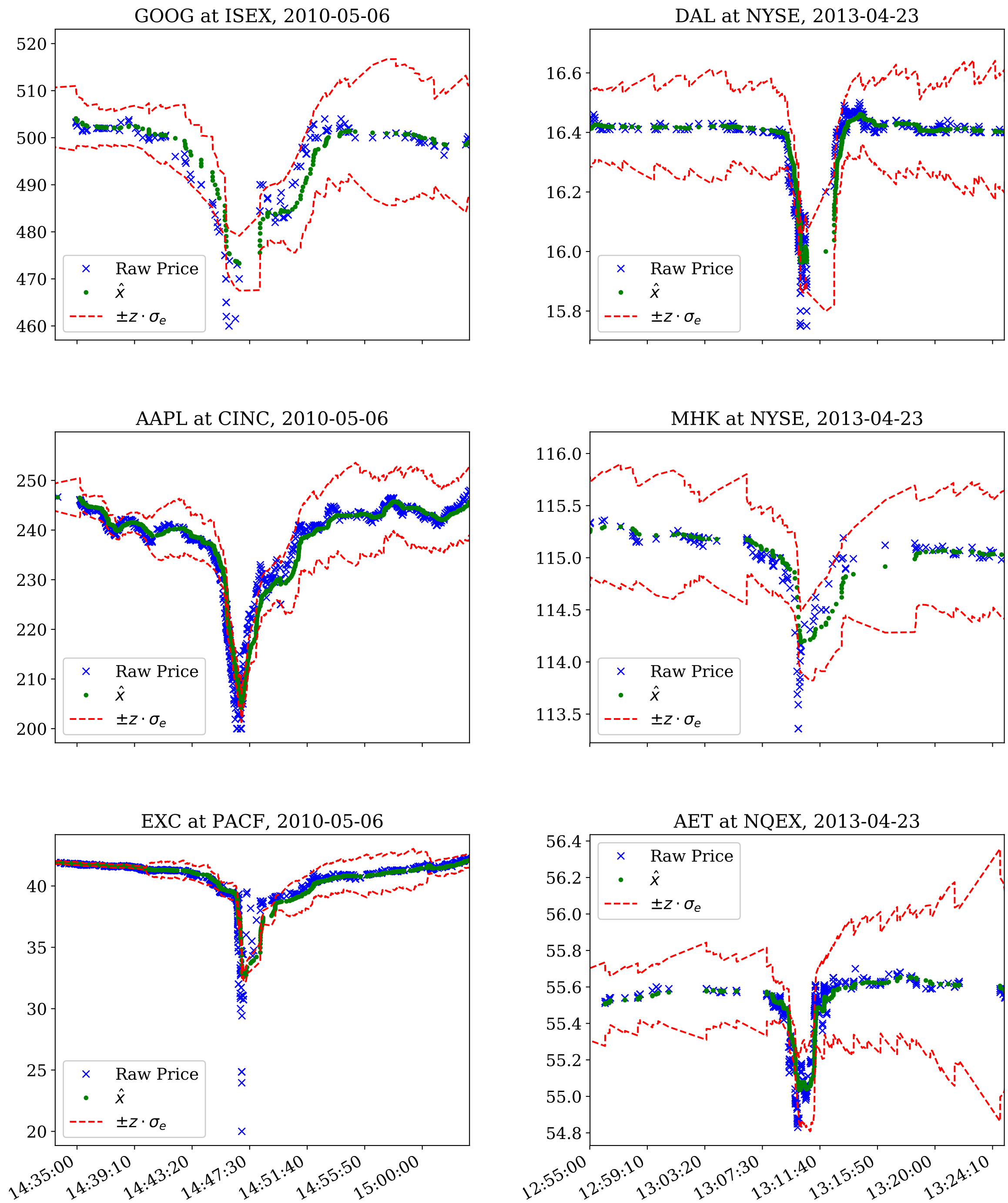

Figure 7: The left column shows 3 randomly selected crashes on May 6 2010 Flash Crash, and the right column shows 3 randomly selected crashes on April 23 2013 hoax tweet crash. The (blue) cross denotes the traded price, while the (green) dot denotes the true price estimate based on our state-space mdel. The dashed (red) lines denote the error estimate around the true price.

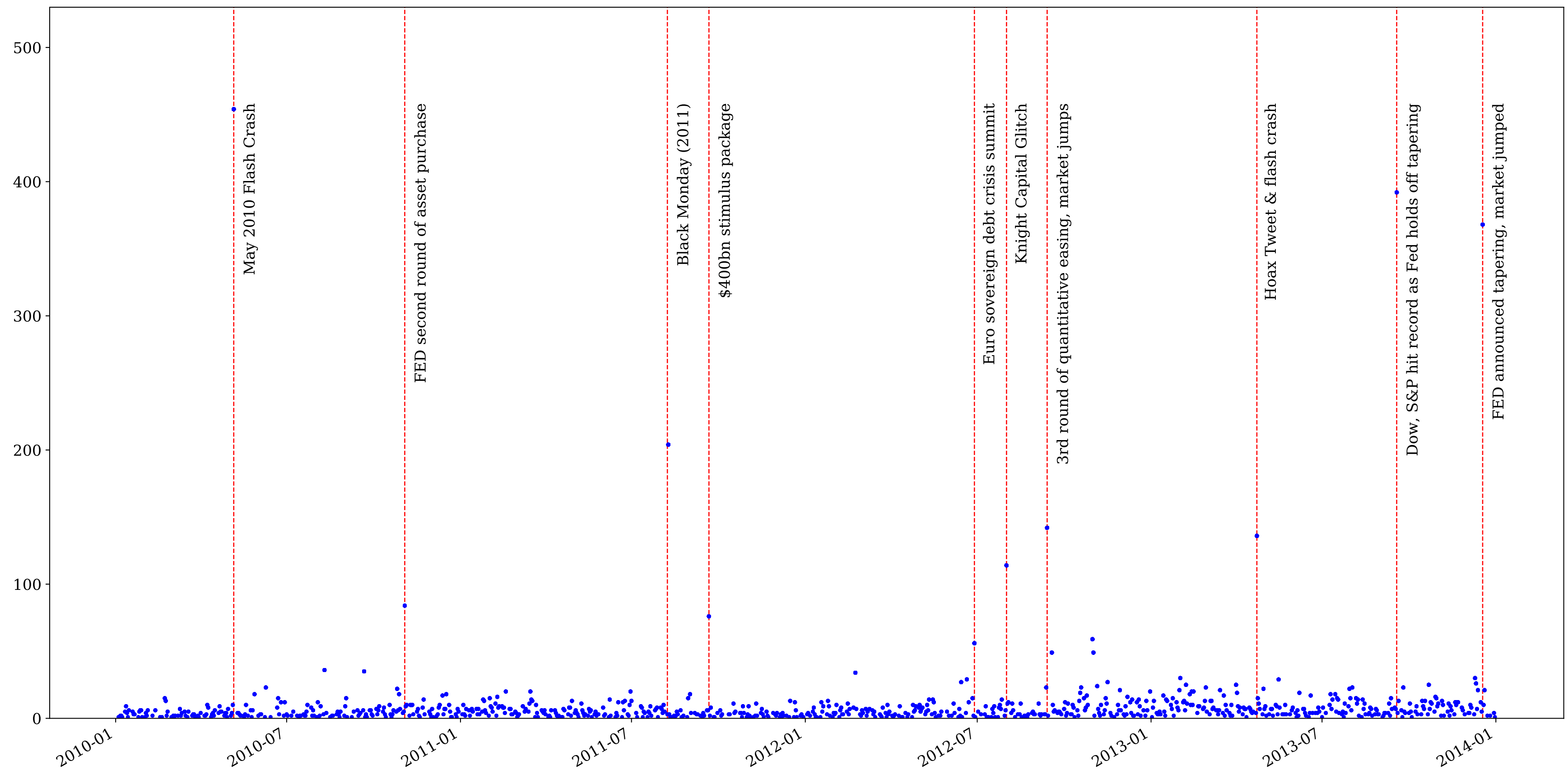

Figure 8: The number of mini flash crashes detected on each day over the period included in this analysis (Jan-2010 through to Dec-2013). Crash detection is based on the state-space model described in this paper. Major economic or market events are labeled with red dashed lines for days with large number of mini flash crashes.