High Frequency Trading: A Simulation

Wei Pan
MIT Media Lab

Alex Sandy Pentland
MIT Media Lab

Ren Cheng
Fidelity Investments

Lisa Emsbo-Mattingly
Fidelity Investments

December 20, 2012

1 The Simulation Model

1.1 The Price Model

Our model is largely an agent-based approach for modeling High Frequency Trading firms (HFTs) together with a discrete price model derived from discrete portfolio execution theories [1]. We divide a time period T into N even short interval of length $\tau = T/N$. S_n is the security price at time $t = n\tau$. We model the security price S_t dynamics according to a discrete arithmetic random walk:

$$S_t = S_{t-1} + \sigma\tau^{1/2}\xi_k - \tau g(\frac{h_t v_t}{\tau}) + \delta\tau, \tag{1}$$

for $t = 1, ..., N$. σ represents the volatility of the price, and $\xi_k \sim \mathcal{N}(0;1)$. h_t represents the net sale volumes of all liquid providing HFT funds. $g(v)$ is the price impact function, which we will discuss later. Based on findings by the Federal Commodity Futures Trading Commission (CFTC) [3], we model the price impact of HFTs using the volume of net trades from HFTs scaled by its percentage in overall trading volume (denoted as v_t in Eq. 1). We also establish a drifting factor in our model δ, which represents the directional change of price due to market conditions, external sell pressures, etc.

In the following content, we define a market crash as a change of price in 10% or more in a 30-minute period.

1.2 The HFTs

Based on observed evidence, it is safe to assume that HFTs (excluding these cross-market arbitrage HFTs) essentially buy at low and sell at high, i.e. a form of the mean reversion strategy based on evidences [3],. We consider in our simulation that the HFT universe is composed of $1, ..., K$ HFT agents. We examine their impact on one single fixed security (in the case of the flash crash it is ESM10). Each HFT system k uses a mean reversion strategy governed by four parameters $< a_k, c_k, b_k, r_k >$. b_k is the budget constraint for the fund k; c_k is the cyclic parameter; the HFT k will rebalance its net holding of the security every $c_n\tau$ seconds. During the rebalancing process, the net holding of HFT k of the security at time t (denoted as w_t^k) is determined by the current price vis-a-vis the price of the previous rebalance and its price sensitivity parameter r_k. We assume across different HFTs $r_1, ..., r_K = r$. To be more specific:

$$\tilde{w_t^k} = \begin{cases} w_{t-1}^k, & \text{if not rebalancing} \\ r_k((1 - \frac{S_{t-c_k}}{S_t}) - m_t^k), & \text{if rebalancing} \end{cases} \tag{2}$$

subject to a boundary limitation:

$$w_t^k = \begin{cases} \tilde{w_{t-1}^k}, & |\tilde{w_t^n}| < b_k \\ -b_k, & \tilde{w_t^n} < -b_k \\ b_k, & \tilde{w_t^n} > b_k \end{cases} \tag{3}$$

In addition, a_k is the phase shifting parameter, which controls at what time each HFT starts their first balance cycle. a_k is used to ensure that all HFT algorithms are running asynchronously. The parameter m_t^k is the market exposure parameter. In a classic equity long short strategy, m_t^k is used to eliminate the overall market exposure in the portfolio when an algorithm is performing mean reversion on multiple different securities. Since this paper focuses on a short period of time around a market crash, we can assume that $m_t^k \approx m^k$, which is time invariant.

Based on our definition, the overall net sell volume of the security of all HFTs at time t is:

$$h_t = \sum_{k=1,...,N} w_{t-1}^k - w_t^k, \tag{4}$$

in which m^k naturally cancels out with each other.

2 Results

2.1 Parameter Selections

[1] We fix the overall number of high frequency trading firms to $n = 20$, a number indicated by the CFTC [3]. For τ, we set τ to 1 second to limit our model computational complexity while providing a reasonable resolution. From multiple sources, we discover that HFTs generally hold stock for around $10 \sim 20$ seconds [3] [2]. We believe our second-level resolution captures most of dynamics for HFTs.

In the following discussion, we will assume that we are modeling the flash crash on May 6th, 2012. The annualized volatility for SPY ETF is around 40%, and the price of ESM10 is around \$1200. Therefore, we have an annualized arithmetic volatility of $\sigma = 40\% \times 1200 = \$480$ for ESM10.

We continue discussing the price impact function, $g(v)$. For this function, we use a linear model by Almgren et al. and CFTC [1, 3]:

$$g(v) = \gamma v, \tag{5}$$

where γ is a linear impact rate. As a rule of thumb, 10% of total volume will result in noticeable change in price, which we interpret as a price movement of more than one standard deviation. From historic data, we notice that the average one-day volume is around 2×10^6 contracts the day before the flash crash, and the one-day price volatility is around $\sigma/\sqrt{250} \approx \$30$. We therefore suggest a $\gamma \approx 5 \times 10^{-6}$. This is a rough estimation similar to Almgren et al. [1], and the readers are recommended to play around this value.

We continue to set r, the sensitivity factor for the weight of the trade based on its price change. The average price change for a 20-second interval with 20-second volatility $\sigma_{20s} = \frac{\sigma}{\sqrt{252 \times 6.5 \times 180}}$ is:

$$\Delta S = \int_0^\infty 2x\mathcal{N}(x; 0, \sigma_{20s})dx = \frac{2\sigma_{20s}}{\sqrt{2\pi}} \approx 0.69. \tag{6}$$

Also, we assume that 30% is the normal percentage of volume for HFTs in total market volume, therefore every second the volume of HFTs is around $2 \times 10^6/3600/6.5 \times 30\% = 25$. Therefore, $r \approx \frac{25}{0.69/1200} \approx 5 \times 10^4$.

The marginal payment for each ESM10 contract is \$2500 (by definition E-Mini contracts provide 50× leverage). We assume that the average fund

[1]Most market data, if not specified, comes from Bloomberg Professional Service.

size dedicated to this one particular future instrument is around 250M, allowing us to set the budget limit of every fund to be 1000 contracts, long or short. This also matches the estimate in empirical observations, where the net holding of contracts for HFTs rarely exceeds a thousand [3].

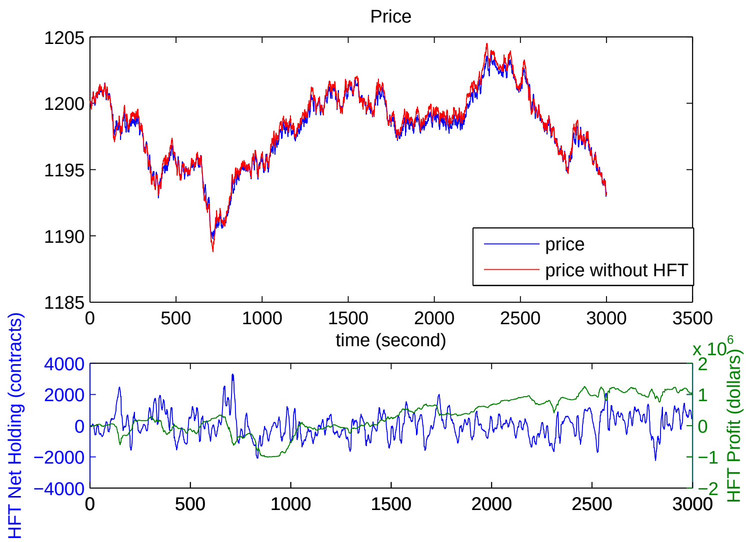
3 The Simulation

With the above parameters, we simulate the stock price movements with the designated 20 HFTs under no market drift ($\delta = 0$). Figure 1a illustrates one instance from our model, and we also show the *hypothetical* price of the market assuming there are no HFTs involved in red. We observe that even when HFTs are taking 30% of the market liquidity, they barely have any impact on market price. The net holdings of HFTs are between $[-2000, 2000]$ contracts, matching empirical observations of the real market [3]. The profit of HFTs, which is defined as the total earned across all HFT firms starting at zero dollar excluding transaction cost, suggests a positive and continuous profitability for all HFTs.

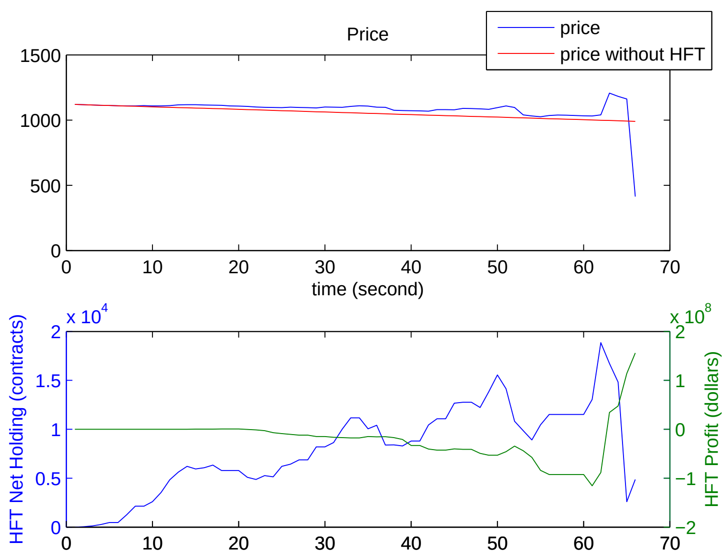
We continue to study the behaviours of HFTs with drift. We show the results with drift $\delta = 1$, which mimics a stressed market with sell pressure during the flash crash week. We show in Figure 1b how HFTs act as a cushion on the market, taking in some of the additional sell pressure. However, after a while, all HFTs start to sell in sync, and the market crashes.

The intuition behind the observation in Figure 1b is not obvious. Our basic assumption is that HFTs are agnostic, and they only react based on market data. Due to their budget limitation, no single HFT can create a market crash. When the market is behaving normally, the actions of other HFT firms will not dramatically impact market prices. However, when the market is under stress and HFTs run at higher capacity, their movements may actually cause price impacts. Without knowing the mechanism or reason behind price changes, other HFTs also quickly react to the price. Therefore, these price movements force all HFTs to act in sync. Such synchronization causes similar trades from a group of agnostic HFTs, and the market is quickly flooded by aggressive one-directional orders from HFTs, which leads to the eventual crash.

In fact, by varying the value of the market stress δ, we can study the system property of a HFT-powered market. As shown in Figure 2, when we increase the market pressure, the system can immediately switch from



(a)



(b)

Figure 1: a): Price chart(top) and HFT net holding and earning chart(bottom) are shown in this plot with default parameter settings and no market drift. The HFTs are not imposing an impact on the price. b): Price chart and earnings chart for an HFT-enabled market with a down drift ($\delta = 0.2$). HFTs provide an initial buffer for the down turn but quickly turn to sharply down in the market.

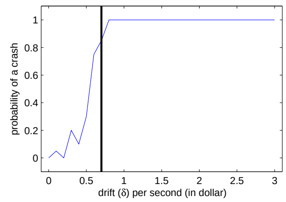


Figure 2: We run 100 trials for each value of δ, and study the likelihood of a crash(10% lost/gain in price in 30 seconds) at different values of sell pressure δ.

stable to unstable. The phase transition demonstrated by Figure 2 suggests a classical non-linear system behaviour.

3.0.1 HFT Participants

Selling pressure is not the only potential cause of a sudden crash, even an increased number of HFT participants ca lead to a crash. In fact, if the number of HFTs hits a threshold, HFTs will create very large market oscillations on their own. We demonstrate such a phenomenon in Figure 3a, which shows a sample from our simulation where we have 60 HFTs rather than 20.

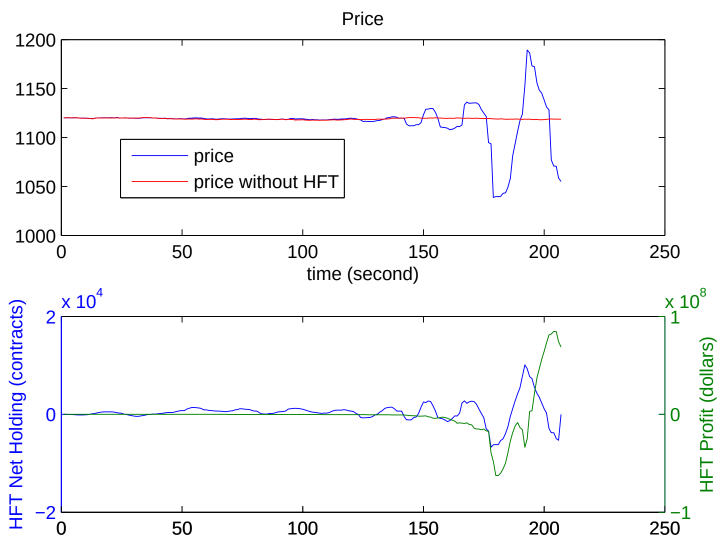
We continue to run our simulations to examine the probability of a crash under normal market condition ($\delta = 0$) in Figure 3b. Notice that the ESM10 market has about 16 HFTs, and our market is clearly close to the cliff shown in our simulation results.

4 Supporting Comments

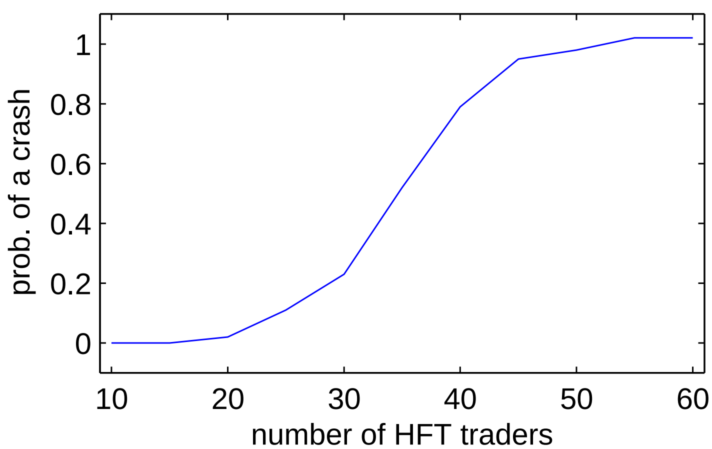
4.1 Can a partition sell algorithm (or just a large order) causes the flash crash directly?

The SEC report suggests that a large sell order (from Waddell&Reed) absorbed all available liquidity and exhausted the trade book, which eventually caused the flash crash. We consider that this argument lacks evidential support.

As Nanex research pointed out, the sell algorithm didn't reach its max. throttle during the crash. The actual sell process went smoothly and gradually until the point at which the futures market crashed [5].



(a)



(b)

Figure 3: a): One instance from our simulation with 60 HFTs. The price will gradually enter into an oscillation process even when $\delta = 0$. b): The probability of a crash in our simulation when we gradually increase the number of HFTs. Each configuration is run for 100 trials, and we keep $\delta = 0$.

4.2 Market Fragmentation

Some researchers argue that market fragmentation contributed to the crash [4]. This is a faulty argument for ESM, as the only market center for E-Mini futures are within the CME's Globex system.

4.3 Oscillation

Our model predicts an oscillation before the actual crash due to the positive feedback loop of the HFT system. We are able to observe such observation in the ESM (E-mini futures June Contract) oscillation (Figure 4). Our model suggests that during each cycle of oscillation, HFTs were increasing their holdings and getting ready to dump all of them in the next oscillation. Within a few cycles, the volumes from HFTs become the market majority, which completely eliminated all other participants on the trade book and inevitably brought down the stock. Our model also predicts an oscillation period to be around 20 seconds, which is exactly the case in the actual crash.

References

[1] R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *Journal of Risk*, 3:5–40, 2001.

[2] J. Creswell. Speedy neew traders make waves far from wall st., May 2010.

[3] A. Kirilenko, A. Kyle, M. Samadi, and T. Tuzun. The flash crash: The impact of high frequency trading on an electronic market. *Manuscript, U of Maryland*, 2010.

[4] A. Madhavan. Exchange-traded funds, market structure and the flash crash. 2011.

[5] L. Nanex. BWorld Robot Control Software. `http://nanex.net/FlashCrashFinal/FlashCrashAnalysis_WR_Update.html`, 2012. [Online; accessed 09-Aug-2012].

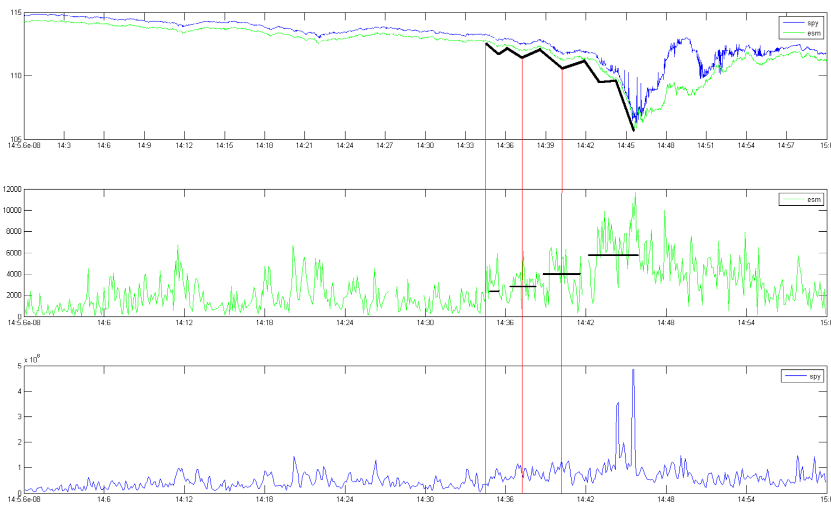


Figure 4: We show the price of SPY (S&P500 ETF) and ESM during the crash in the top row. The second row is the volume for ESM, and the third row is the volume for SPY. We mark the oscillation observed in the ESM price with black lines, and we also mark the gradual increase in volume with black lines in the second row.