INSIGHTS INTO HIGH FREQUENCY TRADING FROM THE VIRTU INITIAL PUBLIC OFFERING

GREG LAUGHLIN

1. Overview

The publication of Michael Lewis' *Flash Boys* in early April, 2014, generated an intense month-long wave of publicity surrounding high frequency trading (HFT) firms, much of it unfavorable. In a typical example, Paul Krugman, writing in a New York Times editorial[1] that appeared just after publication of *Flash Boys*, conflated HFT trading practices with the proliferation of complex derivatives products during the Financial Crisis, and implied that the entire endeavor is "a huge waste". An analysis with Google Trends (see Figure 1) indicates that news volume surrounding HFT spiked in April to levels substantially higher than any previously seen.

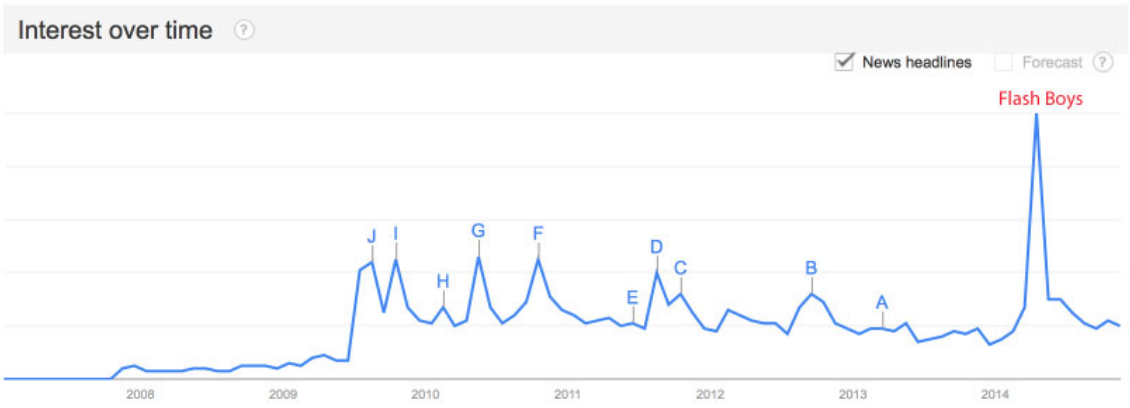

Figure 1: Google Trends chart of "High Frequency Trading"

The swirl of bad publicity surrounding HFT prompted Virtu Financial to indefinitely postpone its initial public offering[2]. The documents associated with Virtu's IPO filing[3],

[1]http://www.nytimes.com/2014/04/14/opinion/krugman-three-expensive-milliseconds.html
[2]http://dealbook.nytimes.com/2014/04/17/virtu-financial-said-to-shelve-i-p-o-plans/
[3]http://www.sec.gov/Archives/edgar/data/1592386/000104746914003044/a2219101zs-1a.htm

however, along with public statements made by company executives[4], provide an interesting window into the overall operation of a modern, large-scale HFT operation.

Analysis of the filing, in conjunction with other data, *also* permit an abundance of new, highly specific insights that can be extremely detailed. For example, in the documentation filed as part of the IPO[5], Virtu lists the subsidiary entities that it the firm controls. These include a Delaware corporation, "Blueline Comm LLC (f/k/a MVC Research LLC)". A scan of the FCC's wireless license database indicates that this subsidiary has registered a Chicago to New Jersey microwave communications network (one of approximately 20 currently in existence). Analysis of the FCC license files indicates that the Blueline/MVC network has a distance excess $d = 62.6$ km over the great circle distance. It enables Virtu to communicate pricing information and order flow between Aurora, IL and Carteret, NJ with a one-way latency of approximately 4.7 ms (compared to the 6.65 ms one-way latency of the Spread Networks fiber-optic path highlighted in *Flash Boys*). The Blueline portion of the network provides connectivity between Washington DC and Chicago.

Overall, one finds that the portrait drawn from the newly public data concerning Virtu's operations is largely at odds with the currently popular viewpoint surrounding HFT.

2. Profitability

In the June 4th, 2014 Bloomberg article, Virtu CEO Doug Cifu states that "51 to 52%" of Virtu's trades are profitable. Given the granularity of trading by a market maker, in which one is trying to win of order one spread ($0.01) per share, there is an unmentioned subtlety in this statistic, because at first glance, it implies that 48 to 49% of Virtu's trades are *unprofitable*. The key point, however, is that if 51 to 52% of trades are profitable (i.e. earn one $0.01 bid-offer spread per share), the remaining 48-49% must be apportioned between trades that are unprofitable and trades that scratch out (have net zero profitability, or which are very slightly negative due to exchange fees and SEC fees). For purposes of order-of-magnitude estimation, we can assume that among the 48-49% of non-profitable trades, approximately half will scratch out exclusive of fees, and half will lose of order one spread.

Assume a 51% win rate, a 24% lose rate, and 25% scratches. This implies an average profit of

$$\langle P \rangle = .51 \times 0.01 - .24 \times 0.01 = \$0.0027/\text{share}\,.$$

Virtu's IPO filing states that trading income from Americas Equities is ∼$440K/day, and this is ∼30% of their recent total daily trading income across all instruments. Given the above profit estimate of $0.0027/share, this implies that Virtu trades ∼160 million shares per day. Current US equity volume is ∼5 billion shares per day[6], which would put Virtu at ∼3% of the US stock market. Volumes were somewhat lower during the period described in

[4]http://www.bloomberg.com/news/2014-06-04/virtu-touting-near-perfect-record-of-profits-backfired-ceo-says.html

[5]http://www.secinfo.com/dVut2.n2vb.w.htm#1stPage

[6]http://batstrading.com/market_summary/

the IPO, and so on many days, Virtu would have been ∼5% of the total US equity market. It also suggests (if other major participants, e.g. Jump Trading, Teza Technologies, Tower Research Capital, etc. have similar per-share profitability), that current HFT US Equity income is of order $10M per day and **$2.5B** per year before fees and technology costs.

In order to gain an understanding of the number of shares per trade, I ran spot probes on a Nasdaq market session by parsing ITCH4.1 binary-format data[7]. In an example ∼1 hour segment, one finds 271,517 trades, a mean of 236.38 shares/trade, a median of 100 shares/trade, and a mode of 100 shares/trade. This suggests that while the both the median trade size and the most common trade size are $N \sim 100$ shares, the tail of larger trades (which can occur in lower-priced high-volume stocks such as BAC) leads to an average number of shares per trade of order $\langle N \rangle \sim 200$ shares.

These considerations suggest that Virtu does of order $N = (160,000,000)/(200) = 800,000$ US equity trades per day. Given that Virtu's US Equity business is $f \sim 30\%$ of their total business, this scales to $N_{\rm tot} = 2.9$ million trades per day across all venues and asset classes. This agrees with the statement in the Bloomberg article that Virtu "executes between 2.5 million and 3.5 million trades every day across all asset classes."

# Trades Per Day	Daily Edge ($p_{\rm d}$ ×100)	Odds of a Profitable Week
1	51.0%	51.9%
100	57.9%	67.0%
1,000	73.5%	92.1%
10,000	97.8%	99.999%
100,000	99.999999%	100%
800,000	100%	100%

Given these statistics, we can readily show that the chances of Virtu's operations experiencing a money-losing day (purely as a result of statistical fluctuations) are vanishingly small. To see this, let p be the probability that a particular trade is profitable. In Virtu's case, we adopt $p = 0.51$ for a 200 share trade. Further, define k as the integer sum of wins observed after n transactions. With a simple (+1, 0) win-loss model, possible values of k thus range from 0 to n. The probability of a *net win* on n transactions is then the sum over probabilities of outcomes with $k > n/2$,

$$P(k > \frac{n}{2}) \;=\; \sum_{k>n/2}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \quad \underset{n\to\infty}{\to} \quad \int_{n/2}^{\infty} \mathcal{N}(x, np, np(1-p))dx\,,$$

where the binomial distribution converges to the normal distribution, $\mathcal{N}(x, \mu, \sigma^2)$, with mean $\mu = np$ and variance $\sigma^2 = np(1-p)$ as the number of trades, n, becomes large[8]. The table below draws on this simple model to illustrate how the probability of a profitable trading day, defined as the "daily edge", $p_{\rm d}$, increases with the number of trades. At Virtu's

[7]For sample data, see: ftp://emi.nasdaq.com/ITCH/S030711-v41.txt.gz

[8]For $p \sim 0.5$, the transition to the continuous normal distribution is effectively complete at $n = 30$.

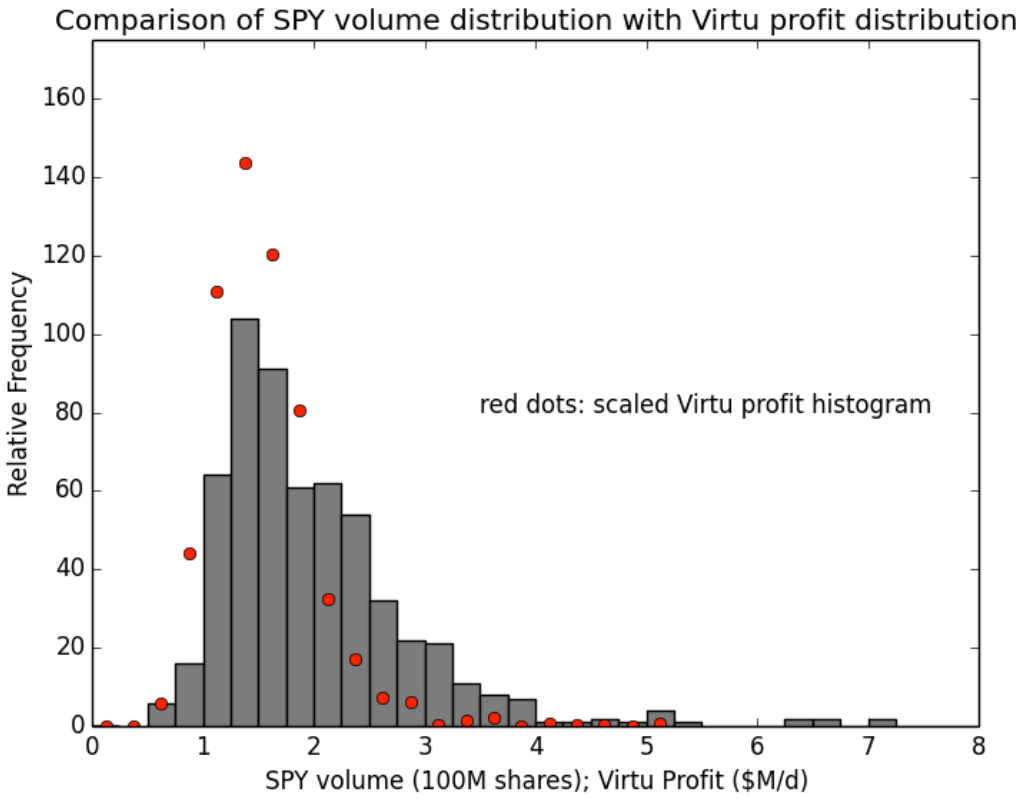

Figure 2: Comparison of SPY daily volume histogram (gray bars) with Virtu daily profitability distribution (red points). Virtu's daily PNL distribution is similar to the SPY volume distribution, suggesting that Virtu's daily profit variation is largely driven by fluctuations in trading volume.

rate of equity trading, the model (taken at face value) indicates an effectively zero chance of having an unprofitable full day of trading. We can therefore conclude that Virtu's single reported losing day (out of $N_{\rm d} = 1,278$ trading days from 2011-2014) resulted from either a technological or human-caused error.

In addition, the realized variance in profit over a large fixed number of trades will be very small. In other words, Virtu's Sharpe ratio during smoothly functioning trading, as measured in tick time rather than wall clock time, should be effectively infinite. As a consequence, the reported variation in Virtu's profit histogram (as shown in their IPO document) should closely resemble the histogram of daily volume of all instruments – shares, contracts, etc. – summed across all the markets in which the firm participates.

Figure 2 shows the result of testing this hypothesis through a direct comparison between the daily volume distribution in the SPY S&P 500 ETF (over 575 days) with Virtu's daily profit distribution (over 1278 days, but with each histogram bar multiplied by $f_{\rm r} =$

$575/1,278 = 0.45$ for purposes of direct comparison). The x-axis simultaneously represents 100M shares/day for SPY volume and \$M/day for Virtu's profits.

Virtu's profit histogram has a modestly sharper peak than the distribution of SPY trading volume, but to zeroth order, these are similar distributions. This supports the idea that Virtu is primarily a market maker with daily profit tied nearly linearly to total market volume. (Note that SPY volume cannot be a perfect proxy for all volume in all markets, and furthermore, there is not a 1:1 overlap in the samples of days being considered.)

3. Conclusion

High frequency trading is a complex phenomenon comprising myriad facets. Our brief analysis here, however, is consistent with a scenario that has also emerged from our two associated academic research studies of high frequency trading's impact on market microstructure.[9,10] At time scales where the finite speed of light is a consideration, equity HFT consists largely of market making, accelerated by arbitrage of fleeting intermarket price discrepancies, and by predictive models that are useful over a characteristic Lyapunov time[11] that is measured in tens of milliseconds.

Address: [1]Dept. of Astronomy and Astrophysics; [2]Center For Analytical Finance, UCSC

E-mail address: glaughli@ucsc.edu

[9]Laughlin, Aguirre & Grundfest (2014) "Information Transmission between Chicago and New York" http://onlinelibrary.wiley.com/doi/10.1111/fire.12036/abstract

[10]Aldrich, Heckenbach & Laughlin (2014) "The Random Walk of High Frequency Trading" http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2481201

[11]Bezruchko, B. & Smirnov, D. "Extracting Knowledge From Time Series: An Introduction to Nonlinear Empirical Modeling", Springer, 2010, pp. 56–57.