DERA

Working Paper Series

High-Frequency Trading Synchronizes Prices in Financial Markets

Austin Gerig

High-Frequency Trading Synchronizes Prices in Financial Markets*

Austin Gerig
Division of Economic and Risk Analysis
U.S. Securities and Exchange Commission†
geriga@sec.gov

High-speed computerized trading, often called "high-frequency trading" (HFT), has increased dramatically in financial markets over the last decade. In the US and Europe, it now accounts for nearly one-half of all trades. Although evidence suggests that HFT contributes to the efficiency of markets, there are concerns it also adds to market instability, especially during times of stress. Currently, it is unclear how or why HFT produces these outcomes. In this paper, I use data from NASDAQ to show that HFT synchronizes prices in financial markets, making the values of related securities change contemporaneously. With a model, I demonstrate how price synchronization leads to increased efficiency: prices are more accurate and transaction costs are reduced. During times of stress, however, localized errors quickly propagate through the financial system if safeguards are not in place. In addition, there is potential for HFT to enforce incorrect relationships between securities, making prices more (or less) correlated than economic fundamentals warrant. This research highlights an important role that HFT plays in markets and helps answer several puzzling questions that previously seemed difficult to explain: why HFT is so prevalent, why HFT concentrates in certain securities and largely ignores others, and finally, how HFT can lower transaction costs yet still make profits.

Keywords: algorithmic trading; automated trading; high-frequency trading; statistical arbitrage.

JEL Classification: G14, G19.

*I thank J.D. Farmer, K. Glover, D. Michayluk, and F. Reed-Tsochas for helpful comments and suggestions. This work was supported by the European Commission FP7 FET-Open Project FOC-II (no. 255987).

I Introduction

Over the past 10 years, high-frequency trading (hereafter HFT) has gone from a small, niche strategy in financial markets to the dominant form of trading. It currently accounts for approximately 55% of trading volume in US equity markets, 40% in European equity markets, and is quickly growing in Asian, fixed income, commodity, foreign exchange, and nearly every other market[1]. Although a precise definition of HFT does not exist, it is generally classified as autonomous computerized trading that seeks quick profits using high-speed connections to financial exchanges.

Policy makers across the globe are spending considerable effort deciding if and how to regulate HFT[2] On the one hand, HFT appears to make markets more efficient. Algorithmic trading in general, and HFT specifically, increases the accuracy of prices and lowers transaction costs[11, 14, 4, 16]. On the other hand, HFT appears to make the financial system as a whole more fragile. The rapid fall and subsequent rise in prices that occurred in US markets on May 6, 2010 (known as the "Flash Crash"), was, in part, due to HFT[13]. Because HFT firms do not openly disclose their trading activities, it has so far been unclear how and why HFT produces these outcomes; a circumstance that has greatly increased the controversy surrounding its existence.

In this paper, using a special dataset supplied by NASDAQ, I present evidence that HFT synchronizes security prices in financial markets. By 'synchronize', I mean the following – to the extent that two securities are related to one another, HFT activity ensures that a price change in the first security coincides nearly instantaneously with a similar price change in the second security. Synchronization is a gargantuan task[3]

[1]Several research firms provide estimates of HFT activity for subscribers; examples are the TABB Group, the Aite Group, and Celent. Publicly, this information is available in articles such as "The fast and the furious", Feb. 25, 2012, *The Economist* and "Superfast traders feel the heat as bourses act", Mar. 6, 2012, *Financial Times*.

[2]See for example the SEC document "Concept Release on Equity Market Structure" available at www.sec.gov, the ESMA document "Guidelines on systems and controls in a highly automated trading environment for trading platforms, investment firms and competent authorities" available at www.esma.europa.eu, the European Commission document "Consultation on financial sector taxation" available at ec.europa.eu, and the BIS Foresight project "The Future of Computer Trading in Financial Markets" available at www.bis.gov.uk.

[3]There are over one thousand transactions per second in US equities alone during the trading

tailor-made for HFT: it is profitable for the firms that do it and can only be done with high-speed computerized trade.

To understand the effects of price synchronization, I modify a standard model of price formation[9] so that it includes multiple related securities. I find that when prices are synchronized, transaction costs are reduced, prices are more accurate, and that informed investors – those who always submit a buy (sell) order when the price will be higher (lower) – make less profits.

The intuition behind these results is straightforward. As an example, suppose that an event occurs which increases the likelihood that country X will default on its sovereign debt. This information is processed by specialized firms who quickly buy securities that track the probability of X's default. The prices of these securities increase, and if markets are synchronized, then the prices of all other securities adjust as well. As a result, an investor who purchases or sells any security in the market receives a more accurate price. Transaction costs are reduced because liquidity providers are more confident in market prices and require less of a price concession to transact with an order. In finance, this is known as a reduction in adverse selection costs[3, 2, 8].

If transaction costs are lower, then average investors benefit from synchronization. So, who loses? When prices are synchronized, information diffuses rapidly from security to security and informed investors are made somewhat redundant. In the model, they make less profit as a result.

Although price synchronization is normally beneficial in markets, it can also have harmful effects. When prices are tightly connected to one another, localized errors quickly propagate through the financial system. In addition, there is potential for incorrect relationships between securities to be enforced, making prices more (or less) correlated than economic fundamentals warrant. Finally, during times of market stress, HFT firms are impelled to leave the market if their systems observe events outside the parameters they are programmed to handle – a circumstance that causes

day (see "U.S. Consolidated tape Data" available at www.utpplan.com).

liquidity to disappear at the precise time it is needed the most.

II Evidence

In Fig. 1, I show the rapid increase in HFT volume and the corresponding increase in market efficiency over the last decade. In Fig. 1(A), HFT estimates are from the TABB Group. In Fig. 1(B), the cost of a buyer initiated transaction is measured as the transaction price, p, minus the current prevailing midpoint price for the security, m, (for a seller initiated transaction, the cost is $m - p$). The transaction error is the absolute difference between the transaction price, p, and the midpoint price 1 minute later, x, (see the diagram in the figure). Data is from Thomson Reuters and includes 35 stocks during the last full week of February in 2000, 2005, and 2010. The 35 stocks are a subset of the 40 large-cap stocks available in the NASDAQ HFT data that is used below (5 of the stocks from the NASDAQ data are not available in the Reuters data for all time periods).

In Fig. 2(A), I show how prices have become more synchronized over the last decade. Using the same data as in Fig. 1(B), I measure the average normalized price response of security i to a price movement in security $j \neq i$ (see the supplemental material for details[1]). In 2000, it took several minutes for a price movement in stock j to be fully incorporated into the price of stock i. In 2005, this occurred in about 1 minute, and in 2010 it took less than 10 seconds. Fig. 2(B) shows that it is HFT activity that keeps prices synchronized. Using data from NASDAQ that flags HFT activity in 120 stocks during the last week of February, 2010 (see the supplemental material for a full description of the data[1]), I take the 40 largest stocks and calculate the average normalized price response of stock i to a price movement in stock j (black curve). I separate this response into the amount due to HFT activity (green curve), non-HFT activity (blue curve), and an amount that could not be categorized either way (red curve). As seen in the figure, an overwhelming majority of the initial price response is due to HFT activity.

III Model

To study the effects of price synchronization in detail, I modify a standard model of price formation[9] so that it includes multiple related securities. In total n securities, $i = 1, 2, \dots, n$, exist and are traded asynchronously over a single period. During the trading period, one unit-sized order to buy or sell is submitted for each security, $O_i \in \{B_i, S_i\}$. Submitted orders are immediately transacted by liquidity providers at the fair price, i.e., at the expected future price of the security, p_i. The original price of security i is m_i, and the final price at the end of trading, x_i, increases or decreases with equal probability, $x_i \in \{x_i^+, x_i^-\}$, where $x_i^+ = m_i + \delta_i$ and $x_i^- = m_i - \delta_i$, with $\delta_i > 0$. So in summary, for each security i, one order is submitted, transacts at price p_i, and the final price is either δ_i higher or δ_i lower than the original price.

In real markets, informed individuals correlate their orders with future price changes, i.e., buying tends to correspond with increases in prices and selling with decreases in prices. To include this effect, I assume that a buy order is more likely when the final price of the security is higher, and that a sell order is more likely when it is lower. For security i, $\mathcal{P}(B_i|x_i^+) = \phi_i > 0.5$ and $\mathcal{P}(B_i|x_i^-) = (1 - \phi_i) < 0.5$. Finally, in real markets, securities are related to one another so that their price changes are correlated. To include this effect, I assume that the price change of security i and j are correlated with correlation coefficient, $\rho_{i,j} = 0$. To complete the model, I assume that the orders for securities are independent of one another, except through the indirect dependence caused by the correlations already assumed, $\mathcal{P}(O_1, O_2, \dots | x_1, x_2, \dots) = \mathcal{P}(O_1|x_1)\mathcal{P}(O_2|x_2)\dots$.

With this simple model, synchronizing the prices of securities (allowing the price of security j to affect the price of security i) lowers transaction costs and increases the accuracy of prices. See the supplemental material for a full proof[1]. Here, I show the result with an example. Assume that the market contains only two securities, $n = 2$, and the initial price of each security is $m_1 = m_2 = 50$ which can increase or decrease by $\delta_1 = \delta_2 = 1$. For both securities $\phi_1 = \phi_2 = 0.75$, and their price changes are correlated with $\rho_{1,2} = 0.8$. The diagram in Fig. 3 analyzes the expected transaction cost, $c(B_1) = E[p_1 - m_1|B_1]$, and average pricing error, $e(B_1) = E[|x_1 - p_1||B_1]$, for a buy order in security 1, when prices are and are not synchronized (see the supplemental material for details[1]). As seen in the diagram, by allowing order flow in security 2 to affect the price of security 1, a buy order for security 1 costs less and is priced more accurately.

When prices are synchronized, transaction costs are lower in the model. For this reason, average investors – individuals who do not correlate their orders with final price changes – do not lose as much money. Who, then, is compensating average investors? When prices are synchronized, information diffuses rapidly from security to security and informed investors who trade different but related securities are forced to compete with one another. They make less profit as a result. In the above example, their average profit is reduced from 0.5 to 0.46875 per transaction[1].

Price synchronization is normally beneficial in markets, but it also can have harmful effects. If shared misconceptions exist in the population of investors within the model – causing for example, a large number of sell orders to be submitted even though the final prices of most securities are higher – then synchronization makes transaction prices less accurate. In addition, when prices are tightly connected to one another, errors quickly spread through the financial system. To mitigate this risk, HFT firms can program their systems to exit the market when errors are detected. But determining the difference between an extreme event and an error is precisely the type of problem that machines find difficult[17]. Machines that continually stay in the market risk propagating errors when they arise. Machines that leave the market at the first sign of an abnormality will often disappear at the precise time they are most needed. The end result is a financial system that becomes unstable during times of stress and behaves very much like US markets did during the Flash Crash[13].

IV HFT Activity

Although HFT firms synchronize prices, this does not imply it is their main activity. Just how important is synchronization to HFT? In Fig. 4A, I plot the relationship between the level of HFT activity within a stock and the correlation strength of that stock to other stocks. Correlations are between the 30 second returns of each stock and the equal-weighted average 30 second returns of all 120 stocks. HFT activity is measured as the fraction of overall share volume attributable to HFT for each stock. The correlation between these variables is 0.80, and the R^2 from a linear fit is 0.64. HFT activity varies significantly from security to security, and synchronization explains the majority of this variance.

To determine if HFT is enforcing plausible economic relationships between securities, I calculate the minimum spanning tree of the correlation network[15] for the 40 large-cap stocks (Fig. 4B). The ticker for each stock is shown on the corresponding node, and nodes are color-coded according to their GICS (Global Industry Classification Standard) sector. The correlation structure of these stocks at 30 second intervals – largely set by HFT – corresponds well with the economic relationships of the companies.

Most HFT firms are run by scientists and engineers, and it is unlikely that they pay close attention to economic fundamentals and create a map of market structure that updates as fundamentals change. Instead, it is more likely that HFT firms are dependent on feedback mechanisms that punish them when the structure they enforce is incorrect. If and how this feedback mechanism works is an important area of future research.

V Analogy to Animal Groups

It is interesting to compare the above results to recent findings in ecology. In animal groups, synchronized behavior facilitates information transfer between individuals, which increases the accuracy of decisions and allows fewer resources to be allocated to information gathering[6, 5]. A simple example is a school of fish. By synchronizing their behavior, fish can scan their environment using "many eyes", which allows them to quickly evade threats or move towards potential food sources.

Financial markets are similar. In markets, the state of the economy is monitored by a large number of investors who quickly broadcast any changes to each other and the rest of society via price movements[10, 7]. By synchronizing prices, HFT allows the "many eyes" of different investors to function as one coherent group, which results in price trajectories that look like the motions of schooling fish (Fig. 5).

Just as in animal groups, synchrony in financial markets leads to an increase in efficiency and a reduction in the resources spent on informed individuals. However, also as in animal groups, synchronization can have harmful effects; shared misconceptions among individuals in a group are amplified when behaviors are synchronized[5, 18].

VI Conclusions

The evidence above suggests that HFT plays an important role in financial markets. By synchronizing prices, HFT facilitates information transfer between investors, which increases the accuracy of prices and redistributes profits from informed individuals to average investors by reducing transaction costs.

Synchronization, however, is not a panacea for markets. When prices are tightly connected to one another, errors can quickly propagate throughout the financial system if safeguards are not in place. In addition, if shared misconceptions exist among investors, they are amplified so that prices are less accurate overall. Finally, synchronization can create spurious structure in markets if information about the changing relationships of securities does not make its way to the high-frequency domain.

In sum, these results help answer several puzzling questions about HFT that previously seemed difficult to explain: (1) why HFT is so prevalent, (2) why HFT increases market efficiency under normal conditions but leads to instability during times of stress, (3) why HFT concentrates in certain securities and not in others, and finally, (4) how HFT can lower transaction costs yet still make profits.

REFERENCES

[1] Supplemental material available at http://www.austingerig.com/research/high-frequency-trading/.

[2] Baruch, S. and G. Saar (2009). Asset returns and the listing choice of firms. *Review of Financial Studies 22*(6), 2239–2274.

[3] Biais, B., L. Glosten, and C. Spatt (2005). Market microstructure: A survey of microfoundations, empirical results, and policy implications. *Journal of Financial Markets 8*, 217–264.

[4] Brogaard, J., T. Hendershott, and R. Riordan. High frequency trading and price discovery. Working paper http://ssrn.com/abstract=1928510.

[5] Conradt, L. (2011). When it pays to share decisions. *Nature 471*, 40–41.

[6] Couzin, I. D., J. Krause, N. R. Franks, and S. A. Levin (2005). Effective leadership and decision-making in animal groups on the move. *Nature 433*, 513–516.

[7] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance 25*(2), 383–417.

[8] Gerig, A. and D. Michayluk. Automated liquidity provision and the demise of traditional market making. Working paper http://ssrn.com/abstract=1639954.

[9] Glosten, L. R. and P. R. Milgrom (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics 14*, 71–100.

[10] Hayek, F. A. (1945). The use of knowledge in society. *The American Economic Review 4*, 519–530.

[11] Hendershott, T., C. M. Jones, and A. J. Menkveld (2011). Does algorithmic trading improve liquidity? *Journal of Finance 66*, 1–33.

[12] Herbert-Read, J. E., A. Perna, R. P. Mann, T. M. Schaerf, D. J. T. Sumpter, and A. J. W. Ward (2011). Inferring the rules of interaction of shoaling fish. *Proceedings of the National Academy of Sciences of the United States of America 108*(46), 18726–18731.

[13] Kirilenko, A., A. S. Kyle, M. Samadi, and T. Tuzun. The flash crash: The impact of high frequency trading on an electronic market. Working paper http://ssrn.com/abstract=1686004.

[14] Litzenberger, R., J. Castura, and R. Gorelick (2012). The impacts of automation and high frequency trading on market quality. *Annual Review of Financial Economics 35*(4), 59–98.

[15] Mantegna, R. N. (1999). Hierarchical structure in financial markets. *European Physical Journal B 11*, 193–197.

[16] Menkveld, A. High frequency trading and the new-market makers. Working paper http://ssrn.com/abstract=1722924.

[17] Mueller, E. T. (2006). *Commonsense Reasoning*. San Francisco, USA: Morgan Kaufmann.

[18] Sumpter, D. J. T. and S. C. Pratt (2009). Quorum responses and concensus decision making. *Philosophical Transactions of the Royal Society B 364*, 743–753.

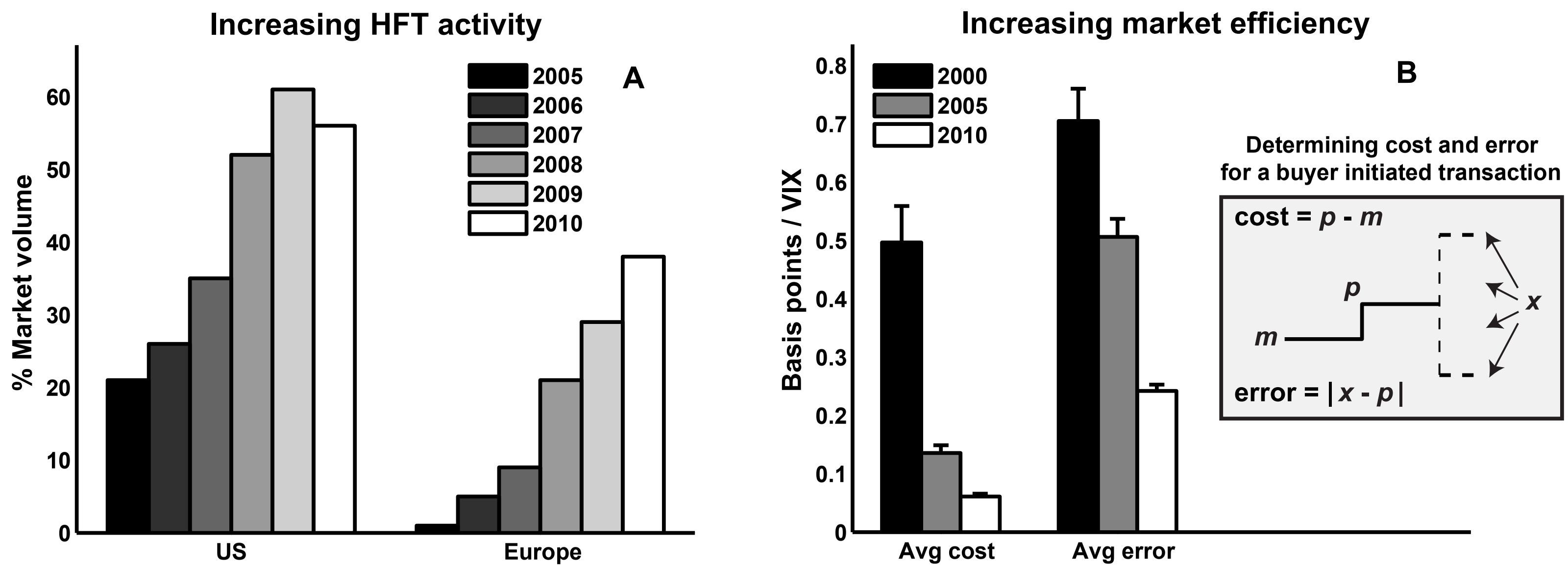

Figure 1: Increase of HFT activity and market efficiency over the last decade. (A) The percentage of HFT volume in US equities (measured in shares) and European equities (measured in value) for 2005 to 2010. Values are estimates from the TABB Group. (B) Average cost and average pricing error of transactions for 35 US stocks during the last full week of February in 2000, 2005, and 2010. The diagram shows how the cost and error of a buyer initiated transaction are calculated (for a seller initiated transaction, the sign of the cost is reversed). To normalize across stocks and time, costs and errors are measured in basis points (1bps=.01%) and divided by the current market volatility as measured by the VIX index. Error bars report the standard error of the mean across the 35 stocks.

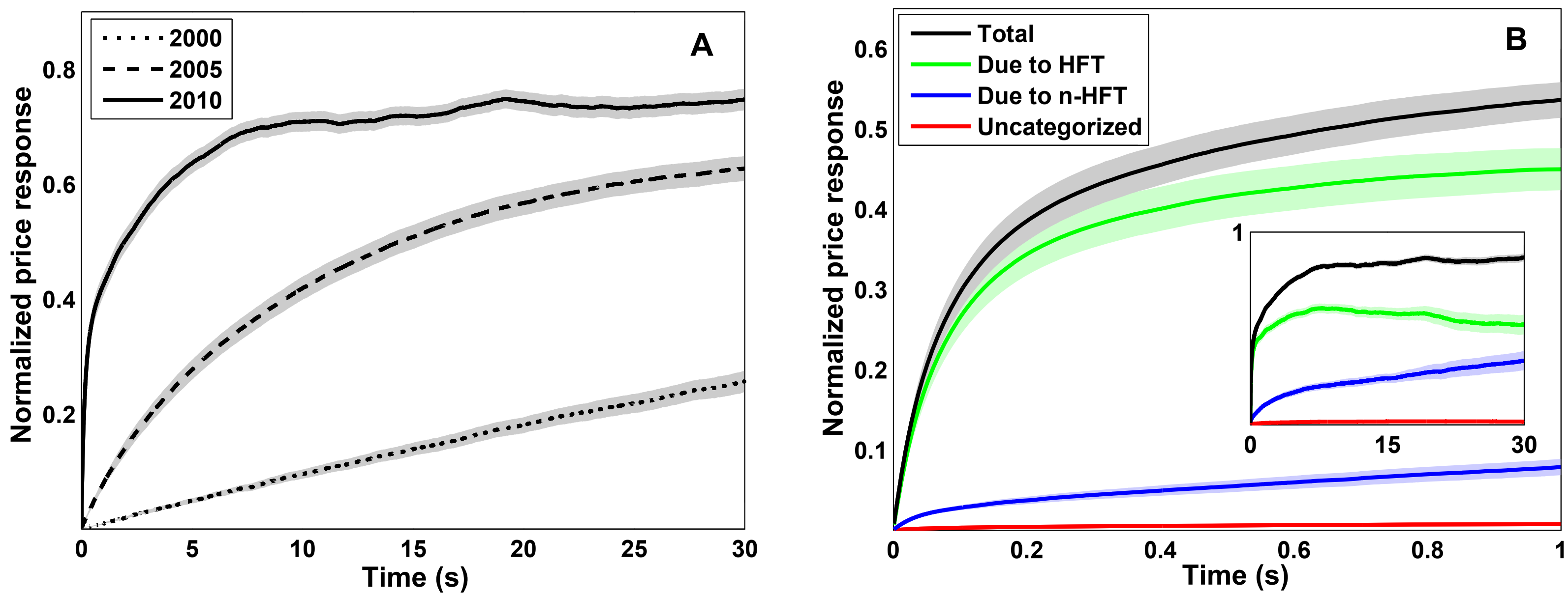

Figure 2: Normalized price response of stock i to a price movement in stock $j = i$. (A) The mean price response for 35 US stocks during the last full week of February in 2000, 2005, and 2010. The standard error of the mean across the 35 stocks is shown in gray. (B) For 2010, the price response of the full 40 US stocks (black) is decomposed into the amount due to HFT activity (green), non-HFT activity (blue), and an amount that could not be categorized (red). Again, standard errors are shown in shaded color.

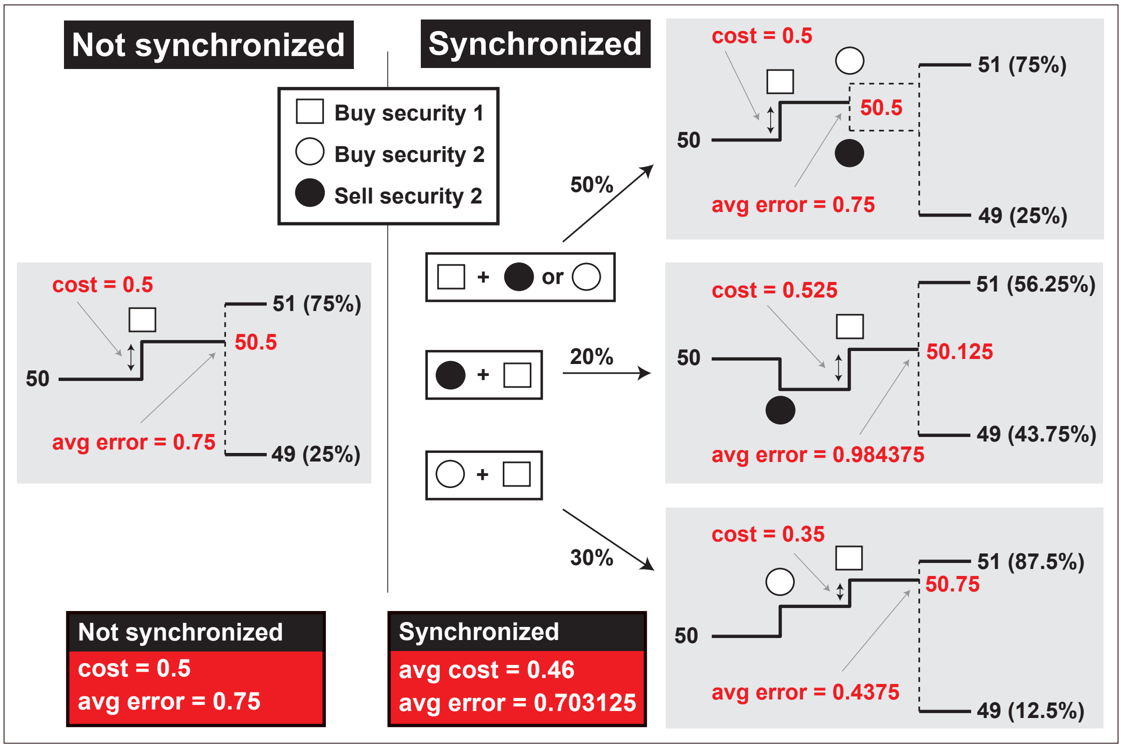

Figure 3: Diagram of the model showing that transaction costs are reduced and prices are more accurate with synchronized prices. Only a buy order for security 1 is analyzed. On the left, the price of security 1 and 2 are not synchronized. The buy order for security 1 transacts at 50.5, giving a transaction cost of 0.5 and an average transaction price error of 0.75. On the right, the prices are synchronized. If the buy order for security 1 arrives before the order for security 2 (top), the analysis is the same as if prices where not synchronized. If a buy or sell order for security 2 arrives first (middle or bottom), this affects the price of security 1 as shown. Transaction costs and errors are calculated for each case and averaged. Final results are shown at the bottom in reverse red highlight.

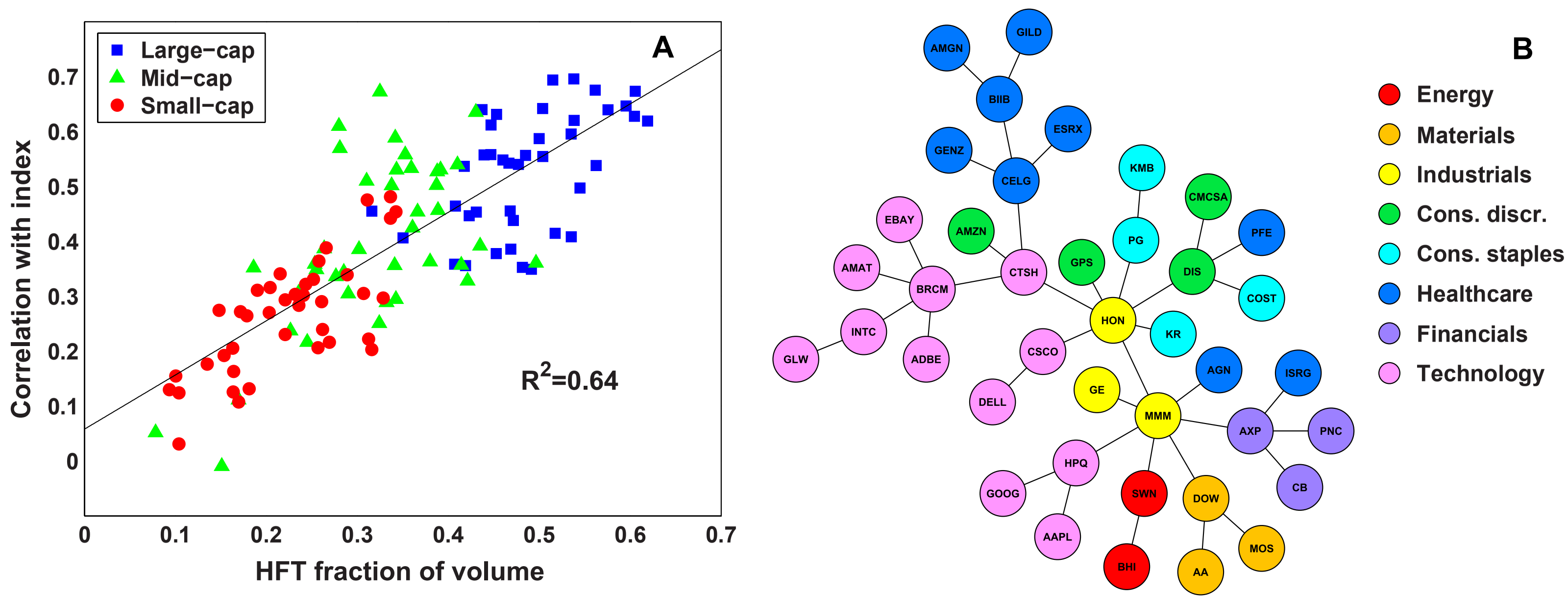

Figure 4: Stocks with higher HFT activity have stronger correlations with other stocks. These correlations correspond with economic structure. (A) Plot of stock correlation vs. the fraction of volume due to HFT for that stock. Correlations are between the 30 second returns of the stock and the equal-weighted average 30 second returns of all 120 stocks. Volume is measured in shares. (B) Minimum spanning tree derived from the 30 second correlation matrix for the 40 large-cap stocks. The ticker for each stock is shown on the corresponding node, and nodes are color-coded according to GICS sector.

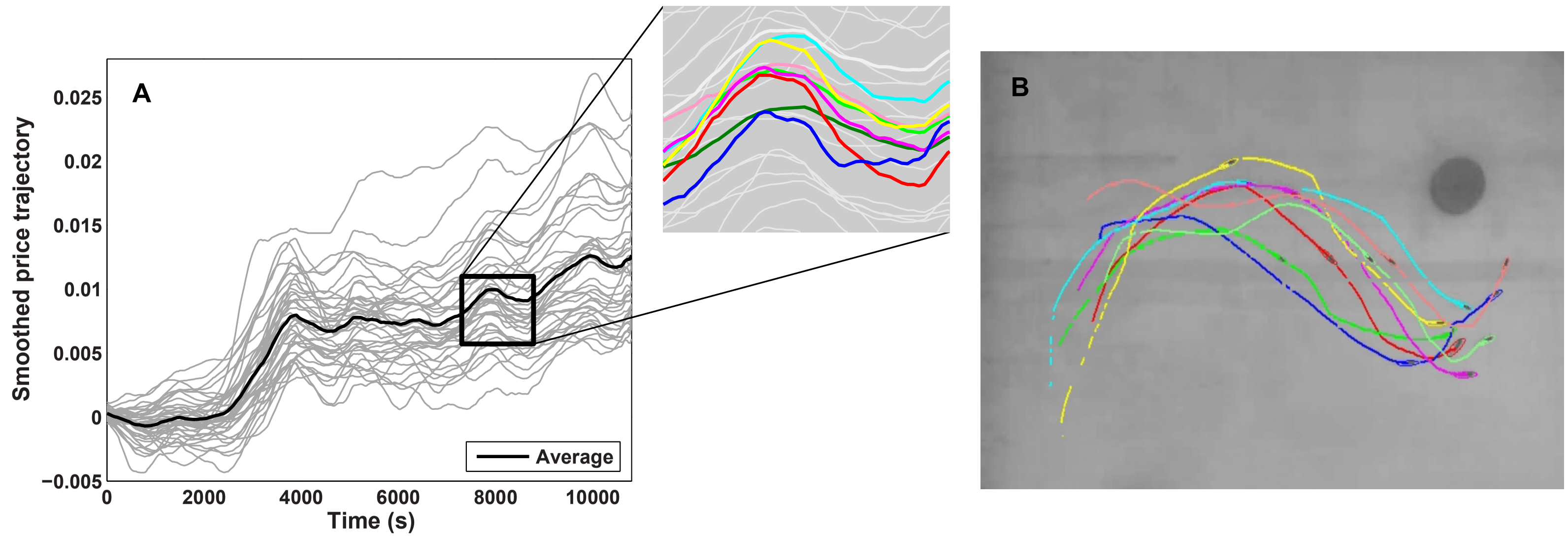

Figure 5: Comparison of the price trajectories of financial securities and the motion of schooling fish. (A) Price trajectories of 40 large-cap US stocks from 1pm to 4pm on February 25, 2010 (from the NASDAQ dataset). The mean trajectory is shown in black. Price trajectories are the change in logarithmic price from 1pm to the current time measured every 30 seconds and are smoothed by taking a moving average with window size 20. The plot is enlarged from 3:02pm to 3:25pm with the following stocks in color: AAPL (red), AMGN (cyan), CMCSA (magenta), HON (yellow), CELG (blue), MOS (dark green), BHI (light green), SWN (pink). (B) Tracked motion of eight schooling mosquitofish. The image consists of two superimposed frames at 15s and 17s of the movie provided as supplemental material in[12].

SUPPLEMENTAL MATERIAL

A Data

The results in the paper are generated using two datasets. The first dataset is provided by NASDAQ and contains all transactions and quotes on the NASDAQ exchange for 120 stocks during the week of February 22-26, 2010. The 120 stocks are a broad representation of the US equity market. Half of the stocks are listed on NASDAQ, half on the NYSE (NYSE listed stocks are also traded on the NASDAQ exchange), and 1/3 are stocks from large companies (large-capitalization or large-cap stocks), 1/3 are medium-cap stocks, and 1/3 are small-cap stocks. Table I contains a list of the stocks. The data is unique for three reasons: (1) the timestamps on the data are precise: they are internally generated by the exchange and are to the millisecond, (2) for every transaction, the initiating order is specified, i.e., I know whether it was the buyer or the seller that caused the transaction, (3) all transactions and quotes are categorized as HFT or non-HFT.

NASDAQ defines the HFT and non-HFT categories as follows: there are 26 firms that specialize in HFT and trade on the NASDAQ exchange; activity that originates from these firms is flagged as HFT, and all other activity is flagged as non-HFT. The 26 firms are primarily independent proprietary trading outfits (although I do not have access to their names, typical examples would include Allston Trading, DRW Holdings, Getco, RGM Advisors, Tradebot, Tradeworx, etc.). The most likely bias in NASDAQ's categorization is under-reporting of HFT. HFT activity that originates from large integrated firms, such as investment banks or large hedge funds, cannot be separated from the other activity of these firms and is therefore not categorized as HFT.

The second dataset is taken from the Thomson Reuters Tick History database (TRTH), which includes records of transactions and quotes from numerous financial markets around the world. From the Reuters database, I collect trade and quote information for the 40 large-cap stocks that are included in the NASDAQ data. I choose the same time period as the NASDAQ data (February 22-26, 2010) but also include data from earlier time periods: February 21-25, 2000 and February 21-25, 2005. 5 of the 40 stocks are not included in the Reuters database during all three periods (GOOG, HPQ, MOS, BIIB, and ISRG), so I disregard the Reuters data for these stocks in all time periods. Furthermore, US markets were closed on February 21 in 2000 and 2005 (both were Mondays), so I disregard the Reuters data for Monday in 2010 as well. The main differences between the Reuters data and the NASDAQ data are the following: (1) the NASDAQ data flags HFT, (2) both the Reuters and NASDAQ data are timestamped to the nearest millisecond, but the Reuters times are not as precise due to delays in transmitting information from exchanges to Reuters, (3) the Reuters data does not specify whether the buyer or seller initiated the transaction, (4) the Reuters data does not include transactions that are less than 100 shares, (5) the Reuters data includes trades and quotes from all major US exchanges, whereas the NASDAQ data includes information only from the NASDAQ exchange.

B Methods

To generate the results in Figs. 1(B) and 2, transactions must be classified as buyer or seller initiated. Unfortunately, the Reuters data does not include this information. I estimate the initiator of a transaction for the Reuters data as follows: for each transaction, I determine whether the transaction price is closer to the bid or the ask price. The bid price is the price at which you can immediately sell in the market and the ask price is the price at which you can immediately buy. If closer to the bid, I assume it is seller initiated. If closer to the ask, I assume it is buyer initiated. If the transaction price is at the midpoint of the bid and ask price, I leave the transaction unclassified and ignore it when generating the figures.

In Fig. 1(A), HFT estimates are from the TABB group as reported in "High-frequency trading: Up against a bandsaw", Sept. 2, 2010, *Financial Times*. It is difficult to ascertain how the TABB group calculates these estimates (They did not respond when contacted for clarification). In the NASDAQ data, 49% of share volume is attributable to HFT, which is close to the estimate of 56% in 2010 by TABB. Also, the Aite Group provides similar estimates for HFT ("The fast and the furious"', Feb. 25, 2012, *The Economist*), and both estimates correspond well with the increase in message traffic used as a proxy for algorithmic trading in [11].

In Fig. 1(B), the cost of a buyer initiated transaction is measured as the transaction price, p, minus the current prevailing midpoint price for the security, m, (for a seller initiated transaction, the cost is $m-p$). The midpoint price is the midpoint between the quoted price at which you can immediately buy (called the ask) and the quoted price at which you can immediately sell (called the bid) in the market. The transaction error is measured as the absolute difference between the transaction price, p, and the midpoint price 1 minute later, x, (see the diagram in the figure). To standardize across securities, costs and errors are measured in basis points (1bps=0.01%) where the cost is divided by m and the error by p. To standardize across time periods with different volatilities, costs and errors are divided by the average value of the VIX for the time period (obtained from http://www.cboe.com/micro/vix/historical.aspx). The figure reports the mean and standard error of the mean for the 35 stocks during each time period.

The curves in Fig. 2(A) and 2(B) are calculated as follows: the average price response of stock i, conditioned on a price movement in j at time 0, is first determined for all $j = i$ (for negative price movements of j, the sign of the response is reversed). Price movements are defined as any change in the midpoint prices. Each response curve is normalized by dividing the response by the difference between its maximum and minimum value. The normalized curves are averaged over all j for each i, and then the final curve is generated by averaging over all stocks i, with the standard error of the mean shown in shaded color.

In Fig. 2(B), I separate out the price response into an amount due to HFT, an amount due to non-HFT, and an amount that was uncategorized. To perform this separation, I determine the cause of each midpoint price change for each security. If the midpoint for a stock increases (decreases) during the same millisecond that a buyer (seller) initiated transaction occurs in that stock, I assume the transaction caused the price change. If the initiator of the transaction was a HFT firm, I assign the midpoint change to HFT. Likewise, if the initiator was non-HFT, I assign the midpoint change to non-HFT. If the midpoint changes without a corresponding transaction, then I first determine whether the bid and/or ask was improved or removed. If improved (i.e., the ask lowered or the bid increased), then I assign the price change to the initiator of the new quote (HFT or non-HFT). If removed (i.e., the ask is increased or the bid is lowered), then I assign the price change to the initiator of the original quote (HFT or non-HFT). I am unable to categorize several instances which are rare in the data: (1) if multiple buyer or seller initiated transactions occur during the same millisecond that the midpoint changed, and if these transactions are mixed between HFT and non-HFT, then I cannot ascertain whether it was HFT or non-HFT that caused the price change, (2) if both the bid and the ask change during the same millisecond, but one change is due to HFT and the other to non-HFT, I cannot ascertain whether the cause was HFT or non-HFT.

I use the NASDAQ data to create the plots in Fig. 4. In Fig. 4(B), the weight between nodes i and j in the full network is $\sqrt{2(1-\rho_{i,j})}$ where $\rho_{i,j}$ is the correlation between 30 second returns for stock i and stock j. From this network, I determine the minimum spanning tree using Prim's algorithm. GICS sectors are taken from the CRSP (Center for Research in Security Prices) database.

C *Discussion of Model*

In the model, liquidity providers are market participants who facilitate trade by transacting with investors' orders. By comparing non-synchronous to synchronous pricing, the model implicitly assumes that liquidity providers were previously unable to synchronize prices, but that HFT (now acting as the de facto liquidity providers) can easily perform this task. Of course, before HFT existed, human liquidity providers would have done their best to keep prices aligned. The point is that computers are much better than humans at performing this task.

In the model, liquidity providers are assumed perfectly competitive so that they set fair prices and make zero profit. In real markets, liquidity providers will require a price concession to transact with an investor's order. Adding a small price concession that allows for liquidity provider profits does not change the results.

In real markets, investors can place aggressive orders that transact immediately at the best available price (called market or marketable orders) or passive orders that specify a price but are not guaranteed to transact (called limit orders). The model does not specifically account for these different order types, but could be adjusted to include either or both. If an investor's order provides a large enough price concession, then regardless of type, liquidity providers will transact with the order. This means the term "liquidity provider" should not be restricted to firms that only transact using limit orders. In fact, in the NASDAQ dataset, HFT firms transact almost equally using market and limit orders, but they can provide liquidity in either case. For example, suppose that identical securities are traded on two different exchanges and that these exchanges have crossing limit orders. An arbitrageur can connect the orders by placing offseting marketable orders in the two markets, which in effect, provides liquidity to both. This example can easily be generalized to two economically related rather than identical securities, and it shows that liquidity providers connect investors not only through time but also across exchanges and securities.

D Calculations

D.1 Cost and pricing error

As in the main text, assume that the market contains only two securities, $n = 2$, and the initial price of each security is $m_1 = m_2 = 50$ which can increase or decrease by $\delta = 1$. For both securities $\phi_1 = \phi_2 = 0.75$, and their price changes are correlated with $\rho_{1,2} = 0.8$.

I will analyze the transaction cost, $c(B_1) = E[p_1 - m_1|B_1]$, and pricing error, $e(B_1) = E[|x_1 - p_1||B_1]$, for a buy order in security 1, B_1, when prices are and are not synchronized (for a sell order, the sign of the cost is reversed). First, if prices are not synchronized, then the buy order transacts at the expected price of the security conditioned on the placement of the buy order

$$
\begin{aligned}
p_1 &= E[x_1|B_1], \\
&= 49\,\mathcal{P}(x_1 = 49|B_1) + 51\,\mathcal{P}(x_1 = 51|B_1), \\
&= 50.5,
\end{aligned}
$$

which is calculated by applying Bayes' Rule, $\mathcal{P}(x_1|B_1) = \mathcal{P}(B_1|x_1)\mathcal{P}(x_1)/\mathcal{P}(B_1)$,

$$
\begin{aligned}
\mathcal{P}(x_1 = 49|B_1) &= 0.25 \\
\mathcal{P}(x_1 = 51|B_1) &= 0.75.
\end{aligned}
$$

The cost is therefore,

$$c(B_1) = p_1 - m_1 = 0.5,$$

and the average pricing error is,

$$
\begin{aligned}
e(B_1) &= (51 - p_1)\mathcal{P}(x_1 = 51|B_1) + (p_1 - 49)\mathcal{P}(x_1 = 49|B_1), \\
&= 0.75.
\end{aligned}
$$

If the market is synchronized, then the transaction price can be one of three values. If the buy order for security 1 arrives before the order for security 2, which occurs 50% of the time, then the cost and pricing error are the same as if prices were not synchronized, 0.5 and 0.75 respectively.

When a sell order for security 2 arrives first, which occurs $\mathcal{P}(S_2|B_1)/2 = 20\%$ of the time,

$$\begin{aligned}
\mathcal{P}(S_2|B_1) &= \frac{\mathcal{P}(S_2, B_1)}{\mathcal{P}(B_1)}, \\
&= \frac{1}{\mathcal{P}(B_1)} \sum_{x_1, x_2} \mathcal{P}(S_2, B_1|x_1, x_2)\mathcal{P}(x_1, x_2), \\
&= \frac{1}{\mathcal{P}(B_1)} \sum_{x_1, x_2} \mathcal{P}(S_2|x_2)\mathcal{P}(B_1|x_1)\mathcal{P}(x_1, x_2), \\
&= 0.40.
\end{aligned}$$

then the transaction price of the buy order is,

$$\begin{aligned}
p_1 &= E[x_1|B_1, S_2], \\
&= x_1^+ \, \mathcal{P}(x_1^+|B_1, S_2) + x_1^- \, \mathcal{P}(x_1^-|B_1, S_2), \\
&= 51 \times 0.5625 + 49 \times 0.4375, \\
&= 50.125
\end{aligned}$$

where again, Bayes' Rule is used, $\mathcal{P}(x_1|B_1, S_2) = \mathcal{P}(B_1, S_2|x_1)\mathcal{P}(x_1)/\mathcal{P}(B_1, S_2)$,

$$\begin{aligned}
\mathcal{P}(x_1 = 49|B_1, S_2) &= 0.4375 \\
\mathcal{P}(x_1 = 51|B_1, S_2) &= 0.5625.
\end{aligned}$$

The cost is therefore,

$$c(B_1) = p_1 - m_1 = 0.525,$$

and the average pricing error is,

$$\begin{aligned}
e(B_1) &= (51 - p_1)\mathcal{P}(x_1 = 51|B_1, S_2) + (p_1 - 49)\mathcal{P}(x_1 = 49|B_1, S_2), \\
&= 0.984375.
\end{aligned}$$

When a buy order for security 2 arrives first (the remaining 30% of the time),

$$
\begin{aligned}
\mathcal{P}(B_2|B_1) &= \frac{\mathcal{P}(B_2, B_1)}{\mathcal{P}(B_1)}, \\
&= \frac{1}{\mathcal{P}(B_1)} \mathop{}_{x_1,x_2} \mathcal{P}(B_2, B_1|x_1, x_2)\mathcal{P}(x_1, x_2), \\
&= \frac{1}{\mathcal{P}(B_1)} \mathop{}_{x_1,x_2} \mathcal{P}(B_2|x_2)\mathcal{P}(B_1|x_1)\mathcal{P}(x_1, x_2), \\
&= 0.60.
\end{aligned}
$$

similar calculations give $c(B_1) = 0.35$ and $e(B_1) = 0.4375$. Averaging over these three possibilities, the overall expected transaction cost and average pricing error are,

$$
\begin{aligned}
c(B_1) &= 0.5 \times 0.5 + 0.525 \times 0.2 + 0.35 \times 0.3 = 0.46, \\
e(B_1) &= 0.75 \times 0.5 + 0.984375 \times 0.2 + 0.4375 \times 0.3 = 0.703125.
\end{aligned}
$$

D.2 Profit of Informed Traders

In the original Glosten and Milgrom model[9], investors are separated into two groups "informed investors" and "liquidity traders". The former have complete knowledge of the end-of-period price for the security they trade, and they submit buy orders when it is higher and sell orders when it is lower. The latter have no knowledge of the final price and buy or sell randomly (in the paper, I call these investors "average investors"). If the fraction of orders for security i from informed investors is γ_i, then $\mathcal{P}(B_i|x_i^+) = \phi_i = 1/2(1 + \gamma_i)$. In the example given in the paper, $\phi_i = 0.75$ so that $\gamma_i = 1/2$, meaning 1/2 of all orders come from informed investors.

If prices are not synchronized, informed investors make an expected profit of,

$$
E[x_1 - p_1|x_1^+] = 0.5,
$$

per buy transaction (and also an expected profit of 0.5 per sell transaction). Average investors make an expected profit of,

$$0.5E[x_1 - p_1|x_1^+] + 0.5E[x_1 - p_1|x_1^-] = -0.5,$$

per buy transaction (and the same per sell transaction). Notice that the total amount lost by the average investors is gained by the informed investors; a well-known result of the Glosten and Milgrom model.

When prices are synchronized, the informed make an expected profit of,

$$E[x_1 - p_1|x_1^+] = 51 - E[p_1|x_1^+].$$

If the buy order for security 1 arrives first, which occurs 50% of the time, then $E[p_1|x_1^+] = 50.5$. When a sell order in security 2 arrives first, which occurs $\mathcal{P}(S_2|B_1, x_1^+)/2 = 15\%$ of the time,

$$\begin{aligned}
\mathcal{P}(S_2|B_1, x_1^+) &= \frac{\mathcal{P}(S_2, B_1, x_1^+)}{\mathcal{P}(B_1, x_1^+)}, \\
&= \frac{1}{\mathcal{P}(B_1, x_1^+)} \sum_{x_2} \mathcal{P}(S_2, B_1|x_1^+, x_2)\mathcal{P}(x_1^+, x_2), \\
&= \frac{1}{\mathcal{P}(B_1, x_1^+)} \sum_{x_2} \mathcal{P}(S_2|x_2)\mathcal{P}(B_1|x_1^+)\mathcal{P}(x_1^+, x_2), \\
&= 0.30,
\end{aligned}$$

then $E[p_1|B_1, S_2, x_1^+] = 50.125$. When a buy order in security 2 arrives first, which occurs the remaining $\mathcal{P}(B_2|B_1, x_1^+)/2 = 35\%$ of the time,

$$\begin{aligned}
\mathcal{P}(B_2|B_1, x_1^+) &= \frac{\mathcal{P}(B_2, B_1, x_1^+)}{\mathcal{P}(B_1, x_1^+)}, \\
&= \frac{1}{\mathcal{P}(B_1, x_1^+)} \sum_{x_2} \mathcal{P}(B_2, B_1|x_1^+, x_2)\mathcal{P}(x_1^+, x_2), \\
&= \frac{1}{\mathcal{P}(B_1, x_1^+)} \sum_{x_2} \mathcal{P}(B_2|x_2)\mathcal{P}(B_1|x_1^+)\mathcal{P}(x_1^+, x_2), \\
&= 0.70,
\end{aligned}$$

then $E[p_1|B_1, B_2, x_1^+] = 50.75$. Putting it all together, the expected profit per buy transaction for an informed investor when prices are synchronized is,

$$
\begin{aligned}
E[x_1 - p_1|x_1^+] &= 51 - (0.5 \times 50.5 + 0.15 \times 50.125 + 0.35 \times 50.75) \\
&= 0.46875.
\end{aligned}
$$

Because of the symmetry of the example, the expected profit per sell transaction for informed investors is the same (as well as the expected profit per transaction in security 2). The informed therefore make less profits (0.46875 per transaction vs. 0.5 per transaction) when prices are synchronized.

E Proofs

Consider a market with n securities as described in the main text. For the proofs below, I will consider a buy order in security i. The same results hold if considering a sell order. If prices are not synchronized, then the transaction cost, $c(\cdot)$, of a buy order for security i is the following:

$$
c(B_i) = E[x_i|B_i] - E[x_i], \tag{1}
$$

and the average pricing error, $e(\cdot)$, of a buy order for security i is,

$$
\begin{aligned}
e(B_i) &= E\left[|x_i - E[x_i|B_i]||B_i\right], &(2)\\
&= E\left[s_i(x_i - E[x_i|B_i])|B_i\right], &(3)\\
&= E[s_i x_i|B_i] - E[s_i E[x_i|B_i]|B_i], &(4)
\end{aligned}
$$

where $s_i = +1$ if $x_i = x_i^+$ and $s_i = -1$ if $x_i = x_i^-$.

If prices are synchronized, then individual transactions cause price updates in all securities as they occur. Therefore, the transaction cost and pricing error of an order depend on the set of transactions that have occurred prior to the order's arrival.

The expected transaction cost, $c'(\cdot)$, of a buy order for security i when prices are synchronized is the following:

$$
\begin{aligned}
c'(B_i) &= E\left[\left(E[x_i|\omega, B_i] - E[x_i|\omega]\right)|B_i\right], && (5)\\
&= E[x_i|B_i] - E\left[E[x_i|\omega]|B_i\right], && (6)
\end{aligned}
$$

where ω is the set of transactions that occur before B_i and $\omega \in \Omega$ where Ω is the set of all sets of transactions in securities $j = i$ that can occur before B_i. The average pricing error, $e'(\cdot)$, of a buy order for security i when prices are synchronized is,

$$
\begin{aligned}
e'(B_i) &= E\left[|x_i - E[x_i|\omega, B_i]||B_i\right], && (7)\\
&= E\left[s_i(x_i - E[x_i|\omega, B_i])|B_i\right], && (8)\\
&= E[s_i x_i|B_i] - E[s_i E[x_i|\omega, B_i]|B_i]. && (9)
\end{aligned}
$$

The difference in transaction costs when prices are synchronized is,

$$
\begin{aligned}
c(B_i) - c'(B_i) &= E\left[E[x_i|\omega]|B_i\right] - E[x_i], && (10)\\
&= E\left[E[x_i|\omega] - E[x_i]|B_i\right], && (11)\\
&= E[2\delta_i(\mathcal{P}(x_i^+|\omega) - \mathcal{P}(x_i^+))|B_i], && (12)\\
&= 4\delta_i\left(E[\mathcal{P}(x_i^+|\omega)\mathcal{P}(B_i|\omega)] - \frac{1}{4}\right), && (13)\\
&= 4\delta_i cov[\mathcal{P}(x_i^+|\omega), \mathcal{P}(B_i|\omega)]. && (14)
\end{aligned}
$$

Because the covariance is positive, transaction costs are lower when prices are synchronized.

The difference in average pricing error when prices are synchronized is:

$$
e(B_i) - e'(B_i) = -E[s_i E[x_i|B_i]|B_i] + E[s_i E[x_i|\omega, B_i]|B_i], \qquad (15)
$$

which after some algebra is,

$$e(B_i) - e'(B_i) = 4\delta_i \mathcal{P}^2(x_i^+|B_i) \quad E\left[\frac{\mathcal{P}(\omega|x_i^+)}{\mathcal{P}(\omega|B_i)} \middle| \, x_i^+ \right] - 1 \tag{16}$$

Because the expectation of $\mathcal{P}(\omega|x_i^+)/\mathcal{P}(\omega|B_i)$ is larger than 1, the average pricing error is lower when prices are synchronized.

Large-cap				Medium-cap				Small-cap			
AA	AAPL	ADBE	AGN	AINV	AMED	ARCC	AYI	ABD	ANGO	APOG	AZZ
AMAT	AMGN	AMZN	AXP	BARE	BRE	BXS	CBT	BAS	BW	BZ	CBEY
BHI	BIIB	BRCM	CB	CETV	CHTT	CKH	CNQR	CBZ	CCO	CDR	CPSI
CELG	CMCSA	COST	CSCO	COO	CPWR	CR	CRI	CRVL	CTRN	DCOM	DK
CTSH	DELL	DIS	DOW	CSE	CSL	ERIE	EWBC	EBF	FFIC	FPO	FRED
EBAY	ESRX	GE	GENZ	FCN	FL	FMER	FULT	IMGN	IPAR	KNOL	KTII
GILD	GLW	GOOG	GPS	GAS	ISIL	JKHY	LANC	MAKO	MDCO	MFB	MIG
HON	HPQ	INTC	ISRG	LECO	LPNT	LSTR	MANT	MOD	MRTN	MXWL	NC
KMB	KR	MMM	MOS	MELI	NSR	NUS	PNY	NXTM	PBH	PPD	RIGL
PFE	PG	PNC	SWN	PTP	ROC	SF	SFG	ROCK	ROG	RVI	SJW

Table 1: Table of stocks in NASDAQ HFT dataset separated by market capitalization.