Price Discovery and Liquidity Recovery: Forex Market Reactions to Macro Announcements

By MASAHIRO YAMADA AND TAKATOSHI ITO*

This version: August 2019, VERY PRELIMINARY

Abstract. We examine the improvement of Forex market quality, measured by the speed of price discovery and liquidity recovery after macro announcements, with 20 years of high-frequency data. Over the two decades the microstructure of the forex market dramatically changed. Machines have been taking over, first as an electronic order-matching broker and then as sophisticated bank dealers with extremely fast price-making and deal execution. A popular conjecture is that price discovery after new public information became faster and liquidity recovers quicker than before. Our findings are mixed: the speed of price discovery has improved but liquidity recovery has been sluggish. Among the determinants we pin down, we find that an increase in traders after announcements improves liquidity but slows down price discovery. This implies that "fast" traders have a poor interpretation of the news for predicting its influences on prices.

* Masahiro Yamada: Graduate School of Economics, Osaka University, m.yamada@econ.osaka-u.ac.jp,
Takatoshi Ito: School of International and Public Affairs, Columbia University, tito@columbia.edu.

1 Introduction

The foreign exchange market has transformed from human-dominating market to machine-dominating market. Computers loaded with algorithm receive and process information and send limit orders and market orders to the order-matching machine, operated by EBS, Reuters or dark pools. The order-matching machines process orders with a first-come-first-served basis. Bank computers are collocated with the order-matching machine to minimize the loss in transmission time. A whole process of banks' order to a deal is completed within a fraction of one second.

In the world of high-frequency traders, that is machines armed with algorithm, new information should be digested into the prices much quicker than before. Although it is a near consensus that price discovery is faster, it has been debated whether high-frequency traders provide or consume liquidity.[1]

New information can be a macroeconomic news announcement at scheduled time, an unexpected result of known events like an election or referendum, or some economic and political development at unexpected timing. Let $\tilde{v}$ denote the true Forex reaction to the news. Upon arrival of news y, the surprise component of new information $y - \hat{y}$ is digested in market price through trades among market participants, achieving $\tilde{v} = y - \hat{y}$. Some of market participants are informed and others are uninformed. That is, informed traders have the correct expectation $\hat{y}$ which makes the prediction of $\tilde{v}$. Not knowing $\hat{y}$, uninformed traders try to learn correct contents and interpretation of news through trading and observing price and volume movements.

The price level moves to a new equilibrium level after new information arrives. This is price discovery. When traders do not share the same information and interpretation, trading needs in the process of price discovery. It is expected that the volatility becomes higher upon arrival of news and remains high for some time span. That is a phenomenon called volatility clustering.

With tick-by-tick data, the price discovery process can be examined very closely. The impact of widespread use of machines can be analyzed better with tick-by-tick data. Interesting and important questions are whether widespread high-frequency traders have made the price discovery process faster; and whether they provide more liquidity so that volatility is less.

In case of macroeconomic statistics announcement at scheduled time, liquidity becomes less toward the announcement time as traders would like to stay away in the strong presence of uncertainty. After the macro statistics number is announced and news contents are digested, the liquidity is recovered. The liquidity recovery is a sign that the new equilibrium is found.

To answer these questions, we use 20 years of high-frequency data collected in the ICAP EBS order-matching system, which makes it possible to examine long-term changes in the speed of price discovery and liquidity recovery. We focus on price and liquidity reactions to macro announcements, which offers a natural experiment on how a market incorporates information into prices.

---

[1] Many academic research claim that algorithmic high-frequency traders helps price discovery and liquidity provision in regular trading hours (e.g., Brogaard, Hendershott, and Riordan (2014), Chordia, Green, and Kottimukkalur (2018)). In situations of extremely volatile market such as a Flash crash, high-frequency traders can consume rather than provide liquidity (Kirilenko et al. (2017)). Chordia, Green, and Kottimukkalur (2018) investigate stock markets and report that price discovery after macro announcements has become faster in these years. Scholtus, Van Dijk, and Frijns (2014) investigated the U.S. equity market and stress that the speed of trading is important for profitability of trading based on news. In line with the finding on market quality by Scholtus, Van Dijk, and Frijns (2014) , Jiang, Lo, and Valente (2012) reports that high-frequency trading lowers depth on the order book during the post-announcement period but increases price efficiency through the trade.

Our main empirical findings regarding these measures are summarized as follows. After macro announcements,

i. the speed of price discovery looks stagnated in recent years. After considering determinants of price discovery, however, we can estimate the improvement by 30 percentage points from 2006 to 2017. Informativeness of the news is a key driver, but it depends on macro-financial situations.
ii. The increase in the number of market participants does not necessarily facilitate faster price discovery, but it significantly improves the liquidity.

Figure 1 describes the time-series of the speed of price discovery and liquidity recovery for USD/JPY.

[Figure 1 is inserted here]

In this empirical examination, we define speed of price discovery as a variance ratio, i.e., denoting u as the time of news arrival, the variance of return in $[u,u+k]$ over the variance of return in $[u,u+T]$, where k and T are free parameters satisfying $k \le T$. Chordia, Green, and Kottimukkalur (2018) employ a similar measure and report that the price discovery becomes faster in recent years in the U.S. equity market. Our results, however, such improvement is much weaker in the Forex market; the speed of price discovery is almost flat over years. Liquidity is measured by an effective bid-ask spread. Again, we do not observe improvements in the speed of liquidity recovery over years. In fact, the liquidity recovery is fast enough even in 1999, and the room of improvements was small.

Then, what explains the variation of price discovery and liquidity? We set two key variables that are theoretically relevant and empirically available: explanatory power of the news for predicting after-the-news returns after the announcement, and the number of traders who submit limit orders. The first variable, we call it R-squared measure, is obtained by a R-squared when the after-the-news return is regressed on the surprise of the announced macro statistics.[2] This measure is a proxy of the precision of information held by traders. The second variable is "quote counts", or the number of traders who hold limit orders, in our high-frequency dataset. Our regression analysis shows that the price discovery is facilitated as the R-squared measure grows but slowed down as quote counts grows. Conversely, the liquidity recovery is negatively affected by the R-squared measure but improved by the quote counts. Figure 2 graphically describes these main empirical findings.

[Figure 2 is inserted here]

[2] For defining the surprise of news, we use pre-expectations of macro statistics provided by Bloomberg.

Theoretically, our empirical findings coincide with a situation when the market participants behave like an uninformed trader, which is an implication of classical noisy rational expectation models (e.g., Grossman and Stiglitz (1980)). When traders behave like an uninformed trader, they can provide liquidity but cannot add any information into the market price.[3] In reality, however, a trading after macro announcements may be more like an informed trading, as far as market participants trade based on event-period private information (Kim and Verrecchia (1997)). Each trader is endowed with a private expectation with errors $\tilde{y}_i = \hat{y} + \tilde{e}_i$, and makes a prediction for $\tilde{v} = y - \hat{y}$. Trader's own surprise on news, $y - \tilde{y}_i = \tilde{v} - \tilde{e}_i$, is related with $\tilde{v}$, and is considered as private information. As more traders reveal their information, the prediction for $\tilde{v}$ becomes more accurate. Thus, the problem on public information is transformed into a problem about standard microstructure model of private information.

For motivating and explaining our empirical analysis, we draw theoretical implications from Kyle (1989), which provides a general microstructure framework. The model is regarded general, for (1) the model is inhabited with N risk-averse informed, M risk-averse uninformed, and random noise traders, and (2) each informed trader has their own private information, and other informed and uninformed traders extract signals from equilibrium prices, i.e., "non-nested" information structure, and (3) both informed and uninformed traders have market power, and (4) it is a model for a quote-driven market, the same as the structure of EBS interbank market.

We fit the baseline model to the reality of trading after news announcements. As for (1), we consider a market with only informed traders. Because of the property (2), an informed trader has both aspects of informed and uninformed trader, which is useful to calibrate the model. We also assume that interpretation of news by each trader is very noisy, which is, in Kyle (1989), studied as an extension of the baseline model.

Our empirical measures of market quality, variance ratio and liquidity, can be formulated in this theoretical model to pin down their connection with underlying parameters. Under a reasonable calibration, the model can reproduce the empirical findings: informed traders tend to behave like uninformed traders, and the increasing number of informed traders does not facilitate price discovery. The impact of news improves the price discovery but reduces liquidity, because the asymmetry of information becomes severe.

[3] In contrast, Anand, Tanggaard, and Weaver (2009) investigated data on Stockholm Stock Exchange and reports the liquidity provision leads to an improvement in price discovery. Scholtus, Van Dijk, and Frijns (2014) shows the speed of HFTs is an important determinant of the profitability of their news-driven trading strategy, and the activity of HFTs positively affect the market quality defined as depth, bid-ask spread, and price resiliency.

Consistent with the theoretical prediction, our empirical evidence shows that the predictive power of news for after-the-news returns improves the price discovery but delays the liquidity recovery. The predictive power depends on macro-financial environment (Fatum, Hutchison, and Wu (2012)) rather than showing a monotonic improvement. The increase of traders does not facilitate the price discovery, suggesting that such "fast" traders do not have precise interpretation that relates the news statistics to Forex prices.

Considering these determinants and reasonable controls, a time-trend can be obtained. We find that the market quality in the Forex market is mixed: price discovery becomes faster, but liquidity recovery becomes slower. These findings is not only applicable to the specific currency pair, but is a market-wide phenomenon.

Outline of this paper is as follows. Section 2 provides a theoretical model that suggests the empirical measures of price discovery and liquidity as well as proxies for exogenous parameters. Section 3 explains data and defines the empirical measures and their overview, and then we show regression analyses to find determinates of market quality and any exogenous time-trend of improvements. Section 4 wrap up related empirical analyses for checking robustness. We also demonstrate numerical experiments on the model by feeding parameters that match to the reality. Section 5 concludes.

2 A Stylized Model

In this section, we present a model that supports our empirical analyses. Based on classical Kyle (1989) model, we consider the environment where the amount of private information decreases as the number of informed traders becomes large. In this situation, despite that each informed trader has exclusive private information, their marginal contribution to the price informativeness can be little; they behave like an uninformed trader.

Our model features one trading stage, one risky and risk-free asset, and N informed and random noise traders. At $t = 0$, informed trader k is endowed with initial belief about the news statistics $\tilde{y}_k = \hat{y} + \tilde{e}_k,\ k = 1,2,\cdots,N$ as well as a public announcement y. The after-the-news return $\tilde{v}$ is realized in a form of $\tilde{v} = y - \hat{y}$, i.e., it depends on the "true" surprise of announcement. Thus, trader k can generate $\tilde{\iota}_k \equiv y - \tilde{y}_k = \tilde{v} + \tilde{e}_k$ for predicting $\tilde{v}$. Both informed and noise traders trade the risky asset at $t = 1$, and the asset value is realized at $t = 2$. Following Kim and Verrecchia (1994), $\tilde{e}_k,\ k = 1,2,\cdots,N$ is regarded as an interpretation error for each trader. We assume the normality of random variables: $\tilde{v} \sim \mathrm{N}(0, \tau_v^{-1})$ and $\tilde{e}_k \sim \mathrm{N}(0, \tau_e^{-1})$ which are mutually independent. The risk-free asset earns zero interest.

Informed traders have a negative exponential utility function, and they maximize a terminal wealth by choosing an order submission strategy. Thus, informed trader's profit maximization problem is

$$\max_{x_k} \mathrm{E}[u\left((\tilde{v} - p(x_k))x_k\right)|i_k, p] = \max_{x_k} \mathrm{E}[(\tilde{v} - p(x_k))x_k|i_k, p] - \frac{\rho}{2}\mathrm{Var}[(\tilde{v} - p(x_k))x_k|i_k, p].$$

At $t = 1$, informed traders submit their demand/supply functions to an auctioneer and simultaneously observe an aggregate demand/supply curve $p(x_k)$. Noise traders randomly submit market orders $\tilde{z} \sim \mathrm{N}(0, \sigma_z^2)$, which is inelastic to the price and independent of $\tilde{v}$ and $\tilde{u}_k$. Price at $t = 1$ clears the demand/supply of informed and noise traders; i.e., $\sum_k x_k(p) + \tilde{z} = 0$ holds.

This model involves signal extraction and imperfect competition. From the assumptions on information structure, each informed trader receives different news signal; that is, the information is not nested, which enables informed traders to extract additional information from the aggregate demand/supply curve. Namely, $\mathrm{E}(\tilde{v}|i_k, p) \neq \mathrm{E}(\tilde{v}|i_k)$ and $\mathrm{Var}(\tilde{v}|i_k, p) \leq \mathrm{Var}(\tilde{v}|i_k)$. Different from the competitive model of Grossman and Stiglitz (1980), informed traders utilize their monopoly power inherited from their informational advantage. Therefore, compared to the competitive case, traders are less aggressive for informed trading, and incorporation of information into prices can delay and the liquidity can decrease.

In the original model of Kyle (1989), there are uninformed traders who do not receive private signals but optimally characterize their demand/supply function. In our model, they are muted for making the characterization of equilibrium easier. In fact, it is reasonable to assume there are only informed traders after an announcement, because each trader observes a public news and trade based on the interpretation of the news. Even without uninformed traders, however, informed traders learn from prices and act like uninformed traders depending on their precision of information, which is a key of this model.

Readers familiar with the literature might notice that Kyle (1989) is a model for private information, not for public information. Microstructure models regarding public information release are studied by, for example, Kim and Verrecchia (2006), Kim and Verrecchia (1994), Wang (1994), Holden and Subrahmanyam (2002), Llorente and Michaely (2002). Although these papers investigate a market structure after public information releases, their model structures allow private information to incentives informed trading. The information structure of our model can be regarded as one borrowed from that of Kim and Verrecchia (1994): a source of private information is individual information production that relates the public news and terminal asset values.[4] Another source of private

[4] Another important difference from Kim and Verrecchia (1994) is that they employ quote-driven market (Kyle (1985) type model) while we employ order-driven (Kyle (1989) type model). The difference is discussed by Bernhardt and Taub (2006), but it is not in a scope of this paper.

information can be idiosyncratic liquidity shocks in pre-announcement, which is studied in Llorente and Michaely (2002) and Tetlock (2010). But this reasoning is more suitable to modeling longer-time horizon situations, not like intraday second-by-second transaction that we analyze.

Overall, our theoretical framework is not original. We use the Kyle (1989) model to derive reasonable predictions of intraday order-driven market transactions in a tractable and general manner. However, the model is rich enough to provide theoretical predictions that can be compared with our empirical analyses.

2.1 Characterization of Equilibrium[5]

We focus on a symmetric linear equilibrium which determines the strategy of informed trader. We start from a conjecture of the strategy that is written by constants β, γ, μ;

$$x_n(p, i_n) = \mu + \beta i_n - \gamma p. \tag{1}$$

The market clearing condition gives the equilibrium price as

$$p = \lambda\left(N\beta\tilde{v} + \beta\sum_n \tilde{e}_n + \tilde{z} + N\mu\right), \qquad \lambda \equiv \frac{1}{N\gamma} \tag{2}$$

λ is price impact, a price change per unit unexpected order flow, and it is interpreted as a measure of illiquidity. To derive the expression of β, γ, μ, we solve a profit maximization problem and a signal extraction problem for informed traders. Profit maximization gives a following demand function

$$x_n(i_n, p) = \frac{E(\tilde{v}|i_k, p) - p}{\lambda_I + \rho Var(\tilde{v}|i_k, p)}, \qquad \lambda_I \equiv \frac{N}{N-1}\lambda, \tag{3}$$

and a second order condition $2\lambda_I + \rho Var(\tilde{v}|i_k, p) > 0$. The signal extraction problem gives a form of conditional expectation and conditional variance

$$\begin{gathered} E(\tilde{v}|i_k, p) = \frac{(1-\varphi)\tau_e}{\tau_I} i_n + \frac{\varphi\tau_e}{\beta\lambda\tau_I} p, \\ \tau_I^{-1} \equiv Var(\tilde{v}|i_k, p) = (\tau_v + \tau_e + (N-1)\varphi\tau_e)^{-1}, \\ \varphi \equiv \frac{(N-1)\beta^2}{(N-1)\beta^2 + \sigma_z^2\tau_e}. \end{gathered} \tag{4}$$

φ is a measure of informational efficiency of the price. $\mathrm{E}(\tilde{v}) = 0$ leads $\mu = 0$. Thus, φ and τ_I are a function of β. Plugging these expressions into the demand function and matching the strategic constants β and γ with the conjecture (1), we obtain

$$\beta = \frac{(1-\varphi)\tau_e}{\lambda_I\tau_I + \rho}, \qquad \gamma = \frac{\beta\lambda\tau_I - \varphi\tau_e}{\beta\lambda(\lambda_I\tau_I + \rho)}. \tag{5}$$

[5] The derivation of the equilibrium follows the Theorem 4.1, 5.1 and 5.2 of Kyle (1989). The proofs of these theorem are provided by Kyle (1989) and we do not repeat the details.

The second order condition implies that $\beta > 0$. Substituting out $\lambda, \lambda_I, \varphi$, these two equations are combined to yield a cubic equation for endogenous variable β,

$$\rho\beta^3 + \frac{N\tau_e}{N-1}\beta^2 + \frac{\sigma_z^2\tau_e}{N-1}\rho\beta - \frac{(N-2)\sigma_z^2\tau_e^2}{(N-1)^2} = 0. \tag{6}$$

By definitions, all the coefficients on β are positive, and the constant term is negative. Thus, the left-hand side of equation (6) is monotonically increasing in β if $\beta > 0$, and a unique solution exists for $N \geq 3$. By determining β, we obtain the solutions of other endogenous variables. From equation (5), λ is a function of β and exogenous parameters. Rewriting (5) yields

$$\lambda = \frac{1}{N\beta}\left(1 - \frac{\tau_v}{\tau_I}\right) \tag{7}$$

Signal extraction problem (4) implies that $1 - \frac{\tau_v}{\tau_I}$ means the goodness of fit of the regression for $\tilde{v}$ on i_k and p. Substituting out λ and λ_I from equation (5), γ is a function of β and exogenous parameters. Some comparative statics are shown straightforwardly:

$$\frac{d\varphi}{d\beta} > 0, \qquad \frac{d\beta}{d\tau_e} > 0 \tag{8}$$

The first result is a natural consequence of (4), and the second result is obtained by applying implicit function theorem to (6). In short, the more aggressive trading on private information (i.e., large β) leads more informative price (i.e., large φ).

For this baseline model, we further assume $\tau_e = \tau_E/N$ for some constant τ_E. Under this assumption, for large N, informed speculator's signal becomes so noisy that it contains only a small amount of information. Even though, each informed trader still executes their tiny monopolistic power, and in this sense, this is a monopolistic competition model. We regard this specification fits more to the reality: since each trader has similar views on the announcement results, increasing number of informed traders does not necessarily lead to more information in the market. We set this monopolistic competition model as a benchmark.

The equation for β is now rewritten:

$$\rho\beta^3 + \frac{\tau_E}{N-1}\beta^2 + \frac{\sigma_z^2\tau_E}{N(N-1)}\rho\beta - \frac{(N-2)\sigma_z^2\tau_E^2}{N^2(N-1)^2} = 0. \tag{9}$$

We also assume $\rho = 1$ for normalization.

2.2 Implication of the model

Two endogenous variables are of our interest: liquidity and price discovery. Liquidity is defined as the inverse of price impact λ, which is easy to find its empirical equivalent. We use the effective bid-ask spread as a proxy of illiquidity. Price discovery is measured as

$$V \equiv \frac{\mathrm{Var}(p - \lambda\sigma_z^2)}{\mathrm{Var}(\tilde{v})} \tag{10}$$

Here we employ $p - \lambda\sigma_z^2 = \lambda\sum_k \beta\tilde{\imath}_k$ instead of p, for omitting the effect of noise trading. In order to construct an empirical proxy to this term, we use mid-price variance. Note that, substituting out p with (2), the expression of V boils down to a function of $R^2 \equiv \left(1 - \frac{\tau_v}{\tau_I}\right)$: $V = \left(1 + \frac{\tau_v}{\tau_E}\right)(R^2)^2$.

Two exogenous parameters are of interest: total precision of informed traders' information τ_e and the number of informed traders N. Empirically, we use R-squared measure, a goodness of fit of the surprise of news for predicting after-the-news return, for proxying τ_e. We explore comparative statics for τ_e instead of directly investigating R^2, because R^2 is endogenously obtained in the model. Theoretically, a R-squared measure and τ_e are related in a way $R^2 \equiv \left(1 - \frac{\tau_v}{\mathrm{Var}(\tilde{v}|i_k)^{-1}}\right) = \left(1 - \frac{\tau_v}{\tau_v + \tau_e}\right)$, that is, a predictive power of private information only, not counting the signal extracted from prices, for after-the-news return.

Traders are assumed to trade on their interpretation of the news just after the announcements, the number of informed traders N is regarded as the number of market participants. We use quote counts as a proxy for N. In precise, however, the same trader can submit multiple limit orders and the quote counts may not the same as the number of traders.

Regarding these variables, implications of the model are summarized in following propositions.

Proposition 1. Comparative statics for the number of informed traders N

In the monopolistic competition model,

(1) Order agressiveness: $\frac{d\beta}{dN} < 0$ if $N \geq 4$.

(2) Informational efficiency: $\frac{d\varphi}{dN} > 0$.

(3) Price infromativenesses: $\frac{d\tau_I}{dN} < 0$ and $\frac{dR^2}{dN} < 0$ if $\tau_E < \frac{(N-1)\sigma_z^2\rho}{2(N-2)}$.

(4) Price impact: $\frac{d\lambda}{dN} < 0$ if $\tau_E < \frac{(N-1)\sigma_z^2\rho}{2(N-2)}$.

(5) Price discovery: $\frac{dV}{dN} < 0$ if $\tau_E < \frac{(N-1)\sigma_z^2\rho}{2(N-2)}$.

Proof. Lemma 1, 2, 3 and Corollary 1 and 2.

Equation (5) indicates why the increase of *N* result in the decrease of β. There are two channels through which *N* affects the behavior of traders; its effect on liquidity and price informativeness. Liquidity channel is straightforward. Because the aggregate demand/supply function is linear in *N*, the increase of *N* makes the price inelastic to the unit order flow: improvement of liquidity. In the imperfect competition model, the less price becomes inelastic to the quantities (or more liquid), the traders are more eager to trade, i.e., β increases (this is implied by equation (5)).

Price informativeness channel is more complicated because it can increase as well as decreases beta. If price becomes more informationally efficient, each trader becomes reluctant to trade on their signals: β decreases. Because informational efficiency is increasing in *N* (i.e., $d\varphi/dN > 0$), the larger *N* decreases β. On the other hand, if total available information $N\tau_e$ increases in *N*, and precision of signal for traders τ_I increases. Then, trading on price information becomes less risky, traders are more aggressive and β increases.

If we assume a monopolistic competition model, the second effect of informational channel is shut down, because total available information is fixed: $N\tau_e = \tau_E$. Therefore, information channel only reduces β, which can overcome the liquidity channel and the increase of *N* result in the decrease of β.

In Proposition 1-(2), informational efficiency increases because $N\beta$ increases even when β is decreasing in *N*. In Proposition 1-(3), precision of total signal decreases because the decrease of τ_e in larger *N* cannot overcome the increase of precision from the price information. Liquidity improves as *N* becomes large because of the direct effect of thicker demand/supply curve and reduction of asymmetric information. The worsen of price discovery is a direct consequence of 1-(3).

In short, in the monopolistic competition model, the informed traders are more like an uninformed trader because their private information becomes trivial. As a result, the increase of informed trader has a similar effect as that of uninformed trader.

Proposition 2. Comparative statics for the precision of signal τ_E

In the monopolistic competition model,

(1) Order agressiveness: $\dfrac{d\beta}{d\tau_E} > 0$.

(2) Informational efficiency: $\dfrac{d\varphi}{d\tau_E} > 0$.

(3) Price infromativenesses: $\dfrac{d\tau_I}{d\tau_E} > 0$ and $\dfrac{dR^2}{d\tau_E} > 0$.

(4) Price impact: $\dfrac{d\lambda}{d\tau_E} < 0$ if τ_E is sufficiently large.

(5) Price discovery: $\dfrac{dV}{d\tau_E} > 0$.

Proof. Lemma 4, 5, 6, 7 and Corollary 3 and 4.

The implications of Proposition 2 are straightforward except for 2-(4). The more precise signal makes traders more aggressive on their information-based trading, leading the price more informative, and better price discovery realizes.

For 2-(4), it is useful to examine the price impact function $\lambda = \frac{R^2}{N\beta}$. Note that changes in τ_E affect λ in two channels: a negative one through β and a positive one thorough R^2, but the magnitude of these two channels depends on the level of τ_E. Because R^2 must be less than 1, $\frac{dR^2}{d\tau_E}$ approaches to zero as τ_E becomes large. Thus, the negative effect through β eventually dominates. The cutoff value, however, depends on other exogenous parameters.

In summary, these two propositions show that price discovery will be facilitated by the informativeness of the private signal (Proposition 2-(5)). The effect from the number of traders (Proposition 1-(4) and 1-(5)) and the effects on liquidity (Proposition 2-(4)) can depend on the level of parameters. In section 4.1, we fit exogenous parameters to data, and demonstrate the comparative statics numerically.

3 Empirical Analysis

3.1 High-Frequency Data

In this section, our dataset and their treatment are described. Our main dataset, spot Forex market rates, is obtained as firm quotes and actual deal prices at the trading platform of ICAP EBS. Data cover the transactions from January 1998 to December 2017, but the detail of records depends on the technical development of the platform. From January 1998 to December 2005, EBS provides "EBS Ticker Historical Data" which record the best prices of deals and quotes; from January 2006 to December 2017, it provides EBS Level-5 data which additionally record deal volumes as well as information of limit order book. Entire dataset allows us to examine the long-term changes in the Forex market quality, and the recent detailed data are utilized to examine the determinants of the quality.

ICAP EBS expands the universe of currency pairs over years, but we focus on EUR/USD for this analysis. Unless otherwise noted, all the empirical result is for EUR/USD. As a robustness check, we us other seven currency pairs to execute a panel regression; AUD/USD, GBP/USD, NZD/USD, USD/CAD, USD/CHF, USD/CNH, USD/JPY.

Data on deals. The order submission and matching in the actual trading process is in real-time, but the recording omits certain deals that show multiple transactions in a time slice. The observation at a time-stamp *t-1* contains the deals that occur between *t-1* and *t*. The minimum time slice, the difference between of *t-1* and *t*, varies from a second to 100 milliseconds depending on the period.[6] The recorded transaction prices at the time slice are the most extreme ones (highest paid and lowest given) during the time slice. The minimum tick size also varies depending on the period.[7]

[6] The grid of time-slices has changed during the following periods: "one second" before January 22, 2008, "a quarter-second" from January 22, 2008 to August 31, 2009, and "a 100 millisecond" from August 31, 2009 to present.

[7] The minimum tick size was traditionally four decimal places for EUR/USD (or 0.1 cent), which is called one pip. It was decimalized (i.e., five decimals, 0.01 cent) on March 7, 2011 and then rolled back to half pips after September 24, 2012.

Data on quotes. The information of quote contains the limit order prices and volumes up to ten steps (tenth best) of the limit order book. This observation is a snapshot of the limit order book, which are recorded when any change occurs in the book. The dataset also contains quote counts, the number of traders who are submitting limit orders at each step of the book.

There are some caveats regarding the dataset. First, EBS allows negative spreads: the best ask price can be lower than the best bid price. This situation happens when the two entities at the book do not have credit lines. An arbitrage opportunity for this negative spread is discussed in Ito et al. (2012). Second, when an observation has both a deal and a quote at the same time slice, the dataset does not specify the order of each transaction. We need to estimate the orders of transactions. Third, several observations in the first zero-GMT times drops accurate snapshot of the limit order book. Because we focus on the trading around the announcements, this does not affect our research. Lastly, the minimum time slice in EBS market is 100millisecond, which is bigger than that of the U.S. equity market. Potentially this discourages the high-frequency traders for trading around macro-announcement, because the speed is important for their profitability (Scholtus, Van Dijk, and Frijns (2014)). Also, with this wide time-grid, it is difficult to construct high-frequency trading activity. For studying the influence of high-frequency traders in the Forex market, we need more detailed dataset. We leave it for future research.

Table 1 shows the descriptive statistics of the data around macro announcements.

[Table 1 is inserted here]

In recent years, the frequency of quote revision has increased, but it is partly due to the technical developments in data acquisition placed in 2008. After 2008, the frequency is mostly flat. Quote counts and depth dropped in 2011, but again, this is ascribed to the changes in minimum tick size (from 0.1 cent to 0.01 cent, effective from March 2011). Subsequent recovery is also explained by the roll back of minimum tick size (from 0.01 cent to 0.5 cent, effective from September 2012). Such discontinuous changes in data potentially affect our analysis. For its treatment, we divided the quote counts and depth by the max of offer limit order minus the min of bid limit order.

Another notable change is that the deal volume decreases in recent years. This is not only for the ICAP EBS market, but BIS Central Bank Survey shows that OTC forex exchange turnover of EUR/USD also dropped globally. BIS survey points heightened activity in Japanese yen against the background of monetary policy developments (BIS Triennial Central Bank Survey 2016). The level of liquidity, measured by bid-ask spread and price impact, barely shows time trends but relative imbalance of order flows heightened after 2014.

3.2 News announcements

Government agencies in US, Japan, and EU announce macroeconomic statistics (e.g. GDP, unemployment, inflation rate and others) at pre-announced day and time. Some are monthly; some are quarterly. While some announcements are published when the US equity markets are closed, Forex markets are open 24-hour a day, and we can obtain rich sample of the market reaction to news.

For gauging the market reaction, we need a pre-announcement expectation to each news. At this point, several days before the announcement, Bloomberg compile "forecasts" by market participants and disclose the median. With these forecasts, we define the news surprise as

$$\frac{X_{t,\mathrm{actual}} - X_{t,\mathrm{forecast}}}{Std_{t-1}\left(X_{t,\mathrm{actual}} - X_{t,\mathrm{forecast}}\right)}, \tag{12}$$

where X_t is a certain announced statistic made in public at time *t*.[8] The denominator is a standard deviation of the surprise estimated by using the information up to *t-1*.[9]

Impact of surprise on returns. Following literatures, we executed a following regression which relates the Forex market returns and news surprise.[10]

$$\Delta S(t, u + 600) = \sum_{i=1}^{n} \alpha_i N_i(t, u) + \epsilon(t, u). \tag{13}$$

- (t, u) pins down each announcement timing indexed by t and its intraday time u. u is in intraday seconds from 0:00AM New York time.
- $\Delta S(t, u + 600)$: exchange rate return (log-difference) in basis points, from time u to $u + 600$ on day t. The rate is defined by mid-quotes.
- $N_i(t, u)$: surprise of *i*-th macroeconomic news statistics on day t at time u, defined as (12).

---

8 In this paper, timing of each announcement is indexed by *t*. Roughly, *t* corresponds to announcement day, but sometimes there are multiple news announcements at the same day. On the other hand, there are more than one news at the same *t*. We regard such multiple news at the same time as one chunk of sample and do not duplicate the sample by the number of simultaneous news.

9 This is implemented by a rolling estimate with 12-month windows.

10 Andersen et al. (2003), Hashimoto and Ito (2010), Fatum, Hutchison, and Wu (2012), Cheung, Fatum, and Yamamoto (2017).

- n: the number of different indicators.
- The coefficients are estimated by OLS with Newey-West standard errors (lag of 10).

We consider the total n=48 indicators, most of them are consistently available since 1998.[11] The regression results are summarized in Table 2, which is associated with the median of absolute after-the announcement return and the number of quote revision.

[Table 2 is inserted here]

The qualitative results are consistent with existing literature; the most influential statistics includes labor statistics (Changes in non-farm payroll and Unemployment rate), GDP statistics (Advance estimate), and some leading and sentiment indicators such as ISM Manufacturing and consumer confidence, etc. After the announcements of such influential statistics, jumps in the rate and quote revision tend to become large.

As Fatum, Hutchison, and Wu (2012) reports, the explanatory power of the news regression depends on macro-financial environment. We executed a rolling regression of (13) and the result is presented in the last column in Table 1. After the global financial crisis, the predictability has been sluggish until 2013.

3.3 Measures

Price discovery and liquidity recovery. As introduced in (11), we measure the price discovery as a variance ratio:

$$PD_{k,T} = \frac{Var_t\big(\Delta S(t, u+k)\big)}{Var_t\big(\Delta S(t, u+T)\big)}. \tag{14}$$

[11] Note that Chicago PMI and U. of Michigan Sentiment data are systematically released early to sub- scribers. In this research, however, omitting these indicator affect the results by little, and particular treatments are not maid for overall estimates.

k and T maintain $k \leq T$. Alternatively, we can use the weighted price contribution (WPC), the R-squared of unbiasedness regressions.[12] Since the exchange rates is defined by mid-quotes, the negative correlation from bid-ask bounce is mitigated to some extent. As we show in later section, there are slight negative, around 5%, autocorrelation of returns. Post-announcement drift, a positive return autocorrelation after announcements, is not observed.

We measure the liquidity recovery as an inverse of bid-ask spreads:

$$LR_{k,T} = E_t\left[\left\{\frac{\ln\left(ask(t,u+k)\right) - \ln\left(bid(t,u+k)\right)}{\ln\left(ask(t,u+T)\right) - \ln\left(bid(t,u+T)\right)}\right\}^{-1}\right]. \tag{15}$$

We take the inverse because a bid-ask is a measure of illiquidity. We employ bid-ask spread for constructing the long-term statistics; the other measures such as depth and price impacts need information of limit order book or transaction volume, which are not available before 2006. Also, dynamics of depth is not stable after the announcement, and price impacts needs be estimated and not obtained descriptively.

Both measures approach to one as $k \to T$; each path shows the speed of price discovery and liquidity recovery. Note that gradual price discovery corresponds to the settlement of volatility clustering, but it does not mean return predictability. When the price follows a geometric Brownian motion, the volatility ratio is linear in time. Faster price discovery draws a path above the linear line.

Figure 3 plots the year-by-year path of the each.[13] We can observe that liquidity recovery is much faster than the price discovery.

[Figure 3 are inserted here]

Figure 3 shows that the both measures do not necessarily improve over time. For price discovery, the slowest year is 2012 and fastest year is 2015. For liquidity recovery, most of years are slower than the year of 2011. For emphasis, we provide Figure 1 that shows the year-over-year changes in the both measures. In fact, there are little evidence that the both measures improve overtime.

---

[12] The statistical property of each measure is examined by van Bommel (2011). This definition is essentially the same as Chordia, Green, and Kottimukkalur (2018) who use a ratio of absolute return for measuring the speed of market reaction. This variance ratio measures the convergence of volatility clustering. In this sense, an alternate rive method is the GARCH-type modeling introduced and applied by Andersen, Bollerslev, and Das (2001), Andersen et al. (2003). In our study, however, we employ the variance ratio for pin down determinants of the time-varying property of the speed of convergence.

[13] Parameterization: for price discovery, we set T=600; for liquidity recovery, we set T=300.

3.4 Regression Analysis

Benchmark specification. Based on the observation above, a next question is (i) what the determinant of such market quality is, and (ii) whether the market quality have improved even taken account such elements. As we discussed theoretically, we consider two key drivers: precision of trader's private information and the number of traders.

If traders have rich information that helps predicting the impact on the Forex rates, the price discovery would be fast. Empirically, as reported in Fatum, Hutchison, and Wu (2012), explanatory power of the news surprise for the Forex return depends on the monetary and macroeconomic situation, and it is time varying. We can also check this in Table 1. In fact, a rolling regression of equation (13) produces time-varying R-squared, which can be regarded as a proxy for the precision of trader's private information.[14] The precision of information is positively correlated with the speed of price discovery after the announcement (Proposition 2-(5)).

For the proxy of the number of traders, we use the quote counts, the number of traders submitting limit orders. As Proposition 1 suggests, they provide liquidity but whether it contributes to price discovery depends on the quality of private information.

The empirical specification is a simple linear regression model, in its LHS holding the price discovery measure or the liquidity recovery measure. In accordance with (14), an observation by observation price discovery measure is re-defined as $\frac{|\Delta S(t,u+k)|}{|\Delta S(t,u+T)|}$, $t = 1,2,\cdots N$, with k and T are fixed to some constant. Similarly, liquidity recovery measure is $\left\{\frac{\ln\left(ask(t,u+k)\right)-\ln\left(bid(t,u+k)\right)}{\ln\left(ask(t,u+T)\right)-\ln\left(bid(t,u+T)\right)}\right\}^{-1}$, $t = 1,2,\cdots N$. We execute different regression for each k. As Scholtus, Van Dijk, and Frijns (2014) stress, the speed matters for trading based on news, and the demography of market participants may change in different time slice k.

[14] In practice, R-squared measure is calculated in a different way for avoiding non-observatory case. Please refer to the caption of Table 4 for the detailed construction.

For independent variables, other than the two proxies introduced above, we also add control variables: information on the limit order books (depth, and frequency of quote revision, both are normalized by tick size[15]), magnitude of news surprises, linear time-trend, and year dummies for 2009, 2010, 2011, and 2012, order imbalance and cumulative volumes. The time-gap of the data is not regular. For constructing these variables, we interpolated the data by replacing the missing variables with their last observances. Sampling timing is as follows. The quote counts and depth are a snapshot variable at (t, u+k), quote revision is the sum in [(t, u-15min, (t, u+15min)], order imbalance and cumulative volumes are the sum in [(t, u), (t, u+k)]. The correlation coefficients of independent variables are presented in Table 3.

The results are presented in Table 4 and Table 5. For a robustness check, we tried different *k*, the duration defining the snapshot of price discovery and liquidity recovery measure. In the last column, we also presented a result without controls.

[Table 3, Table 4 and Table 5 are inserted here]

In Table 4, a regression for price discovery, the coefficients on R-squared measure are significantly positive but those of quote counts are negative. Conversely, in Table 5, a regression for liquidity recovery, the coefficients on R-squared measure are significantly negative but those of quote counts are positive. These results hold true without controls.

As we investigated in section 2.2, R-squared measure represents the asymmetric information in the market. When it is large, informed traders have larger informational advantage on the true reaction of price to the news. Hence, the incorporation of information into price is more rapid, but uninformed traders are reluctant to trade and the liquidity becomes less. If informed traders had distinguished interpretation of news, the increase of traders might facilitate the price discovery process. But this is not the case. Conversely to this conjecture, the increase of traders delays the price discovery, implying that informed traders have poor information and act more like uninformed traders.

Notably, the trend coefficient is positive significant. The trend variable is normalized by the total time-period, and it lets us know the improvement of price discovery/liquidity recovery that is not ascribed to the variables considered in this regression. From 2006 to 2017, 30% the improvement in the speed of price discovery is up by 30% points but liquidity recovery is down by 20% points.

[15] See section 3.1 for this treatment.

For control variables, the pattern of coefficients is not very consistent between Table 4 and Table 5. The number of quote revision negatively affects both measures. This is somewhat surprising because it implies that an active trading environment does not necessarily improve market quality. The order imbalance, usually regarded as another measure of informed trading, positively affect price discovery, consistent with theoretical intuition. But its influence on liquidity is weak. It is also difficult to summarize the results for year dummies, but we can confirm that year 2012 was particularly negative for both measures.

Overall, the faster price discovery does not necessarily imply the faster liquidity recovery, vice versa. Since the R-squared measure and quote counts affect each measure in opposite way, a negative correlation between price discovery and liquidity is a possibility. Table 1 give an empirical evidence for supporting this insight.

The result that the increase of traders deteriorates price discovery is contrasted with the positive findings about high-frequency traders in existing literature; algorithmic high-frequency traders helps price discovery and liquidity provision (Brogaard, Hendershott, and Riordan (2014), Chordia, Green, and Kottimukkalur (2018)).

3.5 Extended analyses on price discovery

Some arguments should be discussed for justifying the robustness of the analysis above. First, it is questioned whether the technological development of EBS matching system, accompanied with the more frequent quote counts by high-frequency traders, can improve the price discovery or not. The time trend and the frequency of quote revision in Table 4 and Table 5 are negative, implying that such possibility is less likely. For answering this question more directly, time may be counted as tick-time instead of calendar time. In this way, we can measure the changes in market quality by each transaction.

Second, we check whether this empirical result is not only applicable to EUR/USD, but also to other currency pairs. To investigate this, we expand the universe of currency pairs to include other seven USD currency pairs and attempt a panel regression analysis.

Price discovery in tick-time The set up for regression is identical to that in section 3.4 except that the price discovery measure is defined based on tick-time: $\frac{|\Delta S(t,u+n)|}{|\Delta S(t,u+T)|}$, $t = 1,2,\cdots N$, where the numerator is indexed by tick-time *n*. Thus, we measure the speed of price discovery not by the calendar time but by the number of quote revisions. With this formulation, we can measure the ability of price discovery in per-transaction basis. The regression results are shown in Table 6.

[Table 6 is inserted here]

Regression results are similar to that in Table 4, calendar-time regression. R-squared measure have positive and quote counts have negative impact on the price discovery process. The coefficients on time trend, are slightly smaller in its magnitude but positive significant, indicating that the capacity of improving price discovery have moderately increased.

Price discovery: A panel regression analysis For a robustness check, we apply the same analysis on seven currency pairs: AUD/USD, GBP/USD, NZD/USD, USD/CAD, USD/CHF, USD/CNH, and USD/JPY.

Figure 4 corresponds to the analysis in Figure 1 and Table 4. The bottom panel is the result of fixed effect panel regression for the speed of price discovery. For readability, we only present the t-statistics of independent variables in different k-sec sample.

Similar to the results in Table 4, R-squared measure is large positive and quote counts are large negative. Time trend variable is also positive. Thus, we can confirm that major findings are still valid. Contrary to the results in EUR/USD, the number is quote revision is positive in this panel regression. This indicates that, if we gauge the activity of high-frequency traders by this variable, recent developments in electric broking system benefit more to the minor currency pairs than major currency pairs.

The top two panels plot the time-series development of price discovery and liquidity recovery for each currency pair. We can observe the developments in EUR/USD is in line with the other currency pairs: price discovery has been improving but experienced sluggish between 2009 and 2012. Liquidity recovery tends to be slow in recent years for all pairs except for some outlier in USD/CNH. In both panel, EUR/USD does not necessarily show the best market quality. Rather, price discovery of EUR/USD is one of the slowest while liquidity recovery is the quickest. Because the R-squared measure and quote counts affect each market quality in opposite direction, these two market quality measures show a negative correlation over time.

[Figure 4 is inserted here]

4 Robustness Analysis

In this section, we examine supplemental analyses that support the robustness of the results in section 3. First, we show numerical experiments of the theoretical model, with fitting the model parameters to the data. Second, we check whether there is market inefficiency after the news announcements.

4.1 Numerical experiments

Our empirical findings are summarized as follows. After the macro announcements,

- the increase of traders does not result in the improvement of price discovery but increases the liquidity.
- the higher explanatory power of the news for after-the-announcement return helps the price discovery but reduces the liquidity.

But the theoretical analysis shows that some of the results depend on the level of parameter values. For checking whether the theory correctly predicts the empirical results, we conduct numerical experiments for these findings.

Numerical experiments on price discovery We explore the informational advantage for informed trader matters for explaining the first result. For a numerical experiment, we set up a smooth transition from the monopolistic competition case to the imperfect competition case. The precision of interpretation noise τ_e is now replaced as

$$\tau_e(N,k) = x(k)\frac{\tau_E}{N_{initial}} + \left(1 - x(k)\right)\frac{\tau_E}{N}, \qquad x(k) = \frac{1}{1 + \exp(-(k - k_c))}.$$

Here, $N_{initial} = 4$ and $k_c = 5$ is assumed to be fixed. k is a parameter that shift the information structure from the monopolistic competition case (i.e., $\tau_e(N) = \frac{\tau_E}{N}$ when $k = -\infty$) toward the imperfect competition case (i.e., $\tau_e = \frac{\tau_E}{N_{initial}}$ when $k = \infty$). Other model parameters are specified as follows.

- $\tau_v = 25^{-2}$, which accounts for the standard deviation of total return of 25bp.
- $\tau_E = \tau_v \times 0.5$, which is set for the full information R-squared being 33%, i.e., $R^2 = 1 - \frac{\tau_v}{\tau_v + \tau_E} = \frac{1}{3}$.
- $\sigma_z^2 = 0.001 \times \sigma_v^2 + 0.00005 \times \tau_E$, which is set for keeping the informed trading ratio $\frac{N\beta\sigma_v + \beta\sum_n \tilde{e}_n}{N\beta\sigma_v + \beta\sum_n \tilde{e}_n + \tilde{z}}$ around 10%.

Under this model calibration, Figure 5 plot the price discovery measure $V(N) \equiv \mathrm{Var}(p(N) - \lambda(N)\sigma_z^2)/\mathrm{Var}(\tilde{v})$ as a function of N.

[Figure 5 is inserted here]

In the imperfect competition when k is large, the variance ratio is initially decreasing but increasing when N becomes large. When k is small, in the monopolistic competition, the variance ratio is decreasing in N, corresponding to our empirical finding.

Numerical experiments on liquidity Proposition 1 states that the liquidity is increasing in N (or price impact λ is decreasing in N) if τ_E is sufficiently small compared with σ_Z^2. On the other hand, as in proposition 2, the price impact can or cannot be decreasing in τ_E, which depends on parameter values. We examine the changes in liquidity around the model calibration specified above. Figure 6 shows the result, which plot the inverse of λ, or a measure of liquidity, as a function of N with different parameter values of τ_E. Now that, $\tau_E = \tau_v \times 0.5 \times k$, and $k = 0.5, 1.0, \cdots, 5.0$.

[Figure 6 is inserted here]

The liquidity is increasing in N, but it is not monotonic in the precision of noise τ_e. When N is small, the increase of τ_e results in the improvement of liquidity, while N is large, the effect reverses. Empirically, we observe the latter effect, namely, the upper-right region applies.

4.2 Predictability of after-the-news return

Post-earnings announcement drift is a well-known anomaly in stock markets (Bernard and Thomas (1989)). In this section, we examine whether such drift, or positive autocorrelation of returns, is observed after the macro announcement in the Forex market.

Testing methodology is straightforward: calculating cumulative returns and autocorrelation of returns. Following the standard methodology for showing the post-announcement drift, we can calculate the cumulative return conditional on news surprise. Each announcement sample is stratified into seven bins by their surprise, then the cumulative return from -20 minutes to +3 hours from the announcement is calculated. An alternative way is calculating simple auto correlation of return after announcements as $Corr_t\big(S(t,u+k) - S(t,u), S(t,u+3\text{hours}) - S(t,u+k)\big), k = 100sec, 200sec, \cdots$. The autocorrelation is calculated by each year sample. Figure 7 describes the two results.

[Figure 7 is inserted here]

The results indicate that there is little evidence of post-announcement drift: the price path conditional on news surprise does not have clear drifts, rather, there is a reversal after the announcements when the magnitude of news surprise is small. In fact, the autocorrelation shows mild negative correlation. For overall sample, the autocorrelation is around -5%.

5 Conclusion

In the paper, we investigate the Forex market reaction to macro announcements, and examine whether the market quality has improved over time. Novelties of our research are three-fold: first, data are long-term (from 1999 to 2017) as well as high-frequency (second-by-second transaction data); second, relating the market impact of news, a major interest of past literature, to the issues of market quality; and third, the empirical methodology and interpretation are fully grounded on a classical market microstructure theory.

Overall, empirical evidence shows that the market quality in the Forex market has not necessarily improved in recent years: the speed of price discovery has improved but liquidity recovery has not. Liquidity recovery was fast enough in 1999 and leaving only a small room of improvements. The fast price discovery is led by the more precise information in traders, but it also implies the severer asymmetric information and then worsen liquidity. Hence, fast price discovery and liquidity recovery do not necessarily coincide. Contrary to the popular conjecture of “sophisticated high-frequency traders”, increase of traders right after the news release does not facilitate the price discovery. This suggests that such “fast” traders do not have precise interpretation that relates the news statistics and Forex prices.

Appendix A: Proof of propositions

Lemma 1.

$$\frac{d\beta}{dN}\frac{N}{\beta} > -1.\ \text{If } N \geq 4, \qquad \frac{d\beta}{dN}\frac{N}{\beta} < 0.$$

Proof. Multiplying the cubic equation $\rho\beta^3 + \frac{\tau_E}{N-1}\beta^2 + \frac{\sigma_Z^2\tau_E}{N(N-1)}\rho\beta - \frac{(N-2)\sigma_Z^2\tau_E^2}{N^2(N-1)^2} = 0$ by N^3, we obtain

$$\rho(N\beta)^3 + \frac{N\tau_E}{N-1}(N\beta)^2 + \frac{N\sigma_Z^2\tau_E}{(N-1)}\rho N\beta - \frac{N(N-2)\sigma_Z^2\tau_E^2}{(N-1)^2} = 0.$$

As N increases, the coefficients on the first three terms decrease, while the last term increases. Therefore $N\beta$ should be increasing in N. Equivalently, $\frac{d\beta}{dN}\frac{N}{\beta} > -1$.

Next, applying the implicit function theorem to $F(N,\beta) \equiv \rho\beta^3 + \frac{\tau_E}{N-1}\beta^2 + \frac{\sigma_Z^2\tau_E}{N(N-1)}\rho\beta - \frac{(N-2)\sigma_Z^2\tau_E^2}{N^2(N-1)^2} = 0$, we have

$$\frac{d\beta}{dN} = -\frac{\frac{\partial F}{\partial N}}{\frac{\partial F}{\partial \beta}} = -\left(\frac{\partial F}{\partial \beta}\right)^{-1}\left\{-\frac{\tau_E}{(N-1)^2}\beta^2 - \frac{(2N-1)\sigma_Z^2\tau_E\rho}{N^2(N-1)^2}\beta - \frac{(-3N^2+9N-4)\sigma_Z^2\tau_E^2}{N^3(N-1)^3}\right\},$$

where $\frac{\partial F}{\partial \beta} = 3\rho\beta^2 + \frac{2\tau_E}{N-1}\beta + \frac{\sigma_Z^2\tau_E}{N(N-1)}\rho > 0$ is straightforwardly obtained because $\beta > 0$. Using $F(N,\beta) = 0$, the last term of $\frac{\partial F}{\partial N}$ is substituted out. The resulting expression is a polynomial of β without constant terms, and each coefficient of β is greater than zero if $N \geq 4$. Thus, we have $\frac{\partial F}{\partial N} > 0$ if $N \geq 4$, and $\frac{d\beta}{dN}\frac{N}{\beta} < 0$ can be obtained. ■

Lemma 2.

$$\frac{d\varphi}{dN} > 0.$$

Proof. Differentiating $\varphi = \frac{(N-1)\beta^2}{(N-1)\beta^2 + \sigma_Z^2\tau_e}$ with respect to N, we have

$$\frac{d\varphi}{dN} = \left\{(N-1)\beta^2 + \frac{\sigma_Z^2\tau_E}{N}\right\}^{-2}\left\{\beta^2 + 2(N-1)\frac{d\beta}{dN}\beta - (N-1)\beta^2\left(2(N-1)\frac{d\beta}{dN} - \frac{\sigma_Z^2\tau_E}{N^2}\right)\right\}.$$

The first parenthesis is clearly positive. The second parenthesis reduces to $(2-\frac{1}{N})\beta^2 + 2(N-1)\frac{d\beta}{dN}\beta$. After rearranging terms, we have

$$\frac{d\varphi}{dN} > 0 \iff \frac{d\beta/\beta}{dN/N} > -\frac{2N-1}{2(N-1)}.$$

Lemma 1 claims that this is satisfied in $N \geq 4$. ■

Lemma 3.

$$\frac{d\tau_I}{dN} < 0 \text{ if } \tau_E < \frac{(N-1)\sigma_Z^2\rho}{2(N-2)}.$$

Proof. By definition, $\tau_I \equiv \tau_v + \frac{\tau_E}{N} + (N-1)\varphi\frac{\tau_E}{N}$. Differentiating with respect to N, we have

$$\frac{d\tau_I}{dN} = -\frac{\tau_E}{N^2} + \frac{d\varphi}{dN}(N-1)\frac{\tau_E}{N} + \varphi\frac{\tau_E}{N^2}.$$

Rearranging terms, we have

$$\frac{d\tau_I}{dN} < 0 \iff \frac{d\varphi}{dN} < \frac{1-\varphi}{N(N-1)}.$$

Plugging $\frac{d\varphi}{dN}$ in, this is rewritten:

$$\frac{d\tau_I}{dN} < 0 \;\Leftrightarrow\; \frac{d\varphi}{dN} < \frac{1-\varphi}{N(N-1)} \;\Leftrightarrow\; \frac{d\beta/\beta}{dN/N} < -1 + \frac{\sigma_z^2\tau_E}{2N(N-1)^2\beta^2}.$$

We have the expression for $\frac{d\beta}{dN}$ in Lemma 1, rearranging it yields

$$\frac{d\beta/\beta}{dN/N} = -\frac{-\frac{N\tau_E}{(N-1)^2}\beta^2 - \frac{(2N-1)\sigma_z^2\tau_E\rho}{N(N-1)^2}\beta - \frac{(-3N^2+9N-4)\sigma_z^2\tau_E^2}{N^2(N-1)^3}}{-\frac{\tau_E}{(N-1)}\beta^2 - \frac{2\sigma_z^2\tau_E\rho}{N(N-1)}\beta + \frac{3(N-2)\sigma_z^2\tau_E^2}{N^2(N-1)^2}}.$$

Here, the denominator is obtained by substituting $3\rho\beta^3$ out by $F(N,\beta)=0$. Plugging $\frac{d\beta/\beta}{dN/N}$ in, after some calculations, we have

$$\frac{d\beta/\beta}{dN/N} < -1 + \frac{\sigma_z^2\tau_E}{2N(N-1)^2\beta^2}$$

$$\Leftrightarrow\; \frac{\tau_E}{(N-1)}\beta^3 + \left(1-\frac{3}{2}\right)\frac{\sigma_z^2\tau_E\rho}{N(N-1)}\beta^2 + \left(\frac{2}{N}-1\right)\frac{\sigma_z^2\tau_E^2}{N(N-1)^2}\beta - \frac{\rho\sigma_z^4\tau_E^2}{2N^2(N-1)^2} < 0.$$

Again, the last term is substitute out by using $F(N,\beta)=0$. The coefficients for β^2 and β are clearly negative. The coefficients for β^3 is $\frac{\tau_E}{(N-1)} - \frac{\sigma_z^2\rho}{2(N-2)}$, which can be negative if $\tau_E < \frac{(N-1)\sigma_z^2\rho}{2(N-2)}$. ■

Corollary 1.

$$\frac{dR^2}{dN} < 0 \; and \; \frac{dV}{dN} > 0 \text{ if } \tau_E < \frac{(N-1)\sigma_z^2\rho}{2(N-2)}.$$

Proof. By definition, $R^2 = 1 - \frac{\tau_v}{\tau_I}$. The sign of $\frac{dR^2}{dN}$ coincide with $\frac{d\tau_I}{dN}$ because $\frac{dR^2}{dN} = \frac{\tau_v}{\tau_I^2}\frac{d\tau_I}{dN}$. $\frac{dV}{dN} > 0$ immediately follows. ■

Corollary 2.

$$\frac{d\lambda}{dN} < 0 \text{ if } \tau_E < \frac{(N-1)\sigma_z^2\rho}{2(N-2)}.$$

Proof. The price impact is $\lambda = \frac{1}{N\beta}\left(1-\frac{\tau_v}{\tau_I}\right)$. Lemma 1 suggests that $\frac{1}{N\beta}$ is decreasing in *N*, and Corollary 1 suggests $\left(1-\frac{\tau_v}{\tau_I}\right)$ is also decreasing in *N*. Thus λ is decreasing in *N*. ■

Lemma 4.

$$\frac{d\beta}{d\tau_E} > 0.$$

Proof. Applying the implicit function theorem to $F(N,\beta) \equiv \rho\beta^3 + \frac{\tau_E}{N-1}\beta^2 + \frac{\sigma_z^2\tau_E}{N(N-1)}\rho\beta - \frac{(N-2)\sigma_z^2\tau_E^2}{N^2(N-1)^2} = 0$, we have

$$\frac{d\beta}{d\tau_E} = -\frac{\frac{\partial F}{\partial \tau_E}}{\frac{\partial F}{\partial \beta}} = -\left(\frac{\partial F}{\partial \beta}\right)^{-1}\left\{\frac{1}{(N-1)}\beta^2 + \frac{\sigma_z^2\rho}{N(N-1)}\beta - \frac{2(N-2)\sigma_z^2\tau_E}{N^2(N-1)^2}\right\},$$

and $\frac{\partial F}{\partial \beta} > 0$ as in Lemma 1. The last term is substitute out by $F(N,\beta)=0$, and we have

$$\frac{\partial F}{\partial \tau_E} = -\frac{2\rho\beta^3}{\tau_E} + \left(\frac{1}{N-1} - \frac{2}{N-1}\right)\beta^2 + \left(\frac{1}{N(N-1)} - \frac{2}{N(N-1)}\right)\sigma_z^2\rho\beta < 0,$$

because each term in parentheses is negative and $\beta > 0$. Finally, we obtain $\frac{d\beta}{d\tau_E} > 0$. ■

Lemma 5.

$$\frac{d\varphi}{d\tau_E} > 0.$$

Proof. By definition, $\varphi \equiv \frac{(N-1)\beta^2}{(N-1)\beta^2+\sigma_z^2\tau_E/N}$. Differentiating with respect to by τ_E and rearranging terms, we have

$$\frac{d\varphi}{d\tau_E} = \left((N-1)\beta^2 + \frac{\sigma_z^2\tau_E}{N}\right)^{-2}\left\{2(N-1)\frac{d\beta}{d\tau_E}\beta\frac{\sigma_z^2\tau_E}{N} - (N-1)\beta^2\frac{\sigma_z^2}{N}\right\}.$$

It is enough to show the numerator is positive. Rearranging terms, we can show the numerator is positive if and only if

$$\frac{d\beta}{d\tau_E}\frac{\tau_E}{\beta} > \frac{1}{2}.$$

In the LHS, substituting out $\frac{d\beta}{d\tau_E}$ yields

$$\frac{d\beta}{d\tau_E}\frac{\tau_E}{\beta} = -\frac{\frac{\tau_E}{(N-1)}\beta^2 + \frac{\sigma_z^2\tau_E\rho}{N(N-1)}\beta - \frac{2(N-2)\sigma_z^2\tau_E^2}{N^2(N-1)^2}}{-\frac{\tau_E}{(N-1)}\beta^2 - \frac{2\sigma_z^2\tau_E\rho}{N(N-1)}\beta + \frac{3(N-2)\sigma_z^2\tau_E^2}{N^2(N-1)^2}}.$$

Note that the denominator is positive. Rearranging terms, we have

$$\frac{d\beta}{d\tau_E}\frac{\tau_E}{\beta} > \frac{1}{2} \iff \frac{-\tau_E}{(N-1)}\beta^2 + \frac{(N-2)\sigma_z^2\tau_E^2}{N^2(N-1)^2} > 0 \iff \rho\beta^3 + \frac{\sigma_z^2\tau_E\rho}{N(N-1)}\beta > 0.$$

Last equivalence holds from $F(N,\beta) = 0$. ■

Lemma 6.

$$\frac{d\tau_I}{d\tau_E} > 0.$$

Proof. Differentiating $\tau_I \equiv \tau_v + \frac{\tau_E}{N} + \frac{(N-1)\tau_E}{N}\varphi(\tau_E)$ with respect to τ_E, we have $\frac{d\tau_I}{d\tau_E} = \frac{1}{N} + \frac{(N-1)}{N}\varphi(\tau_E) + \frac{(N-1)\tau_E}{N}\frac{d\varphi}{d\tau_E}$. Applying Lemma 5, with $\varphi > 0$ by definition, we have the desired result. ■

Corollary 3.

$$\frac{dR^2}{d\tau_E} > 0.$$

Proof. Differentiating of $R^2 = \left(1 - \frac{\tau_v}{\tau_I(\tau_E)}\right)$ by τ_E, we have $\frac{dR^2}{d\tau_E} = \frac{\tau_v}{\tau_I^2}\frac{d\tau_I}{d\tau_E}$. Applying Lemma 6, we have the desired result.

Corollary 4.

$$\frac{d\lambda}{d\tau_E} < 0 \text{ if } \tau_E \text{ is sufficiently large.}$$

Proof. From Lemma 4, as τ_E goes to infinity, $\frac{d\beta}{d\tau_E}$ remains positive. On the other hand, from Corollary 3 and that R^2 is less than 1 by definition, $\frac{dR^2}{d\tau_E}$ goes to zero as τ_E goes to infinity. Because $\lambda = \frac{R^2}{N\beta}$ and $\frac{d\lambda}{d\tau_E}\frac{\tau_E}{\lambda} = \left(\frac{dR^2}{d\tau_E}\frac{1}{R^2} - \frac{d\beta}{d\tau_E}\frac{1}{\beta}\right)\tau_E$, $\frac{d\lambda}{d\tau_E}$ becomes negative (positive) if τ_E is sufficiently large (small). The cutoff value depends on other exogenous parameters. ■

Lemma 7.

$$\frac{dV}{d\tau_E} > 0.$$

Proof. $V = \left(1 + \frac{\tau_v}{\tau_E}\right)(R^2)^2 = \left(1 + \frac{\tau_v}{\tau_E}\right)\left(1 - \frac{\tau_v}{\tau_I}\right)^2$. Substituting out τ_I, we obtain

$$V = \left(1 + \frac{\tau_v}{\tau_E}\right)\left(\frac{1 + (N-1)\varphi}{1 + (N-1)\varphi + \frac{N\tau_v}{\tau_E}}\right)^2.$$

Since $V > 0$, for obtaining $\frac{dV}{d\tau_E} > 0$, it is enough to show $\frac{d\ln V}{d\tau_E} > 0$.

$$\frac{d\ln V}{d\tau_E} = \frac{d\varphi}{d\tau_E}\left(\frac{2(N-1)}{1 + (N-1)\varphi} - \frac{2(N-1)}{1 + (N-1)\varphi + \frac{N\tau_v}{\tau_E}}\right) - \frac{N\tau_v}{\tau_E^2}\left(\frac{1}{N + \frac{N\tau_v}{\tau_E}} - \frac{2}{1 + (N-1)\varphi + \frac{N\tau_v}{\tau_E}}\right)$$

From Lemma 5 we have $\frac{d\varphi}{d\tau_E} > 0$, and the first parenthesis is positive because $\varphi > 0$. Also, we can show the second parenthesis is negative. This results in $\frac{d\ln V}{d\tau_E} > 0$. ■

Appendix B: A competitive model

Competitive equilibrium. In a competitive equilibrium, the demand function of (3) is replaced with $x_n(i_n, p) = \frac{E(v|i_k, p) - p}{\rho Var(v|i_k, p)}$. In this case, the equation for β is rewritten

$$\rho\beta_c^3 + \frac{\sigma_z^2\tau_e}{N-1}\rho\beta_c - \frac{\sigma_z^2\tau_e^2}{N-1} = 0. \quad (16)$$

We denote the subscript c for indicating results for competitive equilibrium. The other endogenous variables are obtained accordingly. As suggested by Theorem 6.1 of Kyle (1989), in this competitive equilibrium, presence of uninformed trader does not alter the strategy of informed traders. Therefore, the increase of uninformed traders does not contribute to the price discovery; since the equilibrium price is a weight average of informed's expectation and uninformed's expectation, an increase of uninformed traders makes the price distant from the true fundamentals.

Comparing (6) to (16), we find that $\beta_c > \beta$ always holds. This means that informed traders are driven to be more aggressive on trading based on their own information. Aggressive information trading leads to more informative price, i.e., $\varphi_c > \varphi$, which can be obtained from (4). An imperfect competition equilibrium converges to a competitive equilibrium as $N \to \infty$ and $\sigma_z^2\rho/N\tau_e \to \infty$ (Theorem 9.2 of Kyle (1989)). $N \to \infty$ means the market becomes "large", and $\sigma_z^2\rho/N\tau_e \to \infty$ means that informed traders do not dominate the market; per unit noise σ_z/N kept at constant.

References

Anand, Amber, Carsten Tanggaard, and Daniel G. Weaver, 2009, Paying for Market Quality, *Journal of Financial and Quantitative Analysis* 44, 1427–1457.

Andersen, Torben G. TG, Tim Bollerslev, Francis X. FX Francis X. Diebold, and Clara Vega, 2003, Micro effects of macro announcements: Real-time price discovery in foreign exchange, *The American Economic Review* 93, 38–62.

Andersen, Torben G, Tim Bollerslev, and Ashish Das, 2001, Variance-ratio Statistics and High-frequency Data: Testing for Changes in Intraday Volatility Patterns., *The Journal of Finance* 56, 305–327.

Bernard, Victor, and Jacob Thomas, 1989, Post-Earnings-Announcement Drift : Delayed Price Response or Risk Premium ?, *Journal of Accounting Research* 27, 1–36.

Bernhardt, Dan, and Bart Taub, 2006, Kyle v.s. Kyle ('85 v. '89), *Annals of Finance* 2, 23--38.

Brogaard, Jonathan, Terrence Hendershott, and Ryan Riordan, 2014, High-Frequency Trading and Price Discovery, *Review of Financial Studies* 27, 2267–2306.

Cheung, Yin-wong, Rasmus Fatum, and Yohei Yamamoto, 2017, Working Paper No. 305 - The Exchange Rate Effects of Macro News after the Global Financial Crisis - Dallas Fed, *Journal of International Money and Finance*.

Chordia, Tarun, T Clifton Green, and Badrinath Kottimukkalur, 2018, Rent seeking by low-latency traders: Evidence from trading on macroeconomic announcements, *Review of Financial Studies* 31, 4650–4687.

Fatum, Rasmus, Michael Hutchison, and Thomas Wu, 2012, Asymmetries and State Dependence : The Impact of Macro Surprises on Intraday Exchange Rates, *Journal of the Japanese and International Economies* 26, 542--560.

Grossman, Sanford J., and Joseph E. Stiglitz, 1980, On the impossibility of informationally efficient markets, *The American Economic Review* 70, 393–408.

Hashimoto, Yuko, and Takatoshi Ito, 2010, Effects of Japanese macroeconomic statistic announcements on the dollar/yen exchange rate: High-resolution picture, *Journal of the Japanese and International Economies* 24, 334–354.

Holden, Craig W., and Avanidhar Subrahmanyam, 2002, News Events, Information Acquisition, and Serial Correlation, *The Journal of Business* 75, 1–32.

Ito, Takatoshi, Kenta Yamada, Misako Takayasu, and Hideki Takayasu, 2012, Free Lunch! Arbitrage Opportunities in the Foreign Exchange Markets, *NBER working paper (18541)*.

Jiang, George, Ingrid Lo, and Giorgio Valente, 2012, High-Frequency Trading in the U.S. Treasury Market Around Macroeconomic News Announcements, *Working Paper, Washington State University*.

Kim, Oliver, and Robert E. Verrecchia, 1994, Market liquidity and volume around earnings announcements, *Journal of Accounting and Economics* 17, 41–67.

Kim, Oliver, and Robert E. Verrecchia, 1997, Pre-announcement and event-period private information, *Journal of Accounting and Economics* 24, 395–419.

Kim, Oliver, and Robert E Verrecchia, 2006, Trading Volume and Price Reactions to Public Announcements, *Journal of Accounting Research* 29, 302.

Kirilenko, Andrei, Albert S Kyle, Mehrdad Samadi, and Tugkan Tuzun, 2017, The Flash Crash: High Frequency Trading in an Electronic Market, *The Journal of Finance* 72, 967–998.

Kyle, AS, 1989, Informed speculation with imperfect competition, *The Review of Economic Studies* 56, 317–355.

Llorente, Guillermo, and R Michaely, 2002, Dynamic volume-return relation of individual stocks, *Review of Financial Studies* 15, 1005–1047.

Scholtus, Martin, Dick Van Dijk, and Bart Frijns, 2014, Speed, algorithmic trading, and market quality around macroeconomic news announcements, *Journal of Banking and Finance* 38, 89–105.

Tetlock, Paul C., 2010, Does public financial news resolve asymmetric information?, *Review of Financial Studies* 23, 3520–3557.

van Bommel, Jos, 2011, Measuring price discovery: The variance ratio, the R2, and the weighted price contribution, *Finance Research Letters* 8, 112–119.

Wang, Jiang, 1994, A Model of Competitive Stock Trading Volume, *Journal of Political Economy* 102, 127.

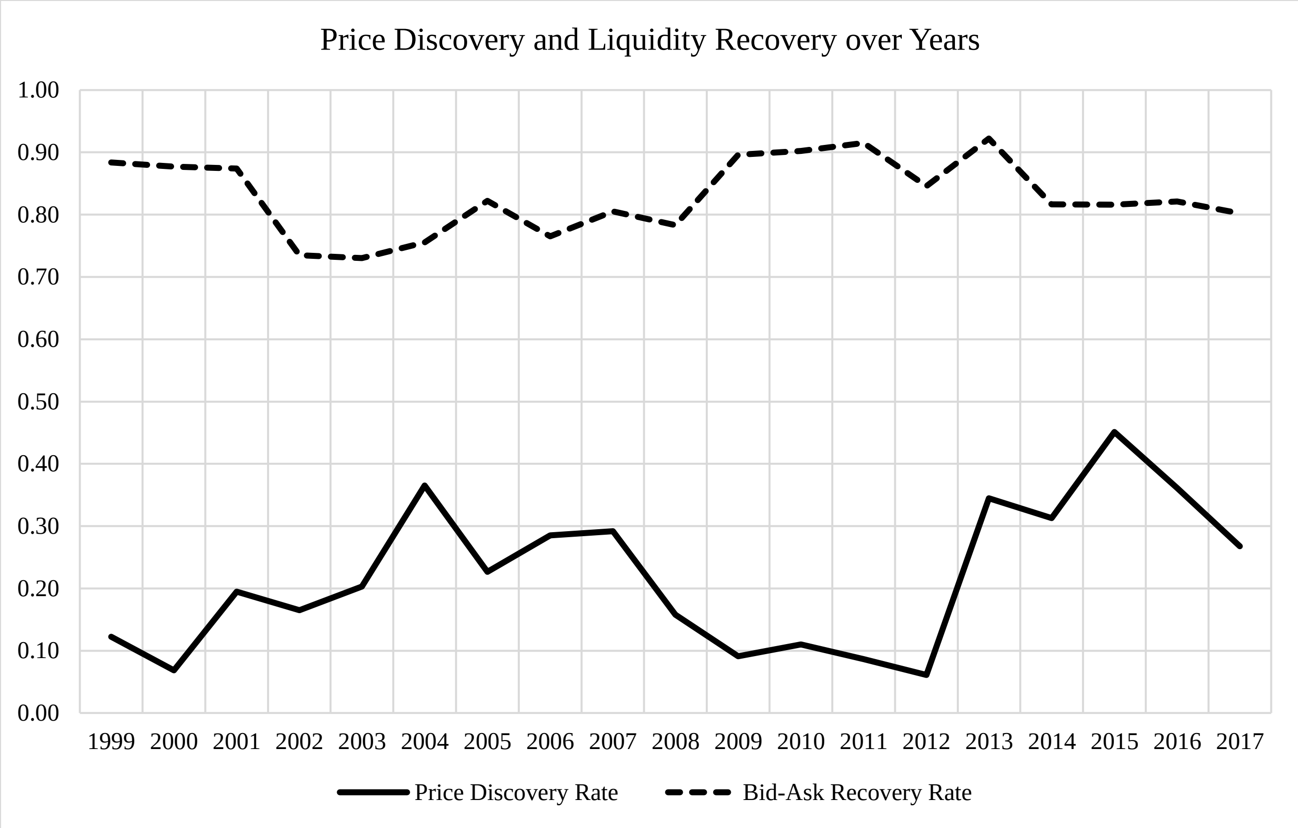


Figure 1: Yearly changes in the speed of price discovery and Liquidity recovery

The definition of price discovery and liquidity recovery are provided at equation (14) and (4). We calculate $PD_{k=60,T=1800}$, $LR_{k=5,T=60}$ for each yearly sub-sample. Both measures do not show evident improvement over years; the market quality does not necessarily become better in recent years.

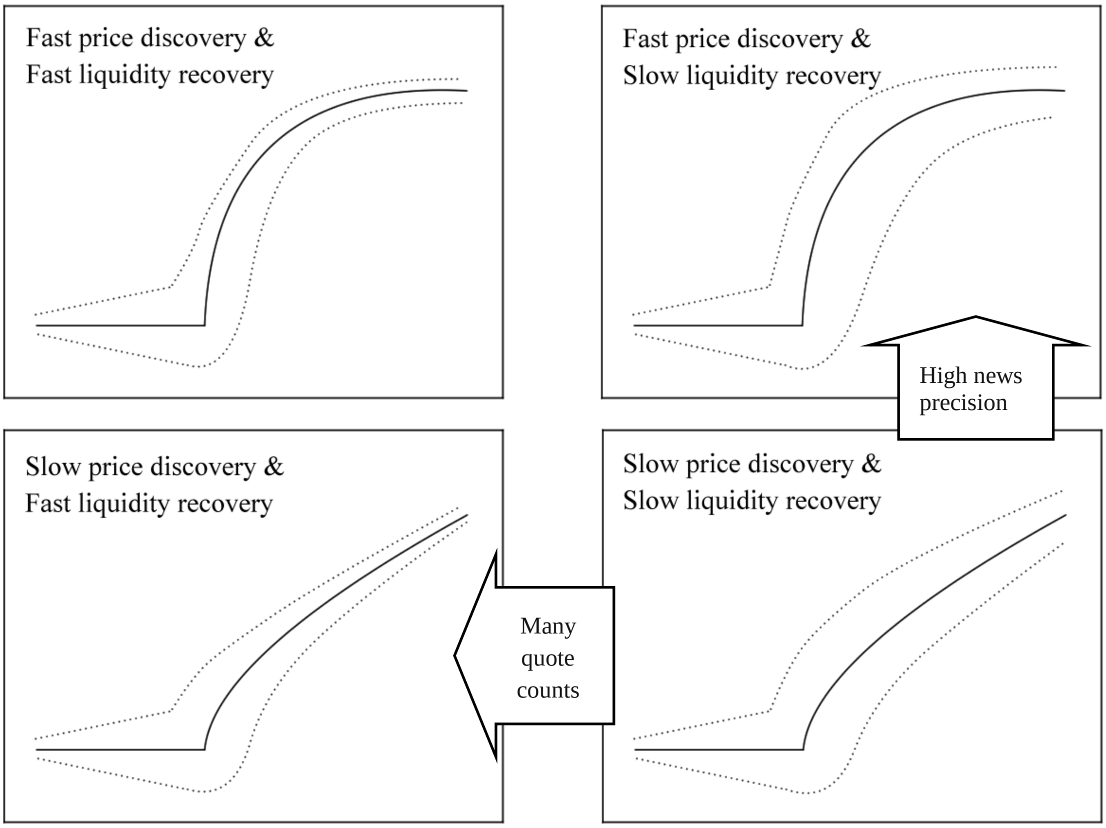


Figure 2: Diagram of empirical findings

In each panel, the solid line describes a sample path of mid-quotes, which is associated with the bid prices (the line below) and ask prices (the line above). When the price discovery is fast, the price path converges to the stable level quickly. When the liquidity recovery is fast, bid-ask spreads become narrow quickly. Our interest is what effects two key variables, the explanatory power of news for after-the-news return, and the number of traders have on these.

Table 1: Selected descriptive statistics after macro announcements

year	# of quote revision	quote counts	depth (n<=5)	depth (n>=5)	depth all	bid-ask spread	|return|	# of trade	volume	|OIB|/VOL	ret/OIB	R^2
1999	9	-	-	-	-	0.95	1.32	-	-	-	-	-
2000	11	-	-	-	-	1.13	1.58	-	-	-	-	0.15
2001	14	-	-	-	-	1.14	2.36	-	-	-	-	0.33
2002	14	-	-	-	-	1.10	2.51	-	-	-	-	0.16
2003	17	-	-	-	-	0.91	2.60	-	-	-	-	0.21
2004	19	-	-	-	-	0.82	2.48	-	-	-	-	0.54
2005	17	-	-	-	-	0.83	2.07	-	-	-	-	0.46
2006	28	84	146	96	252	0.80	2.94	26	202	0.15	0.05	0.27
2007	28	108	175	145	335	0.74	1.81	23	124	0.18	0.02	0.23
2008	78	74	91	75	173	0.74	2.77	36	104	0.20	0.05	0.29
2009	96	103	88	66	156	1.35	2.25	36	87	0.17	0.07	0.06
2010	130	165	125	99	233	0.79	1.96	47	102	0.15	0.04	0.03
2011	96	36	21	23	45	0.95	2.19	44	83	0.18	0.05	0.11
2012	78	39	23	27	51	0.77	1.33	31	53	0.16	0.02	0.05
2013	104	86	49	57	108	0.76	2.11	43	78	0.18	0.06	0.36
2014	82	85	54	57	116	0.77	1.63	32	69	0.25	0.04	0.48
2015	100	74	39	48	89	1.33	3.36	36	72	0.28	0.07	0.47
2016	71	81	43	51	97	1.33	2.00	20	37	0.36	0.03	0.40
2017	69	103	56	61	115	0.89	1.78	16	29	0.35	0.04	0.32

Quote counts, depth, bid-ask spread are a snapshot at (t,u+30sec). Other indicators are the sum between (t,u) to (t,u+30sec). "depth (n<=5) "is the sum of limit order up to 5 steps. Each statistic is the median in each year. "OIB" stands for order imbalance, or the sum of deals initiated by buyers minus deals initiated by sellers. Returns are defined as log return in basis points. Before 2006, detailed limit order book data (quote counts and depth) and deal data are not available. Note that the minimum tick size is revised in 2011 from 1 pip to 1/10 pip, and in 2012 from 1/10 pip to 1/2 pip, showing discrete jumps in quote counts and depth. R-squared is calculated from the 50-day window rolling regression of (13).

Table 2: Contemporaneous News Response Coefficients, t-values, and Average return magnitudes

y = (log(S(t,u+600sec) - log(S(t,u))x10,000	coefficient	t-stat	Median of abs(y) 10-min after the news	Median quote revision in 10-min after the news	Frequency (monthly UON)	Availability (from 1999 UON)	Significance reported in Andersen et al. (2003)
Chg. Nonfarm Payrolls	-20.187	-6.460	22.9	1995			†
GDP_A	-11.073	-4.446	13.4	1252	Quarterly		†
ISM Manufacturing	-6.839	-4.446	10.0	784			†
Retail Sales	-6.054	-4.431	9.4	751			†
Trade Balance	-6.454	-4.298	9.8	589			†
Empire Manufacturing	-3.873	-4.121	6.4	954		from 2004	
Consumer Confidence	-4.642	-3.443	8.9	556			†
New Home Sales	-2.781	-3.311	7.0	595			
Durable Goods Orders	-3.054	-3.096	7.5	587			†
ADP	-4.984	-2.962	6.2	1461		from 2007	
Chicago Purchasing Manager	-2.599	-2.899	8.1	577			
Avg. Hourly Earnings	-7.963	-2.839	22.9	1995			
Current Account	-3.683	-2.798	6.5	549	Quarterly		
Factory Orders	-2.423	-2.724	8.4	590			†
Phil. Fed Outlook	-3.228	-2.233	7.0	594			
CPI	-2.347	-2.148	9.6	1415			
Housing Starts	-1.706	-1.917	6.9	579			
Industrial Production	-2.021	-1.837	6.6	544			†
PPI. Ex. Food Energy	-1.660	-1.501	8.0	691			
U of Mich Sentiment_F	-1.245	-1.413	8.8	648		from 2001	
GDP_S	-2.367	-1.403	8.8	830	Quarterly		
U of Mich Sentiment_P	-1.421	-1.357	8.7	852		from 2001	
Personal Consumption_T	-3.193	-1.318	7.3	847	Quarterly	from 2004	
Leading Index	-0.949	-1.254	7.1	629			
NAHB Housing Market Index	-1.313	-1.203	4.7	656		from 2004	
Personal Consumption_S	-4.416	-1.022	8.1	1164	Quarterly	from 2004	
Existing Home Sales	-1.116	-0.916	8.1	576			
PCE Core	-0.863	-0.821	6.4	868		from 2005	
GDP Def_A	-1.753	-0.793	13.4	1332	Quarterly	from 2001	
Import Price	-1.181	-0.739	7.6	897		from 2005	
GDP Def_T	-1.056	-0.704	7.3	570	Quarterly	from 2001	
Chg. Manufact. Payrolls	-1.715	-0.625	22.9	1995			
GDP Def_S	-0.723	-0.530	8.8	830	Quarterly	from 2001	
Personal Consumption_A	-1.278	-0.391	13.4	1631	Quarterly	from 2004	
Budget Statement	-0.141	-0.318	3.6	143			
Whole Sale Inventories	-0.207	-0.234	7.0	571			
Personal Spending	-0.174	-0.217	6.4	538			
GDP_T	-0.079	-0.041	7.0	561	Quarterly		
Unit Labor Cost_F	-0.053	-0.030	7.6	708	Quarterly	from 2002	
Employment Cost Index	-0.054	-0.017	9.4	615	Quarterly		
Construction Spending	0.186	0.150	9.4	633			
Building Permits	0.466	0.471	6.9	846		from 2004	
Consumer Credit	0.287	0.773	2.5	127			
Unit Labor Cost_P	2.360	0.855	7.5	904	Quarterly	from 2002	
Business Inventories	0.748	0.862	7.7	566			
Capacity Utilization	1.367	1.291	6.6	548			
Personal Income	1.251	1.621	6.4	537			
Unemployment Rate	7.819	3.661	22.9	1995			
Obs.		5497					
R-squared		0.105					

Note: Suffix A, S, T, F, P denotes 'Advance', Second', 'Third', 'Final', and 'Preliminary'

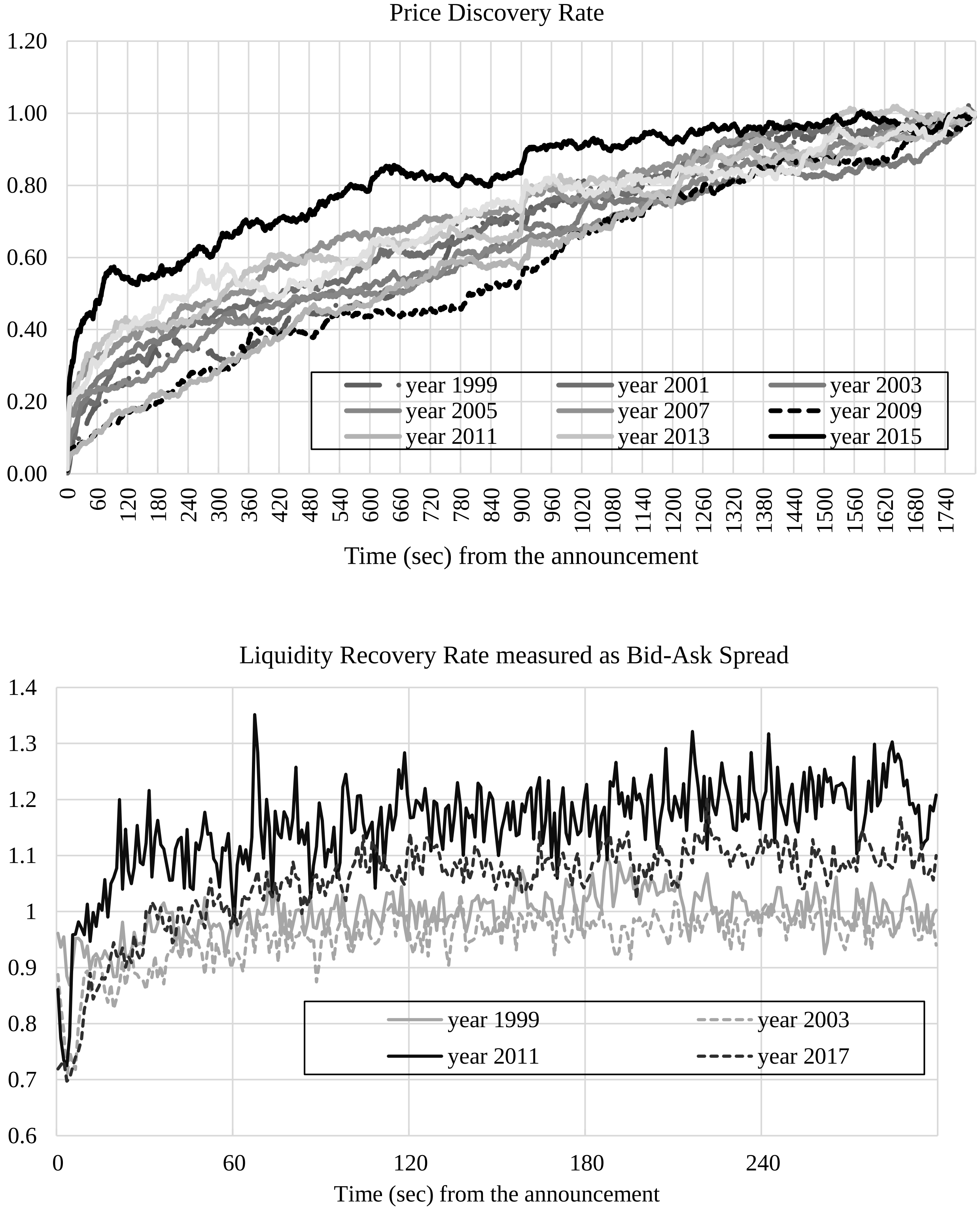


Figure 3: Price discovery and liquidity recovery after the news

The definition of price discovery and liquidity recovery are presented in equation (14) and (15). The panel above shows the $PD_{k,T=1800}$ and the one below shows $LR_{k,T=600}$.

Table 3: Correlation coefficients of independent variables

	R^2	Quote Count	Liquidity recovery	Depth	Surprise of news	# of quote revision	Time trend	Y2009	Y2010	Y2011	Y2012	|OIB|/VOL	VOL
R^2	1.00	-0.09	0.03	0.06	0.00	0.03	-0.01	-0.06	-0.07	-0.06	-0.11	-0.08	0.18
Quote Count at (t,u+k)	-0.09	1.00	-0.02	0.27	-0.05	0.16	0.49	-0.12	0.17	-0.09	0.12	0.11	-0.23
Liquidity recovery rate at (t,u+k) x 100	0.03	-0.02	1.00	0.04	0.00	0.02	0.00	0.00	0.01	-0.03	-0.06	0.04	0.17
Depth at (t,u+k)	0.06	0.27	0.04	1.00	-0.02	-0.22	-0.33	0.03	0.21	-0.39	-0.36	0.00	0.18
Surprise of news	0.00	-0.05	0.00	-0.02	1.00	0.00	-0.02	0.04	-0.06	-0.01	-0.01	0.00	0.04
# of quote revision	0.03	0.16	0.02	-0.22	0.00	1.00	0.18	0.19	0.24	0.20	0.10	-0.26	0.42
Linear time trend	-0.01	0.49	0.00	-0.33	-0.02	0.18	1.00	-0.19	-0.12	-0.03	0.05	0.26	-0.41
Y2009	-0.06	-0.12	0.00	0.03	0.04	0.19	-0.19	1.00	-0.09	-0.08	-0.09	-0.05	0.04
Y2010	-0.07	0.17	0.01	0.21	-0.06	0.24	-0.12	-0.09	1.00	-0.09	-0.09	-0.07	0.09
Y2011	-0.06	-0.09	-0.03	-0.39	-0.01	0.20	-0.03	-0.08	-0.09	1.00	-0.09	-0.08	0.02
Y2012	-0.11	0.12	-0.06	-0.36	-0.01	0.10	0.05	-0.09	-0.09	-0.09	1.00	-0.06	-0.11
|OIB|/cumulative volume(t,u+k)	-0.08	0.11	0.04	0.00	0.00	-0.26	0.26	-0.05	-0.07	-0.08	-0.06	1.00	-0.38
cumulative volume(t,u+k)	0.18	-0.23	0.17	0.18	0.04	0.42	-0.41	0.04	0.09	0.02	-0.11	-0.38	1.00

Table 4: Regression results for the determinants of speed of price discovery

	$y = PD^R_{t,k,T=600}$	k=5sec	k=40sec	k=120sec	k=240sec	k=500sec	k=40sec
(1)	R^2	1.298 ***	0.6933 *	0.9399 **	1.121 **	-0.3764	1.799 ***
		[2.56]	[1.31]	[1.65]	[1.88]	[-0.587]	[3.28]
(2)	Quote Count at (t,u+k)	-7.862 ***	-6.423 ***	-7.822 ***	-6.606 ***	-4.917 **	-11.86 ***
		[-7.48]	[-5.26]	[-4.11]	[-2.97]	[-1.78]	[-11.1]
(3)	Liquidity recovery rate at (t,u+k) x 100	0.01055	0.00609	-0.007077	0.008149	0.02976 ***	
		[1.05]	[0.681]	[-0.752]	[0.923]	[2.54]	
(4)	Depth at (t,u+k)	-0.1873	0.6437	1.306	0.6716	0.6772	
		[-0.174]	[0.53]	[0.901]	[0.406]	[0.365]	
(5)	|Surprise of news|	0.2227	0.9373	-0.2003	-0.7416	-0.3685	
		[0.229]	[0.923]	[-0.186]	[-0.65]	[-0.315]	
(6)	# of quote revision	-9.137 ***	-8.659 ***	-8.765 ***	-12.3 ***	-11.61 ***	
		[-7.63]	[-7.1]	[-6.58]	[-8.84]	[-7.42]	
(7)	Linear time trend	15.17 ***	28.37 ***	33.8 ***	30 ***	31.35 ***	17.96 ***
		[5.22]	[8.81]	[9.08]	[7.65]	[7.04]	[6.77]
(8)	Y2009	-2.368	4.581 **	3.781 *	6.314 **	6.461 **	
		[-1.02]	[1.75]	[1.33]	[2.14]	[2.13]	
(9)	Y2010	-2.121	-4.602 **	2.799	10.45 ***	8.965 ***	
		[-0.941]	[-2.03]	[1.08]	[3.8]	[2.85]	
(10)	Y2011	-5.92 ***	-1.073	3.451	4.672 *	1.087	
		[-2.58]	[-0.437]	[1.25]	[1.4]	[0.336]	
(11)	Y2012	-7.169 ***	-5.827 ***	-0.1205	1.378	2.526	
		[-3.45]	[-2.52]	[-0.0424]	[0.457]	[0.764]	
(12)	|OIB|/cumulative volume(t,u+k)x100	0.0897 ***	0.0332 *	0.0956 ***	0.1685 ***	0.2681 ***	
		[4.82]	[1.48]	[2.87]	[4]	[5.18]	
(13)	cumulative volume(t,u+k)	10.52 ***	10.7 ***	12.14 ***	13.1 ***	13.66 ***	
		[18.3]	[18.2]	[16.9]	[16.9]	[14.9]	
	constant	85.91 ***	70.51 ***	62.19 ***	87.07 ***	70.15 ***	49.56 ***
		[7.6]	[6.26]	[5.36]	[7.08]	[5.68]	[28.5]
	N	3017	3022	2905	2820	2737	3022
	R-squared	0.180	0.165	0.126	0.107	0.087	0.0513

Note: Independent variable (1) is logit transformation of R^2, variables (2), (4), (6), (13) are taken log. *, **, *** indicate significance at 10%, 5%, 1% level.

The dependent variable is a measure of price discovery for each announcement time t, defined as

$$PD^R_{t,k,T=600} = \frac{|\Delta S(t,u+k)|}{|\Delta S(t,u+T)|}.$$

This expression is a realized version of the price discovery measure (14). Independent variable (1) is a R-squared of following regression over *t*:

$$\Delta S(t,u+600) = constant + \alpha_i N_i(t,u) + \epsilon(t,u), i = 1,2,\cdots,\text{\# of different news}.$$

This is a news-by-news version of regression equation (13). LHS takes a value only when each news i is observed. This measures the information held by informed traders. We estimate the regression equation by a rolling regression with window of 12 periods. This procedure generates R-squared for each *i* at each *t*. For each *t*, we employ a max of such R-squared over *i* for the independent variable (1). (2) quote count (t,u+k) is the number of traders submitting limit order at (t,u+k). This is a proxy of the number of traders inhabited with the market at each time. (6) the number of quote revision plus and minus 10 minutes around the announcement. (7) time trend is the linear trend from 2006 to 2017, scaled by N. Year-by-year dummies are introduced for four years after the global financial crisis.

Table 5: Regression results for the determinants of speed of liquidity recovery

	$y = LR^R_{t,k,T=300}$	k=5sec	k=10sec	k=30sec	k=60sec	k=120sec	k=10sec
(1)	R^2	-1.201 **	-0.754 *	-0.4087	-0.6816 *	-1.267 ***	-0.7593 *
		[-2.31]	[-1.55]	[-0.786]	[-1.44]	[-2.55]	[-1.55]
(2)	Quote count at (t,u+k)	22.87 ***	17.52 ***	16.21 ***	13.72 ***	7.381 ***	15.85 ***
		[24.4]	[16.5]	[11.2]	[7.03]	[3.5]	[17]
(3)	PD rate at (t,u+k) x 100	0.01586 *	0.01411 *	0.02576 ***	0.01276 *	0.02112 ***	
		[1.49]	[1.33]	[2.63]	[1.41]	[2.41]	
(4)	Depth at (t,u+k)	-5.618 ***	-4.976 ***	-2.973 **	-3.725 ***	-2.076 *	
		[-5.22]	[-4.51]	[-2.27]	[-2.59]	[-1.42]	
(5)	Surprise of news	-0.887	0.3883	1.474 *	1.435 *	-0.4578	
		[-0.913]	[0.398]	[1.56]	[1.61]	[-0.5]	
(6)	# of quote revision	-1.611 *	-2.904 ***	-1.676 *	-2.196 **	-2.884 ***	
		[-1.43]	[-2.63]	[-1.51]	[-1.96]	[-2.34]	
(7)	Linear time trend	-25.9 ***	-25.31 ***	-24.66 ***	-16.27 ***	-6.57 **	-20.72 ***
		[-10.3]	[-8.81]	[-8]	[-5.15]	[-1.9]	[-9.71]
(8)	Y2009	-0.2331	-1.002	-7.944 ***	-4.998 **	-5.014 **	
		[-0.102]	[-0.402]	[-3.29]	[-2.18]	[-2.09]	
(9)	Y2010	4.327 **	5.465 ***	-2.361	-2.657 *	-2.636 *	
		[2.02]	[2.6]	[-1.19]	[-1.33]	[-1.32]	
(10)	Y2011	-6.585 ***	-5.94 ***	-4.889 **	-3.694 *	-3.782 *	
		[-2.62]	[-2.36]	[-1.81]	[-1.36]	[-1.37]	
(11)	Y2012	-10.26 ***	-9.991 ***	-7.148 ***	-5.675 **	-8.284 ***	
		[-4.39]	[-4.48]	[-2.67]	[-2.14]	[-3.03]	
(12)	|OIB|/cumulative volume(t,u+k)x100	0.01856	0.02847 *	0.03055	-0.06997 **	-0.08433 **	
		[0.975]	[1.36]	[1.06]	[-2.04]	[-2.08]	
(13)	cumulative volume(t,u+k)	0.1053	0.1449	-0.2615	-0.4121	-0.04158	
		[0.18]	[0.255]	[-0.405]	[-0.623]	[-0.0558]	
	constant	85.46 ***	101.4 ***	89.13 ***	101.4 ***	108.2 ***	58.53 ***
		[8.04]	[10.4]	[9.07]	[10.8]	[11.3]	[38.6]
	N	2904	2859	2673	2660	2553	2859
	R-squared	0.206	0.118	0.068	0.037	0.022	0.102

Note: Independent variable (1) is logit transformation of R^2, variables (2), (4), (6), (13) are taken log. *, **, *** indicate significance at 10%, 5%, 1% level.

The dependent variable is a measure of liquidity recovery for each announcement time t, defined as

$$LR^R_{t,k,T=300} = \left\{\frac{\log\big(ask(t,u+k)\big) - \log\big(bid(t,u+k)\big)}{\log\big(ask(t,u+T)\big) - \log\big(bid(t,u+T)\big)}\right\}^{-1}.$$

In fact, this is an element of liquidity recovery measure (14). The definition of independent variables is the same as Table 4.

Table 6: Regression results for the determinants of speed of price discovery (tick-time)

	$y = PD^R_{t,\mathrm{n},T=600}$	n=5tick	n=50tick	n=100tick	n=200tick	k=100tick
(1)	R^2	0.6585 *	-0.02922	0.8013 *	1.204 **	1.317 **
		[1.51]	[-0.0507]	[1.31]	[1.9]	[2.27]
(2)	Quote Count at (t,u+k)	-10.57 ***	-11.05 ***	-11.71 ***	-10.26 ***	-8.358 ***
		[-12.2]	[-8.32]	[-6.41]	[-4.47]	[-7.68]
(3)	Liquidity recovery rate at (t,u+k) x 100	-0.008865 *	0.0111 *	0.009928 *	0.002355	9.237 ***
		[-1.32]	[1.46]	[1.4]	[0.254]	[3.48]
(4)	Depth at (t,u+k)	4.099 ***	2.338 **	1.509	2.654 *	53.76 ***
		[4.12]	[1.78]	[0.989]	[1.52]	[25.1]
(5)	Surprise of news	1.157	1.431	2.051 **	0.1411	2945
		[1.24]	[1.27]	[1.74]	[0.123]	0.023046
(6)	# of quote revision	-8.784 ***	-8.424 ***	-9.876 ***	-14.01 ***	
		[-8.89]	[-6.37]	[-7.17]	[-10.8]	
(7)	Linear time trend	10.17 ***	32.44 ***	35.24 ***	30.12 ***	9.237 ***
		[3.8]	[9.28]	[9.73]	[7.2]	[3.48]
(8)	Y2009	-2.176	1.487	-0.1288	6.214 **	
		[-1.02]	[0.565]	[-0.0476]	[2.15]	
(9)	Y2010	-3.031 **	-1.743	1.419	6.099 **	
		[-1.71]	[-0.667]	[0.534]	[2.05]	
(10)	Y2011	-0.5391	4.01 *	6.045 **	5.144 *	
		[-0.299]	[1.43]	[1.97]	[1.61]	
(11)	Y2012	-0.5458	-5.914 ***	-0.8906	0.4427	
		[-0.293]	[-2.5]	[-0.305]	[0.156]	
(12)	|OIB|/cumulative volume(t,u+k)	0.2529 ***	0.06272 ***	0.03536 ***	0.01581 ***	
		[3.16]	[9.31]	[9.95]	[7.6]	
(13)	cumulative volume(t,u+k)	-0.07785 ***	-0.3662	-1.321 **	2.391	
		[-4.57]	[-0.979]	[-1.76]	[0.773]	
	constant	91.24 ***	99.99 ***	123.9 ***	164.5 ***	53.76 ***
		[9.31]	[7.72]	[8.88]	[12.1]	[25.1]
	N	3049	2945	2883	2761	2945
	R-squared	0.136	0.108	0.093	0.085	0.0230

Note: Independent variable (1) is logit transformation of R^2, variables (2), (4), (6), (13) are taken log. *, **, *** indicate significance at 10%, 5%, 1% level.

The dependent variable is a measure of price discovery for each announcement time *t*, defined as

$$PD^R_{t,n,T=600} = \frac{|\Delta S(t, u+n)|}{|\Delta S(t, u+T)|}.$$

Contrary to the specification in Table 4, we use tick-time to define the speed of price discovery. The numerator of the measure is indexed by the tick-time *n*, instead of the physical time *k*. Other specifications are identical to that of Table 4.

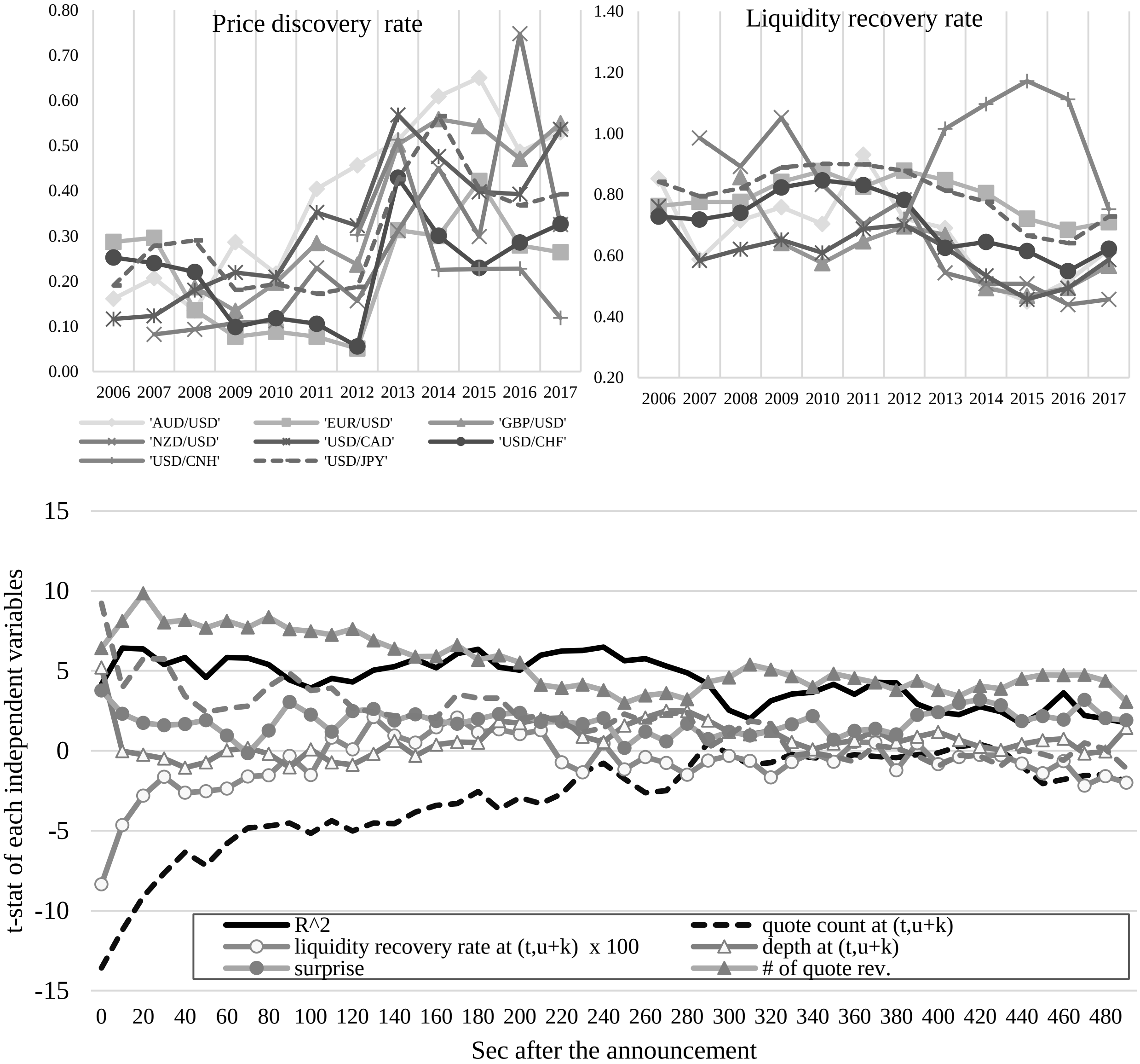


Figure 4: Panel regression (t-stat in different k) and yearly changes in market quality

Top Panels Yearly changes in price discovery rate (k=40, T=1800) and liquidity recovery rate (k=5, T=60) for each eight currency pairs. The way of construction is the same as Table 1.

Bottom Panel In addition to EUR/USD, other seven currency pairs are taken account: AUD/USD, GBP/USD, NZD/USD, USD/CAD, USD/CHF, USD/CNH, USD/JPY. Fixed effect panel regression is employed. The dependent variable is a measure of price discovery for each announcement time *t*, defined as $\frac{|\Delta S(t,u+k)|}{|\Delta S(t,u+T)|}$ where T=600.

The set of independent variables is the same as Table 4. R^2 measure (that of EUR/USD) and surprise of news are common, and other four variables are calculated for each pair. The panel plots the t-statics of independent variables in different k.

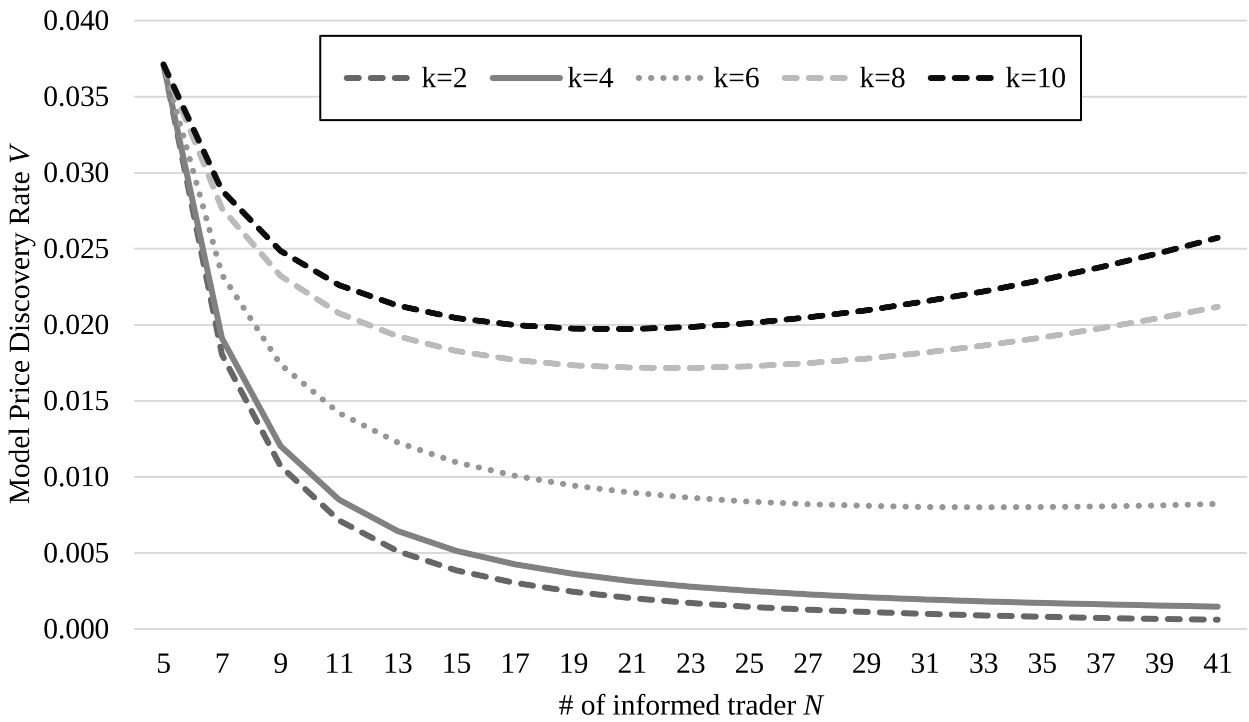


Figure 5: Model Simulation of Price Discovery Rate as a Function of the number of Informed Trader *N*

This plot is the variance ratio $V(N) \equiv \mathrm{Var}(p(N) - \lambda(N)\sigma_z^2)/\mathrm{Var}(\tilde{v})$, a measure of price discovery, with the informed trader's signal precision τ_e shifting from the monopolistic competition (when k is small) toward informed competition (when k is large). Specifically,

$$\tau_e(N,k) = x(k)\frac{\tau_E}{N_{initial}} + \left(1 - x(k)\right)\frac{\tau_E}{N}, x(k) = \frac{1}{1 + \exp(-(k - k_c))}.$$

The downward sloping curve is consistent with our empirical findings.

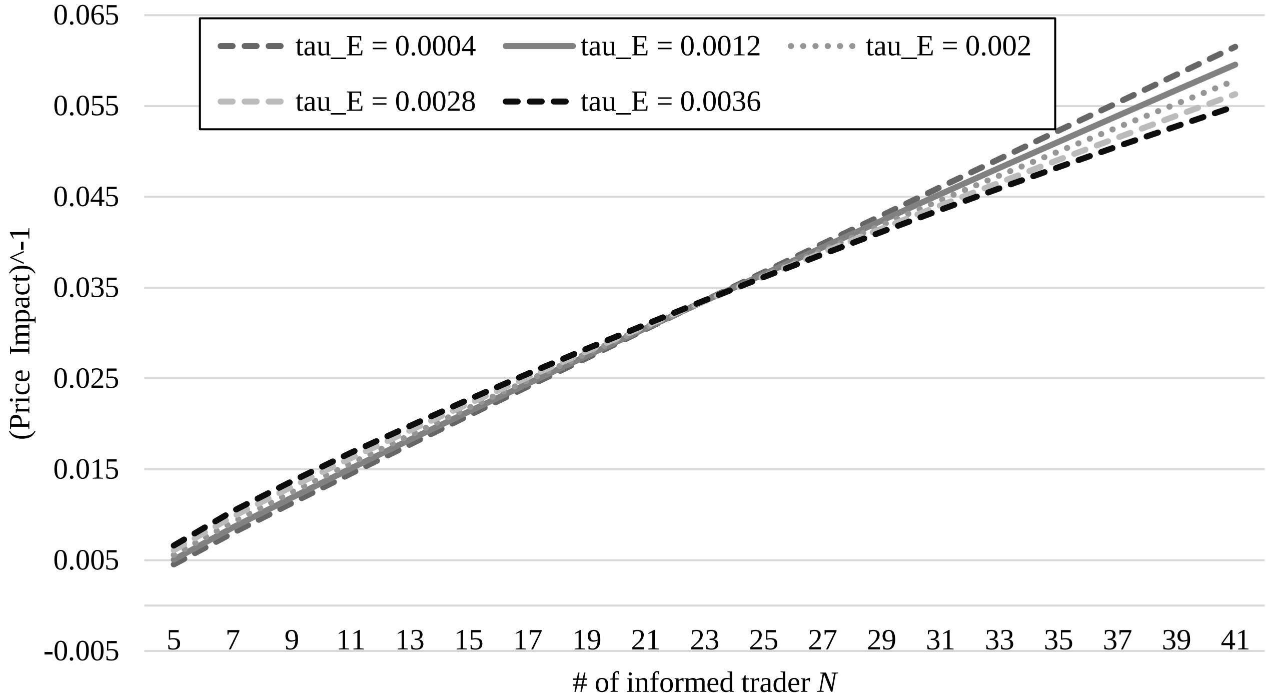
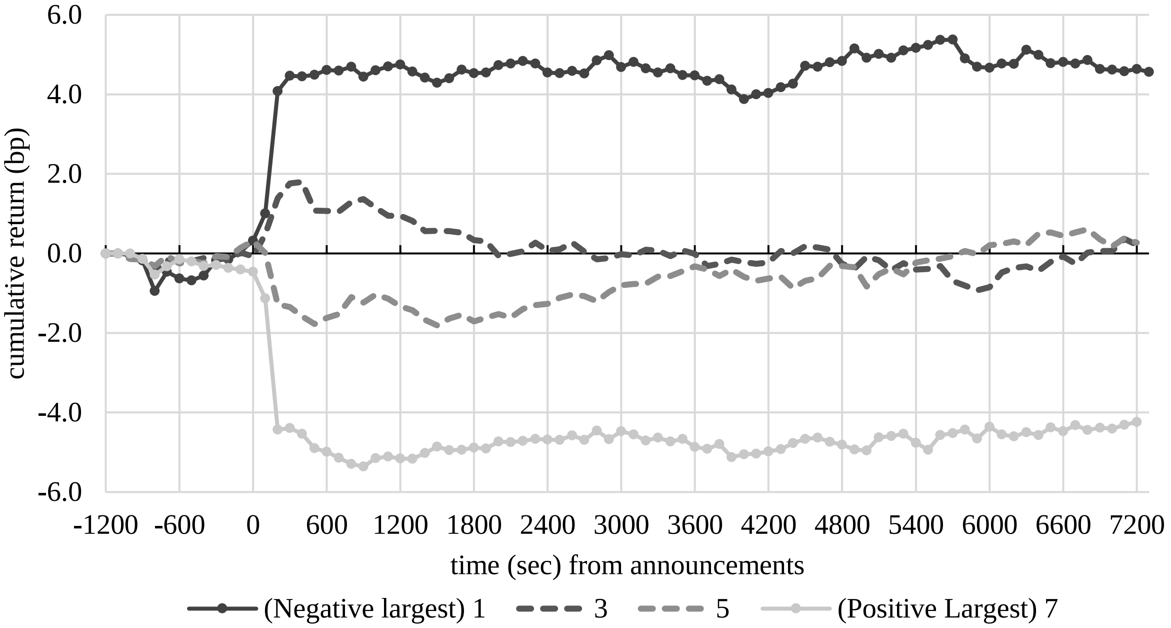


Figure 6: Model Simulation of Liquidity as a Function of the Number of Informed Traders

This plot is the inverse of price impact $\lambda(N)$, a measure of liquidity, as a function of N, with different τ_E. Upward sloping curve is consistent with our empirical findings. Increasing τ_E of reducing liquidity, our empirical findings, is observed in the region of sufficiently high number of informed traders.



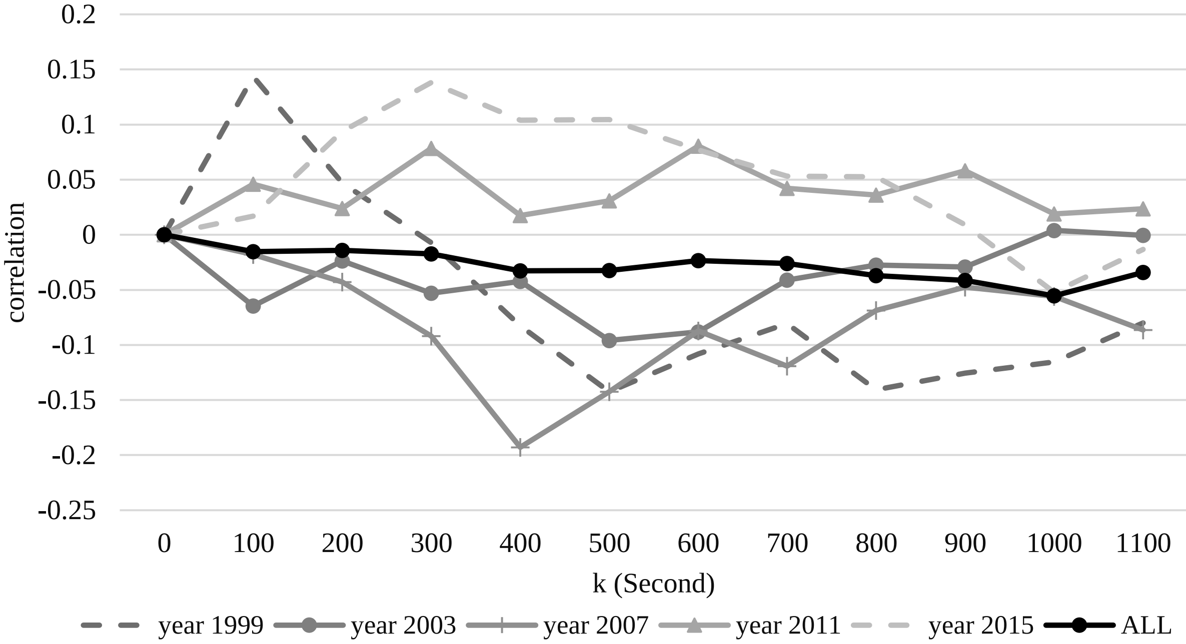


Figure 7: Predictability of return after the announcements

Top figure The figure plots the average cumulative return over two hours (+ 20 minutes) surrounding macro announcements. Each plot is an average of sample stratified by the news surprise defined as (12). Bin 1 represents the largest negative surprise (i.e., depreciates USD and earns positive returns on EUR/USD) and Bin 7 represents the largest positive surprise (i.e., appreciates USD and earns negative returns on EUR/USD). Bin 4 consists of zero surprise sample.

Bottom figure The figure plots the predictability of return after announcements, measured by the autocorrelation of returns defined as $Corr\big(s(t,u+k)-s(t,u), s(t,u+3\text{hours})-s(t,u+k)\big), k = 0, 100sec, 200sec, \ldots$, where t denotes announcement days and u denotes intraday announcement time. Each plot is grouped by years.