Strategic Cross-Trading in the U.S. Stock Market

Paolo Pasquariello and Clara Vega[1]

October 2, 2008

Abstract

We provide a theory and novel empirical evidence of *cross-price impact* — the permanent impact of informed trades in one asset on the prices of other (either related or fundamentally unrelated) assets — in the U.S. stock market. To guide our analysis, we develop a parsimonious model of multi-asset trading in the presence of two realistic market frictions — information heterogeneity and imperfect competition among informed traders — but in which extant channels of trade and price *co*-formation in the literature are ruled out by construction. In that setting, we show cross-price impact to be the equilibrium outcome of strategic trading activity of risk-neutral speculators across many assets to mask their information advantage about some other assets. We find strong, robust evidence of cross-asset informational effects in a comprehensive sample of the trading activity in NYSE and NASDAQ stocks between 1993 and 2004: Net order flow in one industry or random stock has a significant, persistent, and robust impact on daily returns of other industries or random stocks. Our empirical analysis further indicates that, consistent with our stylized model, both *direct* (i.e., an asset's own) and absolute cross-price impact are *i)* smaller when speculators are more numerous; *ii)* greater when marketwide dispersion of beliefs is higher; *iii)* greater among stocks dealt by the same specialist; and *iv)* smaller when macroeconomic news of good quality is released.

JEL classification: D82; G14; G15

Keywords: Equity Market; Market Liquidity; Strategic Trading; Information Heterogeneity; Public News

1   Introduction

There is a large body of research relating frictions to investors' trading activity — such as liquidity, transaction costs, financing and short-selling constraints, information asymmetry and heterogeneity — to stock prices.[1] Within this literature, the process of price *co*-formation in equity markets remains a not well-understood issue.[2] Yet, it is a crucial issue since, e.g., the extent of stock return comovement affects the benefits of portfolio diversification. If equity markets were frictionless, stock prices should immediately adjust to public news surprises. Hence, we should observe prices to comove only during announcement times when news of common information content is released (e.g., King and Wadhwani, 1990).[3] However, stock prices comove significantly during both non-announcement days and announcement times of unrelated news as well.[4]

Our paper contributes to this debate by providing a theory and novel empirical evidence *i)* of the impact of trading activity in one stock on the prices of other, either related or fundamentally unrelated, stocks; and *ii)* of the link between such impact and the number of and dispersion of beliefs among informed traders (henceforth, speculators), as well as the availability and quality of public fundamental information. The basic intuition of our theory — a multi-asset extension of Kyle (1985) based upon Caballé and Krishnan (1994) and Pasquariello (2007) — is as follows. In an interconnected economy, uninformed market-makers (henceforth, MMs) attempt to learn about the liquidation value of one asset from order flow in other assets; thus, speculators, when better-informed about an asset, trade strategically across many assets (even unrelated ones) to attenuate the ensuing dissipation of their information advantage; in equilibrium, MMs' *cross-inference* and speculators' *cross-trading* lead to *cross-asset liquidity* (i.e., cross-price impact), even among unrelated assets.[5] In this setting, we show that both direct and absolute cross-

---

[1]See, e.g., Amihud and Mendelson (1986), Constantinides (1986), De Long et al. (1990), Brennan and Subrahmanyam (1996), Brennan et al. (1998), Vayanos (1998), Vayanos and Vila (1999), Shleifer (2000), Amihud (2002), Huang (2003), Pastor and Stambaugh (2003), Acharya and Pedersen (2005), Duffie et al. (2005, 2006), Sadka and Scherbina (2007).

[2]Recent exceptions — motivated by the spillover of financial crises of local origins across the world financial markets during the 1990s — are Calvo (1999), Kyle and Xiong (2001), Kodres and Pritsker (2002), Yuan (2005), Veldkamp (2006), and Pasquariello (2007).

[3]Accordingly, Lo and MacKinlay (1990), Brennan et al. (1993), Chan (1993), McQueen et al. (1996), Chordia and Swaminathan (2000), and Chordia et al. (2008) attribute the evidence on positive cross-autocorrelations among stock returns to lagged transmission of common information.

[4]E.g., see Campbell et al. (2001), Hasbrouck and Seppi (2001), Kumar and Lee (2006), Green and Hwang (2008), Greenwood (2008), and Kallberg and Pasquariello (2008).

[5]In a related vein, Pasquariello (2007) and Bernhardt and Taub (2008) show that the presence of strategic speculators internalizing the extent to which their trades influence either observable order flow or observable

---

asset liquidity are increasing in the number of speculators, decreasing in the heterogeneity of their private information, and smaller in the presence of public signals (especially when of high quality), for those factors affect the extent of (and compensation for) adverse selection risk for the MMs.

We test our model's implications in the U.S. stock market by analyzing the Trades and Automated Quotations (TAQ) database — the most comprehensive sample of the equity trading activity in the New York Stock Exchange (NYSE) and the National Association of Securities Dealers Automated Quotation System (NASDAQ) — between 1993 and 2004. For the sake of parsimony (given the large number of stocks in each year of the sample period), we separately concentrate on ten industry-sorted stock portfolios and a large number of random stock pairs, and assess the intensity of their fundamental relationships by means of the economic and statistical significance of the correlation of their quarterly earnings. Our empirical analysis provides strong evidence of the *informational* role of trading for the process of price co-formation in the U.S. equity markets as advocated by our theory.

First, we show that measures of *permanent* cross-industry and cross-stock price impact are both economically and statistically significant, even among the least related industries and stocks. For instance, we estimate that a one standard deviation shock to order flow in HighTech stocks increases daily Energy stock returns by an average of roughly 28 basis points, although the historical correlation between those two industries' earnings is statistically indistinguishable from zero. Accordingly, we find that a one standard deviation shock to the order flow in one randomly selected stock moves the returns of another randomly selected stock more often than if due to chance (i.e., to statistical Type I error), on average by no less than 13 basis points (versus an average of 43 basis points in correspondence with a similar shock to that stock's own order flow), and even within quintiles of random stock pairs with absolute earnings correlation of nearly zero. This evidence of cross-asset informational effects — robust to controlling for marketwide trading activity and price fluctuations, inventory considerations, and any public direct and cross-asset information already embedded in past prices — provides indirect support for our theory.

Further, more direct support for our theory comes from testing its unique predictions stemming from speculators' informed and strategic cross-trading activity. In particular, we document that, consistent with our stylized model, direct and cross-asset liquidity are lower when speculators are less numerous, when various measures of marketwide dispersion of beliefs among them are higher, and among stocks dealt by the same specialist; our estimates of direct and cross-

_____

prices may also enhance equilibrium price comovement in the multi-asset noisy rational expectations models of Caballé and Krishnan (1994) and Admati (1985), respectively.

asset liquidity are instead higher in days when macroeconomic news — a trade-free source of marketwide information for the MMs — is released, especially if of good quality. For example, we find that Telecom stock returns increase by an average of either 79 or 93 basis points in correspondence with a one standard deviation shock to order imbalances in Nondurables stocks when either the number of speculators is low or the dispersion of their beliefs is high, while being insensitive to trading activity in those stocks otherwise. Similarly, we find that when information heterogeneity in the U.S. equity markets is high, the daily returns of a randomly selected stock move on average by almost 36 basis points more than when marketwide information heterogeneity is low in correspondence with a one standard deviation shock to its own order flow, and by 2 basis points more in correspondence with a one standard deviation shock to the order flow of another randomly selected stock. We also find that the effect of the availability of macroeconomic news about the U.S. economy, when statistically significant and consistent with our model, is to attenuate absolute cross-industry permanent price impact by an average of 23 basis points in response to one standard deviation shocks to cross-industry order flow. This additional evidence is robust to controlling for alternative channels of trade and price co-formation in the literature (inventory management, correlated information, portfolio rebalancing, correlated liquidity, and price observability) ruled out from our model by construction.

This work is related to two recent studies. Hartford and Kaul (2005) find evidence of strong common effects in order flow and returns among S&P500 stocks and, via the estimation of cross-trading regressions, attribute most of the observed return commonality to order flow commonality. Within a multi-asset extension of Grossman and Miller (1988), Andrade et al. (2008) argue that the hedging needs of risk-averse liquidity providers may lead to cross-asset price impact of *non-informational*, inelastic trading, if asset payoffs are correlated, despite the absence of cross-trading. Using data from margin accounts set up by individual investors with local brokerage firms in the Taiwan Stock Exchange (TSE), Andrade et al. (2008) find support for this implication by showing that individual weekly stock returns are more positively related to trading imbalances in more related industry portfolios. Our analysis differs from those aforementioned studies for we investigate, both theoretically and empirically (using transaction-level data), the properties of cross-asset liquidity in the U.S. stock market in the presence of informational, strategic cross-trading, even when asset payoffs are uncorrelated.

The paper is organized as follows. In Section 2, we construct a stylized model of multi-asset trading to guide our empirical analysis. In Section 3, we describe the data. In Section 4, we present the empirical results. We conclude in Section 5.

2 Theoretical Model

In this section we motivate our investigation of the impact of *i)* the dispersion of beliefs among sophisticated market participants and *ii)* the release of fundamental news on the informational role of trading in the U.S. equity market. We first describe a simple extension of the multi-asset model of trading of Caballé and Krishnan (1994) and Pasquariello (2007) and derive closed-form solutions for the equilibrium market depth, trading volume, and price volatility.[6] Then, we enrich the model by introducing public signals and consider its implications for the equilibrium price and trading strategies. All proofs are in Appendix A.

2.1 Benchmark: No Public Signals

The model is based on Kyle (1985). It consists of a three-date, two-period economy in which N risky assets are exchanged. Trading occurs only at the end of the first period $(t = 1)$. At the end of the second period $(t = 2)$, the payoffs of the risky assets, a $N \times 1$ multivariate normally distributed (MND) random vector v with mean P_0 and nonsingular covariance matrix Σ_v, are realized. The economy is populated by three types of risk neutral traders: a discrete number (M) of informed traders (that we label speculators), liquidity traders, and perfectly competitive market-makers (MMs). All traders know the structure of the economy and the decision process leading to order flow and prices.

At time $t = 0$ there is neither information asymmetry about v nor trading. Sometime between $t = 0$ and $t = 1$, each speculator m receives a private and noisy signal of v, S_{vm}. We assume that each vector S_{vm} is drawn from a MND with mean P_0 and covariance matrix Σ_s and that, for any two S_{vm} and S_{vk}, $cov(v, S_{vm}) = cov(S_{vm}, S_{vk}) = \Sigma_v$. We further parametrize the degree of diversity among speculators' private information by imposing that $\Sigma_s = \frac{1}{\rho}\Sigma_v$ and $\rho \in (0, 1)$. These assumptions imply that each speculator's information advantage about v at $t = 1$, before trading with the MMs, is given by

$$\delta_m \equiv E(v|S_{vm}) - P_0 = \rho(S_{vm} - P_0), \tag{1}$$

where $var(\delta_m) \equiv \Sigma_\delta = \rho\Sigma_v$ is nonsingular. It then follows that any two δ_m and δ_k have a joint multivariate normal distribution and $cov(\delta_m, \delta_k) \equiv \Sigma_c = \rho\Sigma_\delta$, a symmetric positive definite (SPD) matrix. Therefore, $E(\delta_i|S_{vm}) = \rho\delta_m$ and ρ can be interpreted as the correlation

---

[6]Holden and Subrahmanyam (1992), Foster and Viswanathan (1996), Back, Cao, and Willard (2000), and Pasquariello and Vega (2007) develop one-asset extensions of Kyle (1985) to link market liquidity and price volatility to the trading activity of heterogeneously informed traders.

between any two information endowments δ_m and δ_j: The lower (higher) is ρ, the more (less) *heterogeneous* — i.e., the less (more) correlated and of course precise — is speculators' private information about v.

At $t = 1$ both speculators and liquidity traders submit their orders to the MMs, before the price vector P_1 has been set. We define the vector of market orders of speculator m to be X_m. Thus, her profit is given by $\pi_m (X_m, P_1) = X'_m (v - P_1)$. Liquidity traders generate a vector of random demands z, MND with mean $\underline{0}$ (a zero vector) and nonsingular covariance matrix Σ_z. For simplicity, we impose that noise trading z has identical variance and is independent across markets ($\Sigma_z = \sigma_z^2 I$) as well as from any other random vector.[7] MMs do not receive any information, but observe the aggregate order flow for each asset $\omega_1 = \sum_{m=1}^{M} X_m + z$ from all market participants and set the market-clearing prices $P_1 = P_1 (\omega_1)$.

2.1.1 Equilibrium

Consistently with Caballé and Krishnan (1994), we define a Bayesian Nash equilibrium of this economy as a set of $M + 1$ vector functions $X_1 (\cdot), \dots, X_M (\cdot)$, and $P_1 (\cdot)$ such that the following two conditions hold:

1. *Profit maximization*: $X_m (S_{vm}) = \arg \max E (\pi_m | S_{vm})$;

2. *Semi-strong market efficiency*: $P_1 (\omega_1) = E (v | \omega_1)$.

We restrict our attention to linear equilibriums. We first conjecture general linear functions for the pricing rule and speculators' demands. We then solve for their parameters satisfying conditions 1 and 2. Finally, we show that these parameters and those functions represent a rational expectations equilibrium. The following proposition characterizes such an equilibrium, similarly to Pasquariello (2007, Proposition 1).

Proposition 1 *There exists a unique symmetric linear equilibrium given by the price function*

$$P_1 = P_0 + \Lambda \omega_1 = P_0 + \frac{1}{2 + (M - 1) \rho} \sum_{m=1}^{M} \delta_m + \Lambda z \tag{2}$$

and by each speculator m's demand strategy

$$X_m = \frac{1}{2 + (M - 1) \rho} \Lambda^{-1} \delta_m, \tag{3}$$

---

[7]Bernhardt and Taub (2008) explore the implications of correlated liquidity trading across assets for price and order flow commonality.

where

$$\Lambda = \frac{\sqrt{M\rho}}{[2 + (M-1)\rho]\sigma_z}\Sigma_v^{1/2} \tag{4}$$

is a SPD matrix.

The optimal trading strategy of each speculator depends on the private information she receives about v (δ_m) as well as on the depth of the market (Λ^{-1}). These speculators are imperfectly competitive and so, albeit risk neutral, exploit their information advantage in each market cautiously $(|X_m(n)| < \infty)$ to avoid dissipating their informational advantage with their trades, as in the single-asset setting of Kyle (1985). For the same purpose, these speculators also trade strategically across assets. Intuitively, the MMs know the structure of the economy (the covariance matrix Σ_v). Hence, unless all securities' terminal payoffs are fundamentally unrelated (i.e., unless Σ_v is diagonal), they rationally use the order flow for each asset to learn about the liquidation values of other assets when setting the market-clearing price vector P_1. The speculators are aware of this learning process, that Pasquariello (2007) labels *cross-inference*. Thus, they strategically place their trades in many assets — rather than independently trading in each asset — to limit the amount of information divulged by their market orders.

As a result of this effort, that Pasquariello (2007) labels *cross-trading*, Eqs. (2) and (3) represent a noisy rational expectations equilibrium.[8] Further, the equilibrium market liquidity matrix Λ of Eq. (4) is non-diagonal, reflecting both speculators' strategic trading activity $(\frac{\partial X_m(n)}{\partial \delta_m(j)} \neq 0)$ and MMs' attempt to be compensated for the losses they anticipate from it by affecting their expected profits from noise trading $(\frac{\partial P_1(n)}{\partial \omega_1(j)} \neq 0)$.

2.1.2 Testable Implications

The parsimonious multi-asset equilibrium of Eqs. (2) and (3) generates unambiguous predictions on direct $(\Lambda(n,n))$ and cross-asset $(\Lambda(n,j))$ liquidity. In the basic model of Section 2.1, speculators are risk neutral, financially unconstrained, and formulate "fundamentally correct" inference from their private signals $(\frac{\partial \delta_m(n)}{\partial S_{vm}(j)} = 0$ if $\Sigma_v(n,j) = 0)$. Hence, neither correlated information shocks (King and Wadhwani, 1990), correlated liquidity shocks (Calvo, 1999; Kyle and Xiong, 2001; Yuan, 2005; Bernhardt and Taub, 2008), nor portfolio rebalancing (Kodres and Pritsker, 2002) drive their cross-trading decisions, as in Pasquariello (2007). Nonetheless, Proposition 1 implies that if the underlying economy is *fundamentally interconnected* (a non-diagonal Σ_v),

---

[8]This is in contrast with the equilibrium of the multi-asset model of trading with perfect competition of Admati (1985), where prices fully reveal all private information of risk-averse insiders when their risk aversion wanes.

order flow in one security has a contemporaneous impact on the equilibrium prices of many assets $(\Lambda(n,j) \neq 0)$ — even those whose terminal values are unrelated to that security's payoff. Such an impact is due to both rational belief updating by the MMs and strategic trading by the speculators. The speculators place market orders in several assets — even if *fundamentally unrelated* $(\Sigma_v(n,j) = 0)$ — to affect the MMs' inference from the observed aggregate order flow. The MMs use those assets' aggregate trades when setting the price of that security to account for the speculators' attempt.

Remark 1 *If the economy is fundamentally interconnected there exists cross-asset liquidity, even among fundamentally unrelated assets.*

The number of speculators (M) and the correlation among their private information (ρ) affect both direct and cross-asset liquidity. The intensity of competition among speculators influences their ability to attenuate the informativeness of the order flow in each security. A greater number of speculators trade more aggressively — i.e., their aggregate amount of trading is higher — in every asset since (imperfect) competition among them precludes any collusive trading strategy.[9] This behavior reduces the perceived intensity of adverse selection for the MMs in each market, thus leading to greater direct and absolute cross-asset liquidity (lower $\Lambda(n,n)$ and $|\Lambda(n,j)|$).

The heterogeneity of speculators' signals moderates their trading aggressiveness. When information is less correlated (ρ closer to zero), each speculator has some monopoly power on her signal vector, because at least part of it is known exclusively to her. Hence, they trade more cautiously — i.e., their absolute amount of trading is lower — in each asset to reveal less of their own information advantage δ_m.[10] This "quasi-monopolistic" behavior makes the MMs more vulnerable to adverse selection. However, the closer ρ is to zero the lower is the precision of each speculator's private signal of v (since $\Sigma_s = \frac{1}{\rho}\Sigma_v$), hence the less severe is adverse selection for the MMs in all markets. In the presence of few — thus already cautious — speculators (low M), the latter effect dominates the former and both direct and absolute cross-asset liquidity improve (lower $\Lambda(n,n)$ and $|\Lambda(n,j)|$) for lower ρ. In the presence of many — thus already competitive — speculators (high M), the former effect dominates the latter and both $\Lambda(n,n)$ and $|\Lambda(n,j)|$ increase for lower ρ. The following corollary summarizes these empirical implications of our model.

---

[9]For instance, in the limit, if M speculators were *homogeneously* informed — i.e, if $\rho = 1$ such that $\Sigma_s = \Sigma_v$, $S_{vm} = v$, and $X_m = \frac{\sigma_z}{\sqrt{M}}\Sigma_v^{-1/2}(v - P_0)$ — it would ensue that $\frac{\partial |MX_m|}{\partial M} = \frac{\sigma_z}{2\sqrt{M}}\left|\Sigma_v^{-1/2}(v - P_0)\right| > 0$.

[10]In particular, $\frac{\partial |X_m|}{\partial \rho} = \frac{\sigma_z}{2\sqrt{M\rho}}\left|\Sigma_v^{-1/2}(S_{vm} - P_0)\right| > 0$.

Corollary 1 *Direct and absolute cross-asset liquidity are increasing in the number of speculators and decreasing in the heterogeneity of their information (except in the presence of only a few of them).*

To gain further insight on these results, we construct a simple numerical example along the lines of Pasquariello (2007). Specifically, we assume that there are three assets in the economy ($N = 3$), that their liquidation values are related to each other by way of the baseline parametrization of Σ_v reported in Appendix B (Eq. (B-1)), and that $\sigma_z = 1$. According to Eq. (B-1), assets 1 and 3 are fundamentally unrelated ($cov\left[v\left(1\right), v\left(3\right)\right] = 0$) yet both exposed to asset 2 ($cov\left[v\left(1\right), v\left(2\right)\right] > 0$ and $cov\left[v\left(2\right), v\left(3\right)\right] > 0$). We then vary the parameter ρ to study the direct and cross-asset equilibrium liquidity of this economy with respect to private signal correlation. For that purpose, we focus on assets 1 and 3 and plot the resulting $\Lambda\left(1, 1\right)$ and $\Lambda\left(1, 3\right)$ in Figure 1 for $M = 5$ or 500.

As a result of speculators' strategic trading, aggregate order flow in asset 3 impacts the equilibrium price of asset 1, although their terminal payoffs are unrelated: $\Lambda\left(1, 3\right) \neq 0$ in both cases although $cov\left[v\left(1\right), v\left(3\right)\right] = 0$. In the presence of many ($M = 500$) speculators, greater information heterogeneity among them — albeit accompanied by poorer quality of their signals — intensifies such trading activity, thus worsening MMs' perceived adverse selection problems and both direct and cross-asset liquidity in every security (Figures 1B and 1D, respectively). This is also the case in the presence of only a few ($M = 5$) speculators, yet only when the quality of their private information is high (Figures 1A and 1C). Otherwise, when signal quality deteriorates (i.e., when ρ is lower), their cautious and strategic trading activity becomes a less significant adverse selection threat for the MMs, leading to greater direct and cross-asset depths for asset 1.

2.2 Extension: Public Signals

An important characteristic of most financial markets is the frequent release of news about the fundamentals of the securities there traded to the public. At scheduled and frequent intervals, companies report their earnings and government agencies announce data on the macroeconomy. No less frequently, unscheduled news about both is also made available to market participants when it occurs. There is a vast literature showing that the release of public information affects both the dynamics of asset prices and the liquidity of their trading venues.[11] In this paper, we

---

[11]An incomplete list of recent studies on these subjects includes Fleming and Remolona (1997), Brandt and Kavajecz (2004), Green (2004), and Pasquariello and Vega (2007, 2008) in the U.S. Treasury bond market,

are interested in the impact of such releases on direct and cross-asset liquidity as well as on their interaction with the heterogeneity of speculators' private information.

To address these issues, we extend the basic model of Section 2.1 by providing each player with an additional, common source of information about the risky assets before trading takes place. To our knowledge, the resulting theoretical analysis of the relationship between the strategic trading activity of heterogeneously informed, imperfectly competitive speculators, the availability and quality of public information, and market liquidity is novel to the literature.[12] Specifically, we assume that, sometime between $t = 0$ and $t = 1$, both the speculators and the MMs receive a vector of public and noisy signals, S_p, of the N assets' payoffs, v. This vector is MND with mean P_0 and variance $\Sigma_p = \frac{1}{\psi_p}\Sigma_v$, where the signal-to-noise parameter $\psi_p \in (0, 1)$ controls for the quality of the public signals. We further impose that $cov(S_p, v) = cov(S_p, S_{vm}) = \Sigma_v$.

The availability of S_p affects the level and improves the precision of the information of *all* market participants prior to trading at time $t = 1$, with respect to the economy of Section 2.1. The MMs' revised priors about the distribution of v are now given by $P_0^* \equiv E(v|S_p) = P_0 + \psi_p(S_p - P_0)$ and $\Sigma_v^* \equiv var(v|S_p) = (1 - \psi_p)\Sigma_v$. Therefore, each speculator's information advantage about v at $t = 1$, before trading with the MMs, is now given by

$$\delta_m^* \equiv E(v|S_{vm}, S_p) - E(v|S_p) = \rho^*(S_{vm} - P_0^*), \tag{5}$$

where $\rho^* = \rho\frac{1 - \psi_p}{1 - \rho\psi_p} < \rho$. The above assumptions also imply that $var(S_{vm}|S_p) = \frac{1 - \rho\psi_p}{\rho}\Sigma_v$ and $cov(S_{vm}, S_{vk}|S_p) = \Sigma_v^*$, hence that $var(\delta_m^*|S_p) \equiv \Sigma_\delta^* = \rho^*\Sigma_v^*$ is nonsingular, that any two δ_m^* and δ_k^* are jointly MND for which $cov(\delta_m^*, \delta_k^*|S_p) \equiv \Sigma_c^* = \rho^*\Sigma_\delta^*$, a symmetric positive definite (SPD) matrix, and that $E(\delta_k^*|S_{vm}, S_p) = \rho^*\delta_m^*$. We can interpret ρ^* as the true (hence lower) correlation between any two information vectors δ_m^* and δ_k^* when a public signal vector is available, and δ_m^* as the truly private (hence less correlated) component of speculator m's original private information endowment (δ_m). The ensuing unique linear equilibrium of this amended economy mirrors that of Proposition 1 and is summarized below.

Proposition 2 *When a public signal vector of v (S_p) is available, there exists a unique symmetric linear equilibrium given by the price function*

$$P_1 = P_0^* + \Lambda_p\omega_1 = P_0^* + \frac{1}{2 + (M - 1)\rho^*}\sum_{m=1}^{M}\delta_m^* + \Lambda_p z \tag{6}$$

Andersen and Bollerslev (1998) and Evans and Lyons (2002, 2007, 2008) in the foreign exchange market, Berry and Howe (1994) in the U.S. stock market, and Brenner et al. (2008) in the U.S. corporate bond market.

[12] Admati and Pfleiderer (1988), Foster and Viswanathan (1990, 1993), Diamond and Verrecchia (1991), Kim and Verrecchia (1994, 1997), and Pasquariello and Vega (2007) study the effect of the disclosure of public information on liquidity and price volatility in one-asset markets.

and by each speculator m's demand strategy

$$X_m = \frac{1}{2 + (M-1)\rho^*} \Lambda_p^{-1} \delta_m^*,$$ (7)

where

$$\Lambda_p = \frac{\sqrt{M\rho^*}}{[2 + (M-1)\rho^*]\sigma_z} \Sigma_v^{*1/2}$$ (8)

is a SPD matrix.

2.2.1 Additional Testable Implications

Pasquariello and Vega (2007) show that, in a one-asset Kyle (1985) setting, the availability of a public signal improves market liquidity. This is also the case in our multi-asset economy. Intuitively, a public signal vector of v makes the speculators' private information less valuable and their trading activity less cautious, while providing the MMs with a trade-free source of information. In equilibrium, these considerations attenuate adverse selection risk for the MMs in all markets, thus increasing both their direct and cross-asset depth, even among fundamentally unrelated assets ($\Sigma_v(n,j) = 0$) as in Section 2.1. Accordingly, the impact on market liquidity is stronger the better the quality of the available public signals (i.e., the higher is ψ_p) for the less valuable the private signal vectors of v (S_{vm}) become for the speculators.

Corollary 2 *The availability of a public signal vector of v increases both direct and absolute cross-asset liquidity, the more so the greater is the public signal's precision.*

In the simple economy of Appendix B (in Figure 1), both $\Lambda(1,1)$ and $|\Lambda(1,3)|$ decline in the presence of S_p (i.e., $\Lambda_p(1,1) < \Lambda(1,1)$ and $|\Lambda_p(1,3)| < |\Lambda(1,3)|$ for $\psi_p = 0.5$), the more so the less numerous (hence more cautions) the speculators are — i.e., the more so when $M = 5$ (Figures 1A and 1C) than when $M = 500$ (Figures 1B and 1D) — for the more valuable public information about v becomes for the MMs. The extent of this decline is also sensitive to the degree of information heterogeneity among speculators (ρ). As mentioned in Section 2.1, when ρ is low their private signals are not only highly heterogeneous (thus inducing caution in trading) but also less precise (thus less valuable for trading). In the presence of only a few speculators, the latter effect dominates the former, the adverse selection risk for the MMs is relatively low, hence the availability of a public signal of v is marginally less beneficial to them (e.g., $\Lambda(1,1) - \Lambda_p(1,1) > 0$ and $|\Lambda(1,3)| - |\Lambda_p(1,3)| > 0$ in Figures 1A and 1C are smaller) than if ρ were high. In the presence of many speculators and low ρ, the former effect dominates the latter, the adverse

10

selection risk for the MMs is relatively high, hence the availability of a public signal of v is marginally more beneficial to them (e.g., $\Lambda(1,1) - \Lambda_p(1,1) > 0$ and $|\Lambda(1,3)| - |\Lambda_p(1,3)| > 0$ are greater) than if ρ were high.

Remark 2 *The improvement in direct and absolute cross-asset liquidity due to the availability of a public signal vector of v is decreasing in the number of speculators and increasing in the heterogeneity of their information (except in the presence of only a few of them).*

3   Data Description

We test the implications of the model presented in the previous section in a comprehensive sample of U.S. stock market transaction-level data, firm-level characteristics, and U.S. macroeconomic announcements.

3.1   U.S. Stock Market Data

We use intraday, transaction-level data — trades (i.e., market orders) and quotes — during regular market hours (9:30 a.m. to 4:00 p.m.) for all stocks listed on the NYSE and the NASDAQ between January 1, 1993 and June 30, 2004 ($2,889$ trading days). We obtain this data from the NYSE's TAQ database. We exclude Real Estate Investment Trusts (REITs), closed-end funds, foreign stocks, American Depository Receipts (ADRs) since their trading characteristics might differ from those of ordinary equities (Chordia and Subrahmanyam, 2004; Boehmer and Wu, 2008), i.e., we concentrate exclusively on the trading activity in domestic common stocks with Center for Research in Security Prices (CRSP) share code 10 or 11. Corresponding daily price data comes from CRSP. Firm-level accounting information — e.g., quarterly earnings-per-share (EPS) — is from the COMPUSTAT database. Merging TAQ, CRSP, and COMPUSTAT data yields a sample of $3,773$ firms (unique identifiers) over our sample period.

We filter the TAQ data by deleting a small number of trades and quotes representing possible data error (e.g., negative prices or quoted depths) or with unusual characteristics (as listed in Bessembinder, 1999, footnote 5). We then sign intraday trades using the Lee and Ready (1991) procedure: *i)* if a transaction occurs above (below) the prevailing quote mid-point, we label it a purchase (sale); *ii)* if a transaction occurs at the quote mid-point, we label it a purchase (sale) if the sign of the last price change is positive (negative). As in Bessembinder (2003), we do not allow for a five-second lag between trade and quote reports and compare exchange quotes from NYSE (NASDAQ) exclusively with NYSE (NASDAQ) transaction prices — i.e., only consider

order flow taking place in the exchange where the stock is listed — since off-exchange quotations (e.g., from regional stock exchanges) rarely improve on the exchange quote (Blume and Goldstein, 1997).

Our model, a multi-asset extension of Kyle (1985), conjectures a relationship between a firm's stock price changes and both its own and other firms' aggregate order flow. Chordia and Subrahmanyam (2004, p. 486) observe that "the Kyle setting is more naturally applicable in the context of signed order imbalances over a time interval, as opposed to trade-by-trade data, since the theory is not one of sequential trades by individual traders." Accordingly, in this paper we define the net order flow in firm i on day t, $\omega_{i,t}$, as the estimated daily number of buyer-initiated trades minus the number of seller-initiated trades scaled by the total number of trades on day t as follows:

$$\omega_{i,t} = \frac{BUYNUM_{i,t} - SELLNUM_{i,t}}{BUYNUM_{i,t} + SELLNUM_{i,t}}, \tag{9}$$

where $BUYNUM_{i,t}$ and $SELLNUM_{i,t}$ are the aggregate daily number of purchases and sales, respectively (e.g., Chordia and Subrahmanyam, 2004; Boehmer and Wu, 2008). Jones et al. (1994) and Chordia and Subrahmanyam (2004) show that the total number of transactions has greater explanatory power for stock return fluctuations than trading volume. We divide the net order flow by the total number of trades in Eq. (9) to eliminate the impact of total trading activity (Chordia and Subrahmanyam, 2004). In unreported analysis, we find the inference that follows to be nonetheless robust to defining order imbalances as the net scaled trading volume (e.g., Jones et al., 1994) or to assigning the direction of trades via the Hasbrouck (1988, 1991) algorithm.

3.2   Information Heterogeneity

The model of Section 2 postulates a relationship between the intensity of cross-asset liquidity and the extent of *marketwide* information heterogeneity among speculators, ρ. In this paper, we use professional forecasts of individual stocks' future earnings and U.S. macroeconomic announcements to proxy for the beliefs of sophisticated market participants about traded assets' fundamentals. The standard deviation across professional forecasts is a commonly employed measure of aggregate and security-level information heterogeneity unrelated to risk (e.g., Lang and Lundholm, 1996; Diether et al., 2002; Green, 2004; Moeller et al., 2006; Pasquariello and Vega, 2007, 2008; Kallberg and Pasquariello, 2008; Yu, 2008).

We obtain our first set of proxies for ρ by using the unadjusted I/B/E/S summary database of analyst forecasts of the long-term growth of individual stocks' EPS. Long-term growth forecasts

are less likely to be biased by firms' potential "earnings guidance" (Yu, 2008) and normalization for cross-firm comparability (Qu et al., 2004). The inference that follows is nonetheless robust to employing fiscal year EPS forecasts. We define the diversity of opinion about the long-term prospects of each firm i in the TAQ/CRSP/COMPUSTAT sample in each month m between January 1993 and June 2004 as the standard deviation across multiple (i.e., two or more) analyst forecasts of that firm's long-term EPS growth from the I/B/E/S Summary History database (when available), $SDLTEPS_{i,m}$.[13] Following Kallberg and Pasquariello (2008) and Yu (2008), we then compute our measure of marketwide information heterogeneity in month m, $SDLTEPS_m$, as a simple average of firm-level dispersion of opinion in that month,

$$SDLTEPS_m = \frac{1}{N_m} \sum\nolimits_{i=1}^{N_m} SDLTEPS_{i,m}, \tag{10}$$

where N_m is the total number of firms in month m. The equal-weighting scheme in Eq. (10) adjusts for the relatively poor coverage of small stocks in our merged TAQ/CRSP/COMPUSTAT dataset. We discuss this issue in greater detail in Section 4; our inference is nevertheless insensitive to computing $SDLTEPS_m$ as a value-weighted average of individual stock forecast standard deviations (labeled $VWSDLTEPS_m$). Yu (2008) shows that both $SDLTEPS_m$ and $VWSDLTEPS_m$ successfully capture the common component of differences in investors' opinions about individual stocks' future prospects in the U.S. equity markets.

Our second set of proxies for ρ is based upon the professional forecasts of 18 U.S. macroeconomic announcements from the International Money Market Services (MMS) Inc. real-time database, available exclusively between January 1993 and December 2000.[14] We use the standard deviation across those forecasts for each announcement p in each month m, $SDMMS_{p,m}$, to construct an alternative measure of the common dispersion of beliefs across speculators, as in Green (2004) and Pasquariello and Vega (2007, 2008). Specifically, we compute the aggregate degree of information heterogeneity about common macroeconomic fundamentals in month m, $SDMMS_m$, as a scaled simple average of normalized announcement-level dispersions in that

---

[13]Diether et al. (2002) describe similarities and differences between the I/B/ES Summary History and Detailed History databases. In unreported analysis, we find analogous results when obtaining analyst forecast data from the latter.

[14]Since being acquired by Informa in 2003, MMS discontinued its survey services. These surveyed announcements include quarterly (GDP Advance, GDP Preliminary, GPD Final), monthly (Nonfarm Payroll Employment, Retail Sales, Industrial Production, New Home Sales, Durable Goods Orders, Factory Orders, Construction Spending, Trade Balance, Producer Price Index, Consumer Price Index, Consumer Confidence Index, NAPM Index, Housing Starts, Index of Leading Indicators), and weekly news releases (Initial Unemployment Claims). Detailed discussions of the properties of this dataset can be found in Fleming and Remolona (1997) and Andersen et al. (2003).

month,

$$SDMMS_m = 10 + \sum_{p=1}^{18} \frac{SDMMS_{p,m} - \widehat{\mu}(SDMMS_{p,m})}{\widehat{\sigma}(SDMMS_{p,m})}, \tag{11}$$

where $\widehat{\mu}(\cdot)$ and $\widehat{\sigma}(\cdot)$ are the sample mean and standard deviation operators, respectively. The standardization in Eq. (11) is necessary because units of measurement differ across announcements, while shifting the mean of $SDMMS_m$ by a factor of 10 ensures that $SDMMS_m$ is always positive.

Figures 2a and 2b plot each of the measures for marketwide information heterogeneity in the U.S. equity markets described above over our sample period 1993-2004. Overall, these figures suggest that aggregate dispersion of beliefs is large (e.g. roughly 3.4% on average, when measured by the standard deviation of long-term EPS growth forecasts, versus an equal-weighted average of those forecasts of about 14.3%), time-varying, and positively correlated across different proxies, albeit not strongly so (with the exception of $SDLTEPS_m$ and $VWSDLTEPS_m$). Common disagreement is low in the mid 1990s, sharply increases and declines in correspondence with the Internet stock bubble, and stays historically high afterwards. These dynamics are consistent with those reported in recent studies employing similar proxies for information heterogeneity in the U.S. equity markets (e.g., Pasquariello and Vega, 2007; Kallberg and Pasquariello, 2008; Yu, 2008).

4   Cross-Price Impact in the U.S. Stock Market

The model of Section 2 generates several implications for the direct and cross-asset liquidity in the U.S. equity market that we now test in this section. In the context of our model, direct liquidity of any stock i is defined as the marginal *contemporaneous* impact of a trade in stock i on its equilibrium price, $\lambda_{ii,0}$. Similarly, cross-stock liquidity between any two stocks i and h is defined as the marginal contemporaneous impact of a trade in stock h on the equilibrium price of stock i, $\lambda_{ih,0}$. Ideally, we would need to estimate direct and cross-price impact for each of the stocks in our sample. This is a challenging task. When transaction-level data is available (as in our case), measures of direct price impact are typically estimated as the slopes of regressions of stock returns on the aggregate direct order flow (order imbalance) over either intraday or daily time intervals (e.g., see Hasbrouck, 2007). A natural and appealing extension to this procedure in our setting would be to assess the sensitivity of the returns of each stock to both *its own* and *each of all other* stocks' aggregate order flow (order imbalance). The large number of stocks in our database and the relative scarcity of trades for some of them makes the literal implementation

of such a route impractical.

In light of these considerations, we proceed by *i)* aggregating the stocks in our sample into a smaller number of industry portfolios, and *ii)* estimating direct and cross-price impact within this smaller subset of assets. We sort all firms in our merged TAQ/CRSP/COMPUSTAT dataset into either of the ten broad industry groupings proposed by Fama and French (1988): Durables (Cars, TVs, Furniture, Household Appliances), Nondurables (Food, Tobacco, Textiles, Apparel, Leather, Toys), Manufacturing (Machinery, Trucks, Planes, Chemicals, Office Furniture, Paper, Commercial Printing), Energy (Oil, Gas, Coal Extraction and Products), HighTech (Computers, Software, Electronic Equipment), Telecom (Business Equipment, Telephone and Television Transmission), Shops (Wholesale, Retail, Laundries, Repair Shops), Health (Healthcare, Medical Equipment, and Drugs), Utilities, and Other (Mines, Construction, Building Materials, Transportation, Hotels, Business Services, Entertainment, Finance).[15] We then compute daily, *equal-weighted* returns, $r_{n,t}$, for each of the resulting $n = 1, \ldots, 10$ industry portfolios,

$$r_{n,t} = \frac{1}{N_{n,t}} \sum_{i \in n} r_{i,t}, \tag{12}$$

where $r_{i,t}$ is firm i's daily close-to-close mid-point stock return (from TAQ) and $N_{n,t}$ is the number of firms in the industry portfolio n on day t, as well as daily aggregate, equal-weighted, industry-level net order flow, $\omega_{n,t}$,

$$\omega_{n,t} = \frac{1}{N_{n,t}} \sum_{i \in n} \omega_{i,t}. \tag{13}$$

where $\omega_{i,t}$ is firm i's estimated order flow on day t as defined in Section 3 (Eq. (9)). Chordia and Subrahmanyam (2004) document that contemporaneous price impact of daily order imbalance is increasing in firm size. Thus, employing aggregate, value-weighted order flow has the potential to favorably bias our analysis in a systematic way. Value-weighted industry portfolio returns, industry-level order flow, and all other industrywide and marketwide averages in our analysis lead to similar inference. We report summary statistics for daily industry-level price changes and net order flow in Table 1. The mean scaled order imbalance in number of transactions $\omega_{n,t}$ is positive for most industries — suggesting that buying pressure was predominant among market orders over our sample period — with two noteworthy yet unsurprising exceptions (HighTech and Utilities). Average firm-level net order flow $\omega_{i,t}$ is also positive (about 1.73%) and in line with previous studies (e.g., Chordia and Subrahmanyam, 2004; Boehmer and Wu, 2008).[16]

---

[15]The corresponding SIC codes are available on Kenneth French's research website at http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[16]As noted by Chordia and Subrahmanyam (2004), specialists maintaining a constant inventory accomodate

According to the model of Section 2, the strategic cross-trading activity of speculators leads to equilibrium cross-price impact of order flow, even among fundamentally unrelated assets, as long as the covariance matrix of their payoffs (Σ_v) is non-diagonal, i.e., as long as the underlying economy is fundamentally interconnected (see Remark 1). To assess the extent to which the ten industries listed above are fundamentally related over our sample period, we estimate (and report in Table 2) a corresponding correlation matrix of industrywide, equal-weighted averages of firm-level quarterly earnings. We proceed in two steps. First, for every firm i in our merged TAQ/CRSP/COMPUSTAT dataset we obtain its quarterly EPS ($EPS_{i,q}$, basic, excluding extraordinary items, for calendar quarter q) over our sample period 1993-2004 (when available). Second, we estimate Pearson correlations ($\rho_{n,j}$) of equal-weighted averages of $EPS_{i,q}$ within each industry n, $EPS_{n,q}$, defined as

$$EPS_{n,q} = \frac{1}{N_{n,q}} \sum_{i \in n} EPS_{i,q}, \tag{14}$$

where $N_{n,q}$ is the number of firms in industry n in calendar quarter q. Not surprisingly, Table 2 indicates that the U.S. economy, as represented by its stock market, is clearly fundamentally interconnected. Importantly, Table 2 also suggests that cross-industry earnings correlations are not uniformly high, positive, and statistically significant but vary significantly from the highest (0.774 between Manufacturing and HighTech) to the lowest (0.270 between Manufacturing and Nondurables).

As discussed above, the expression for the equilibrium prices of all assets in Proposition 1 translates naturally in the following set of $N = 10$ regression models:

$$r_{n,t} = \alpha_n + \sum_{j=1}^{10} \lambda_{nj,0} \omega_{j,t} + \varepsilon_{n,t}. \tag{15}$$

According to our model, we expect our estimates for direct price impact, $\lambda_{nn,0}$, to be positive and our estimates for cross-price impact, $\lambda_{nj,0}$ for $n \neq j$, to be significant. Even in the absence of the information-based, strategic cross-trading activity described in our model, inventory considerations (first formalized in a one-asset setting by Garman, 1976 and Amihud and Mendelson, 1980) may lead to statistically and economically significant correlation between price changes of our industry portfolios and cross-industry order flow. For instance, dealers' attempts to manage inventory fluctuations correlated across individual assets — because of marketwide dynamics in volatility, inventory carrying costs, or trading volume (e.g., Chordia et al., 2000; Chordia and Subrahmanyam, 2004; Comerton-Forde et al., 2008) — may eventually generate cross-price impact even when order imbalances have no information content. To assess the relevance of these

---

the excess buy or sell market orders for a firm's stock (i.e., nonzero means and medians for $\omega_{i,t}$) in the limit order book.

arguments for our inference, we include lagged values of direct and cross-net order flow in Eq. (15), in the spirit of Hasbrouck (1991). Hasbrouck (1991) argues that trades in an asset have a permanent, direct impact on its prices if due to information shocks, but a transitory impact if due to non-informational (e.g. liquidity or inventory-driven) shocks and other microstructure imperfections (e.g., see also Hasbrouck, 2007). Consistently, we assume that absolute cross-price impact may be deemed permanent if due to cross-inference (as advocated by our theory), but transitory otherwise. Hence, we interpret significant contemporaneous cross-order flow effects in Eq. (15), $\lambda_{nj,0}$, as *transient* if they are later reversed — i.e., if they are accompanied by lagged cross-impact ($\lambda_{nj,l}$) of same cumulative magnitude but opposite sign. On the other hand, we interpret significant estimates for $\lambda_{nj,0}$ as driven by *permanent* information effects (consistent with our model) if they are not subsequently reversed — i.e., if their estimated cumulative magnitude is either insignificant, of the same sign as $\lambda_{nj,0}$, or of the opposite sign but smaller than $\lambda_{nj,0}$.

Non-informational commonality in prices and trading activity may also lead to cross-price impact, even in the absence of the strategic, information-based cross-trading activity advocated by our theory. For instance, Chordia et al. (2000) observe that marketwide trading activity may be sensitive to general swings in stock prices. Hasbrouck and Seppi (2001) suggest that their evidence of common factors in the prices and order flows of the thirty stocks in the Dow Jones Industrial Average (DJIA) may be attributed to such marketwide liquidity shocks as portfolio substitution. Accordingly, Barberis et al. (2005) argue that investors' portfolio rebalancing activity may be triggered by non-informational shifts in the composition of broad categories and indexes (*category* investing) or other fixed subsets of all available securities (*habitat* investing). Alternatively, marketwide information shocks may also cause correlated trading and price changes (e.g., King and Wadhwani, 1990; Hasbrouck and Seppi, 2001). In Section 4.2, we examine in greater detail the potential impact of these and other alternative theories of cross-trading activity within the U.S. equity market on our inference. At this preliminary stage, we control for the extent of portfolio rebalancing and correlated, marketwide information-motivated trading by including daily equal-weighted returns on all NYSE and NASDAQ stocks in our sample ($r_{M,t}$) in Eq. (15), as well as by scaling firm-level net order flow by total trading volume ($\omega_{i,t}$ of Eq. (9)).[17]

---

[17]The inclusion of market returns also allows to reduce potential cross-correlations in error terms across industries (e.g., see Chordia and Subrahmanyam, 2004). The estimated coefficients on $r_{M,t}$, not reported here, are in line with those in the literature for similar industry portfolios (e.g., Table IX in Bernanke and Kuttner, 2005). We obtain qualitatively similar results when omitting $r_{M,t}$, when using market-adjusted industry returns ($r_{n,t} - r_{M,t}$), or when replacing $r_{M,t}$ with the three Fama-French factors to control for both a broader set of systematic sources of risk and other popular forms of category investing (e.g., small-cap versus large cap, or value versus growth).

The ensuing amended regression model for the estimation of direct and cross-industry liquidity in the U.S. equity market is given by:

$$r_{n,t} = \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10} \sum_{l=1}^{L} \gamma_{nj,l} r_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l} \omega_{j,t-l} + \varepsilon_{n,t}, \qquad (16)$$

where $L = 3$ trading days and $\sum_{l=0}^{L} \lambda_{nn,l}$ and $\sum_{l=0}^{L} \lambda_{nj,l}$ are measures of cumulative direct and cross-price impact, respectively.[18] Eq. (16) is based on the presumption — rooted in the theoretical and empirical microstructure literature — of a causal link from trades to price changes. As such, it also includes lagged returns of all industry portfolios in our sample to control for any public direct and cross-industry information already set in those portfolios' recent price change history, in the spirit of Hasbrouck (1991). Our inference is nonetheless robust to (and only weakened by) this inclusion. We efficiently estimate Eq. (16) for each of the ten industries *separately* by Ordinary Least Squares (OLS) — further correcting the standard errors for heteroskedasticity and serial correlation — and report the resulting estimates in Table 3.[19] Unreported analysis indicates that the corresponding adjusted R^2 (R_a^2) are higher than when replacing industry-level net scaled number of transactions ($\omega_{n,t}$) in Eq. (16) with industry-level net scaled trading volume, in line with Jones et al. (1994) and Chordia and Subrahmanyam (2004).

Table 3 provides strong evidence of direct and cross-asset informational effects of net industry order flow on daily industry portfolio returns. Consistent with both our model and previous studies (e.g., Chordia and Subrahmanyam, 2004; Boehmer and Wu, 2008), estimates of the permanent direct price impact of net scaled number of transactions are economically and statistically significant for each of the ten industries in our sample. For example, a one standard deviation shock to daily, direct, net order flow moves its corresponding industry's daily, market-adjusted returns by an average of 28.3 basis points. Importantly, and consistent with our model, nearly half of the estimates of the permanent cross-industry price impact of net scaled number of transactions are economically and statistically significant as well, even among the least fundamentally related industries. For instance, daily Energy returns increase by an average of 27.6 basis points in correspondence with a one standard deviation shock to daily net order flow in HighTech stocks — versus an average of 52 basis points in correspondence with a similar shock to its own order flow — although the historical correlation between quarterly earnings of these two industries is low and statistically insignificant (-0.173, in Table 2).

---

[18] The ensuing inference is qualitatively unaffected by employing more or fewer lags L in Eq. (16).

[19] *Joint* estimation of Eq. (16) by Feasible Generalized Least Squares (FGLS) leads to the same, efficient coefficient estimates since the resulting ten staked regression models have identical explanatory variables (e.g., Greene, 1997, p. 676).

In short, the evidence in Table 3 supports the notion, postulated by our theory, that there is permanent cross-price impact at the level of *industry* groupings of U.S. stocks. We intend to provide further evidence of cross-price impact in the U.S. equity markets at the level on the *individual* stocks, where non-informational and marketwide information-driven commonality in prices and trading activity is likely to be lower than for industry portfolios. Yet, this can be accomplished only once we make our large sample of U.S. stocks more manageable. We do so by repeatedly estimating the direct and cross-price impact for any *two randomly selected* stocks in our sample. Specifically, for any two stocks i and h, randomly drawn from our merged TAQ/CRSP/COMPUSTAT database with a common history of all quarterly earnings (474 firms), we compute the correlation of their earnings ($EPS_{i,q}$ and $EPS_{h,q}$) and estimate the following reduced version of Eq. (16):

$$
\begin{aligned}
r_{i,t} = {} & \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} \\
& + \sum_{l=0}^{L} \lambda_{ii,l} \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l} \omega_{h,t-l} + \varepsilon_{i,t},
\end{aligned}
\tag{17}
$$

again by OLS over each stock pair's longest common trading history in TAQ within our sample period.[20] We repeat this procedure two thousand times, and then compute averages of the ensuing estimates of cumulative direct ($\sum_{l=0}^{L} \lambda_{ii,l}$) and pairwise *absolute* cross-price impact ($\left| \sum_{l=0}^{L} \lambda_{ih,l} \right|$, to prevent signed effects from canceling out) on stock i's return ($r_{i,t}$) in correspondence with a one standard deviation shock to its own ($\sigma(\omega_{i,t})$) and the other stock's order imbalance ($\sigma(\omega_{h,t})$), respectively. We report these averages in basis points (i.e., multiplied by $10,000$) in Table 4, together with averages of those effects within each quintile of stock pairs sorted according to their absolute earnings correlations ($|\rho_{i,h}|$) from the lowest to the highest, as well as averages of those effects when statistically significant. Consistent with Table 3, Table 4 indicates that average cross-price impact among randomly selected stocks is large and statistically significant more often than if due to chance (i.e., to statistical Type I error) — in 14% of the random stock pairs at the 10% level — even within the quintiles of firm pairs displaying the lowest average absolute earnings correlations. For instance, we find that the daily returns of a randomly selected stock within the first such quintile of firm pairs (whose mean $|\rho_{i,h}|$ is 0.04) move by an average of 41.1 basis points in correspondence with a one standard deviation shock to its own order flow — in the 93% of the cases in which such direct impact is statistically significant — and by 13.8 basis points in correspondence with a one standard deviation shock to the order flow of another randomly selected stock — in the 16% of the cases in which such cross-impact is statistically

---

[20]In unreported analysis, we find that both size and industry distributions of the subsample of firms satisfying these criteria are similar to those of the full sample.

significant.

4.1 The Informational Role of Cross-Trading

The evidence reported in Section 4 provides *indirect* support for the basic equilibrium implication of our model, i.e., that the equilibrium matrix for cumulated direct and cross-asset price impact be non-diagonal, even among fundamentally unrelated assets (Remark 1). In particular, Tables 3 and 4 indicate that cross-industry net order flow in the U.S. stock markets has (statistically and economically) significant and persistent cross-price impact, even among only weakly fundamentally related industries. Given this important premise, we now test two additional, unique predictions of our theory that stem from the informational role of speculators' strategic cross-trading in our stylized market setting. These predictions are unique to that theory, i.e., cannot be easily attributed to the alternative theories briefly discussed in Section 1. As such, if validated, they provide further, *direct* support for our model.

The first one (from Corollary 1) states that, ceteris paribus, equilibrium direct and absolute cross-asset liquidity is decreasing in the marketwide heterogeneity of speculators' private information (i.e., is increasing in ρ) because the latter makes their strategic direct and cross-trading activity more cautious and the MMs more vulnerable to adverse selection. We test this prediction parsimoniously by amending the regression models of Eqs. (16) and (17) to include the cross-products of direct and cross-order imbalances with either the average dispersion of analyst EPS forecasts ($SDLTEPS_m$, Eq. (10)) or the average standardized dispersion of macroeconomic forecasts ($SDMMS_m$, Eq. (11)). Specifically, we estimate the following amended versions of the regression models of Eqs. (16) and (17),

$$
\begin{aligned}
r_{n,t} = {} & \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10}\sum_{l=1}^{L}\gamma_{nj,l}r_{j,t-l} + \sum_{j=1}^{10}\sum_{l=0}^{L}\lambda_{nj,l}\omega_{j,t-l} \\
& + \sum_{j=1}^{10}\sum_{l=0}^{L}\lambda_{nj,l}^{x}X_t\omega_{j,t-l} + \varepsilon_{n,t},
\end{aligned}
\tag{18}
$$

and

$$
\begin{aligned}
r_{i,t} = {} & \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L}\gamma_{ii,l}r_{i,t-l} + \sum_{l=1}^{L}\gamma_{ih,l}r_{h,t-l} + \sum_{l=0}^{L}\lambda_{ii,l}\omega_{i,t-l} + \sum_{l=0}^{L}\lambda_{ih,l}\omega_{h,t-l} \\
& + \sum_{l=0}^{L}\lambda_{ii,l}^{x}X_t\omega_{i,t-l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_t\omega_{h,t-l} + \varepsilon_{i,t},
\end{aligned}
\tag{19}
$$

where the variable X_t is either $SDLTEPS_m$ or $SDMMS_m$. Because of data limitations, in the latter case our sample ends in December 2000. As clear from Eqs. (16) and (17), the scale of X_t affects the scale of the ensuing estimates for the cross-product coefficients. Thus, to ease the interpretation of the results, we compute (and report in Tables 5 and 6) the differences between

20

OLS estimates of direct and absolute cross-price impact in days characterized by historically *high* information heterogeneity — i.e., for X_t at the top 70th percentile of its empirical distribution, $X_{t,70^{th}}$ — and those in days characterized by historically *low* information heterogeneity — i.e., for X_t at the bottom 30th percentile of its empirical distribution, $X_{t,30^{th}}$.

These differences are consistent with Corollary 1: Both direct and absolute cross-price impact are generally *higher* when the extent of information heterogeneity among speculators is high (ρ is low). In particular, we find that, in relation to Tables 3 and 4, $\left| \sum_{l=0}^{L} \lambda_{nj,l} + \sum_{l=0}^{L} \lambda_{nj,l}^x X_{t,70^{th}} \right| - \left| \sum_{l=0}^{L} \lambda_{nj,l} + \sum_{l=0}^{L} \lambda_{nj,l}^x X_{t,30^{th}} \right|$ of Table 5, and both $\sigma\left(\omega_{i,t}\right) \left(\sum_{l=0}^{L} \lambda_{ii,l}^x X_{t,70^{th}} - \sum_{l=0}^{L} \lambda_{ii,l}^x X_{t,30^{th}} \right)$ and $\sigma\left(\omega_{h,t}\right) \left(\left| \sum_{l=0}^{L} \lambda_{ih,l}^x + \sum_{l=0}^{L} \lambda_{ih,l}^x X_{t,70^{th}} \right| - \left| \sum_{l=0}^{L} \lambda_{ih,l}^x + \sum_{l=0}^{L} \lambda_{ih,l}^x X_{t,30^{th}} \right| \right)$ of Table 6 are generally positive and statistically significant — especially when ρ is proxied by $SDLTEPS_m$ (Panel A), less so when by $SDMMS_m$ (Panel B).[21] For example, Panel A of Table 5 shows that on average, daily Manufacturing stock returns increase by 224.7 basis points in correspondence with a one standard deviation shock to Manufacturing stocks' order flow when $SDLTEPS_m$ is high (i.e., for $SDLTEPS_{m,70^{th}} = 5.12$), but by just 73.7 basis points if that shock takes place while $SDLTEPS_m$ is low (i.e., for $SDLTEPS_{m,30^{th}} = 3.75$); Panel B of Table 5 further shows that, e.g., Telecom stock returns are generally insensitive to trading activity in Durables stocks unless when $SDMMS_m$ is high (i.e., for $SDMMS_{m,70^{th}} = 10.65$), in which case those returns decrease by 136.2 basis points in response to a one standard deviation shock to net order flow in Durables stocks despite the correlation of their quarterly EPS being small and statistically insignificant (0.128 in Table 2). Consistently, Table 6 indicates that especially direct, but also cross-price impact among random pairs of stocks are statistically significant more often than if due to chance and higher in days when either $SDLTEPS_m$ or $SDMMS_m$ is larger than average, both over the entire sample and within nearly all of the earnings correlation quintiles. For instance, Panel A of Table 6 shows that when marketwide information heterogeneity (proxied by $SDLTEPS_m$) is high, the daily returns of a randomly selected stock move on average by almost 36 basis points more than when marketwide information heterogeneity is low in correspondence with a one standard deviation shock to its own order flow — in the 31% of the cases in which differences in direct impact are statistically significant — and by 2.2 basis points more in correspondence with a one standard deviation shock to the order flow of another randomly selected stock — in the

---

[21] In these and similar subsequent tables, estimated cross-price impact coefficients occasionally change sign depending on the magnitude of X_t. In those circumstances, we report the *distance* between these estimates at $X_{t,70^{th}}$ and $X_{t,30^{th}}$ and sign it depending on its accordance with the model. We also mark differences (or distances) of sums of estimated price impact coefficients with the subscript "$\circ$" when *i)* neither sum is statistically significant but their difference (or distance) is, or *ii)* only one sum is statistically significant and the difference (or distance) is also significant but with the opposite sign.

18% of the cases in which differences in cross-impact are statistically significant.

The second prediction (again from Corollary 1) states that ceteris paribus, the more numerous speculators are in the economy (higher M), the less cautiously they cross-trade with their private signals, the less severe adverse selection risk becomes for uninformed market makers in all assets, hence ultimately the greater is both direct and absolute cross-asset liquidity (i.e., the lower are both direct and absolute cross-price impact of aggregate net order flow). To evaluate this argument, we estimate the amended regression models of Eqs. (18) and (19) after allowing for the cross-product of direct and cross-asset order flow with ANA_m, the equal-weighted average of analyst coverage among the stocks in our sample, as a proxy for the number of informed traders in the U.S. economy. Specifically, in the spirit of Brennan and Subrahmanyam (1995) and Chordia et al. (2007), among others, we define ANA_m (plotted in Figure 2c) as

$$ANA_m = \frac{1}{N_m} \sum_{i=1}^{N_m} ANA_{i,m}, \tag{20}$$

where $ANA_{i,m}$ is the number of analysts covering firm i and reporting their forecasts of that firm's earnings to the I/B/E/S database in month m.[22] We then estimate Eqs. (18) and (19) after setting $X_t = ANA_m$. Again we report the differences between OLS estimates of direct and absolute cross-price impact in days characterized by historically *large* and *small* number of speculators — i.e., for ANA_m at the top 70$^{\text{th}}$ and at the bottom 30$^{\text{th}}$ percentiles of its empirical distribution, $ANA_{t,70^{th}}$ and $ANA_{t,70^{th}}$ — in Tables 7 and 8.

The evidence in these tables indicates that, again consistent with Corollary 1, the magnitude of direct and especially absolute cross-industry price impact is generally *lower* in days when the average number of analysts per firm in the market is large: $\left| \sum_{l=0}^{L} \lambda_{nj,l} + \sum_{l=0}^{L} \lambda_{nj,l}^{x} X_{t,70^{th}} \right| - \left| \sum_{l=0}^{L} \lambda_{nj,l} + \sum_{l=0}^{L} \lambda_{nj,l}^{x} X_{t,30^{th}} \right|$ of Table 7, and both $\sigma(\omega_{i,t}) \left(\sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,70^{th}} - \sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,30^{th}} \right)$ and $\sigma(\omega_{h,t}) \left(\left| \sum_{l=0}^{L} \lambda_{ih,l}^{x} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,70^{th}} \right| - \left| \sum_{l=0}^{L} \lambda_{ih,l}^{x} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,30^{th}} \right| \right)$ of Table 8 are generally *negative* and (in relation to Tables 3 and 4) often statistically significant, albeit less so than in Tables 5 and 6. For instance, Table 7 shows that on average, daily Shops stock returns decrease by 125.4 basis points in correspondence with a one standard deviation shock to HighTech stocks' order flow when ANA_m is historically low ($ANA_{m,30^{th}} = 6.07$), but by only 58.8 basis points if that shock occurs when ANA_m is high ($ANA_{m,30^{th}} = 6.39$); similarly, it is in days when ANA_m is historically low that the daily stock returns of many of the industries in our sample (especially Shops and Other) display the greatest sensitivity to cross-industry trading activity. Along those lines, Table 8 indicates that as postulated by our model, averages of both direct

---

[22]We employ firm-level averages to adjust for the time-varying number of firms in our sample. We obtain similar inference when replacing ANA_m with the *sum* of firm-level $ANA_{i,m}$.

and cross-asset liquidity among randomly drawn pairs of stocks in the U.S. equity markets are lower — more often than if due to chance, but generally by no more than a few basis points — the smaller is the number of speculators in the economy, regardless of the absolute correlation of their earnings. For example, when the marketwide number of speculators in the U.S. equity markets is low, a one standard deviation shock to the order flow of a randomly selected stock within the third earnings correlation quintile of firms significantly moves both that stock's daily returns (in 18% of the random pairs) and the daily returns of another randomly selected stock (in 10% of the random pairs) by an average of almost 3 basis points less than when the number of speculators is high.

Overall, the above results provide additional support for our theory, for they indicate that direct and cross-asset liquidity in the U.S. equity markets is related to the informational role of the strategic direct and cross-trading activity of better-informed speculators in those markets.

4.2 Alternative Theories of Cross-Trading

In this section we assess the importance of alternative theories of the relationship between cross-trading and cross-asset liquidity for the evidence presented above. As in our model, these theories also emphasize the role of *financial linkages*, rather than of real ones, in the process of price co-formation in equity markets. Yet, they propose alternative mechanisms — e.g., related to the extent and dynamics of marketwide fundamental uncertainty, risk aversion, and financial constraints — potentially leading to the equilibrium cross-price impact of speculators' trading activity. Tables 5 and 6 do not explicitly control for any of these mechanisms. These omissions have the potential to bias our inference. For example, our model assumes that all market participants are risk neutral and does not allow for the endogenous entry of informed speculators (e.g., Veldkamp, 2006). It is nonetheless possible that both their equilibrium number and the dispersion of their beliefs may be related to their time-varying risk aversion, fundamental volatility, or financial constraints, since those factors may affect speculators' potential profits from strategic direct and cross-trading. In those circumstances, omitted variable biases might arise, making our previous inference spurious. Current literature groups these alternative channels of transmission into several, often related categories (e.g., see the discussion in Kodres and Pritsker, 2002; Pasquariello, 2007; Kallberg and Pasquariello, 2008). In what follows, we gauge the robustness of our inference to their inclusion. Our analysis indicates that this inference is indeed robust.

The first one, the *correlated information* channel (e.g., King and Wadhwani, 1990), is based upon the idea that in the presence of information asymmetry among investors, cross-trading activity motivated by correlated information shocks may lead to cross-asset liquidity effects.

23

By construction, this mechanism precludes cross-price impact among fundamentally unrelated assets since that impact stems directly from uninformed investors' cross-inference about the terminal payoffs of the traded assets, in absence of financial intermediation. Both our model and empirical results suggest otherwise.[23] However, this mechanism also implies that greater marketwide information asymmetry may lead to greater direct and absolute cross-asset liquidity, consistent with both our model (e.g., see Σ_v in the expression for Λ of Eq. (4)) and most of the extant literature described below. In addition, time-varying information asymmetry may affect other parameters of our model — for instance if we endogenized speculators' participation (M) or the intensity of noise trading (σ_z) — as well as interact with marketwide dispersion of beliefs (ρ). As previously mentioned, ignoring these dynamics may lead to misspecification biases when estimating the regression models of Eqs. (18) and (19). We measure marketwide information asymmetry about U.S. stocks' future prospects with two alternative proxies. The first one is $EPSVOL_q$ (plotted in Figure 2d), the equal-weighted average of firm-level earnings volatility in calendar quarter q,

$$EPSVOL_q = \frac{1}{N_q} \sum_{i=1}^{N_q} EPSVOL_{i,q}, \tag{21}$$

where $EPSVOL_{i,q}$ is the standard deviation of firm i's quarterly EPS ($EPS_{i,q}$) over the most recent eight quarters from COMPUSTAT, as in Chordia et al. (2007), and N_q is the total number of firms in quarter q. The second one is $EURVOL_m$ (plotted in Figure 2e), the monthly average (to smooth daily variability) of daily Eurodollar implied volatility from Bloomberg; $EURVOL_m$ is a commonly used measure of market participants' perceived uncertainty surrounding U.S. monetary policy — an important source of common fundamental uncertainty in the U.S. stock markets (e.g., Bernanke and Kuttner, 2005; Bordo et al., 2008; Brenner et al., 2008; Vega and Wu, 2008).

Within the second category of theories, Fleming et al. (1998) and Kodres and Pritsker (2002) argue that in economies populated by uninformed investors learning from prices, the *portfolio rebalancing* activity of privately informed, price-taking investors — driven by risk aversion — may induce contemporaneous price covariance and cross-price impact, even among assets with uncorrelated payoffs. As mentioned in Section 4, both this intuition and the potential trading patterns ensuing from *style investing* (e.g., Barberis et al., 2005; Boyer, 2006) or correlated information-motivated trading motivate the inclusion of contemporaneous market returns $r_{M,t}$ in the basic empirical specifications of Eqs. (16) and (17). We further measure the extent and

---

[23]Accordingly, Cohen and Frazzini (2006) show that stock returns do not adjust promptly to shocks about economically related firms.

dynamics of marketwide risk aversion, or risk appetite in the U.S. stock market with a model-free proxy suggested by Bollerslev and Zhou (2007), $RISKAV_m$ (plotted in Figure 2f), the monthly difference between the end-of-month Chicago Board Options Exchange (CBOE)'s VIX index of implied volatility of S&P500 options with a fixed 30-day maturity, VIX_m, and the realized volatility of five-minute S&P500 returns, $r_{SP,\tau}$, within the month:

$$RISKAV_m = VIX_m - RV_m, \qquad (22)$$

where $RV_m = \sqrt{\sum_{\tau \in m} r_{SP,\tau}^2}$.[24]

A third set of studies models the cross-trading activity of speculators, even across fundamentally unrelated assets, as the equilibrium outcome of *correlated liquidity* shocks due to financial constraints (Calvo, 1999; Kyle and Xiong, 2001; Yuan, 2005). In the presence of those shocks, speculators' trading activity may also lead to equilibrium cross-asset liquidity by influencing the inferences and trades of other speculators and uninformed investors via prices (Bernhardt and Taub, 2008), rather than via order flow (as in our model). The most direct implication of these arguments for our analysis is that on average, absolute cross-industry price impact may be *asymmetric* — i.e., higher during times when borrowing, short-selling, and wealth constraints are particularly binding (e.g., during periods of liquidity crises) — and/or sensitive to the dynamics of interest rates (Shiller, 1989). We proxy for the former with a dummy, d_t^{CR}, equal to one if day t falls within any of the liquidity crisis periods listed by Chordia et al. (2005), and zero otherwise.[25] We capture the latter with a measure of time-varying risk-free interest rates, $r_{RF,t}$, the daily time series of one-month Treasury Bill rates (from CRSP).

We assess the relevance of these arguments for our inference on direct and cross-asset liquidity parsimoniously by including the cross-products of direct and cross-order imbalances with the various proxies described above in the regression models of Eqs. (18) to (19). We begin by estimating Eqs. (18) and (19) after replacing the information variable X_t with either $EPSVOL_q$, $EURVOL_t$, $RISKAV_t$, $r_{RF,t}$, or d_t^{CR}, *separately*, i.e., while ignoring both the number of speculators and their information heterogeneity. The results of this analysis, not reported here for economy of space, provide (at best) weak support for the notion that both direct and absolute cross-price impact may be increasing in marketwide risk aversion, and little or no evidence of

---

[24]We thank Tim Bollerslev and Hao Zhou for sharing the $RISKAV_m$ data with us. Similar results ensue when replacing RV_m with the standard deviation of $r_{SP,\tau}$ over the same monthly window. Nyberg and Wilhlmsson (2008) provide a theoretical motivation for time-varying risk aversion and discuss the merits of its nonparametric estimation over the formal estimation of parametrized consumption-based asset pricing models.

[25]These periods are: March 1, 1994 to May 31, 1994 (U.S. bond market crisis); July 2, 1997 to December 31, 1997 (Asian crisis); and July 6, 1998 to December 31, 1998 (Russian default crisis).

direct and cross-asset liquidity being sensitive to fluctuations in marketwide fundamental uncertainty, cost of borrowing, or to the occurrence of liquidity crises.

We also amend Eqs. (18) and (19) to include the cross-products of direct and cross-asset order imbalance with both our proxies for either information heterogeneity or the number of speculators, separately, and *all* of the proxies described above, X_t^v, as follows:

$$
\begin{aligned}
r_{n,t} &= \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10}\sum_{l=1}^{L}\gamma_{nj,l}r_{j,t-l} + \sum_{j=1}^{10}\sum_{l=0}^{L}\lambda_{nj,l}\omega_{j,t-l} \\
&\quad + \sum_{j=1}^{10}\sum_{l=0}^{L}\lambda_{nj,l}^x X_t \omega_{j,t-l} + \sum_{v=1}^{5}\left(\sum_{j=1}^{10}\sum_{l=0}^{L}\lambda_{nj,l}^v X_t^v \omega_{j,t-l}\right) + \varepsilon_{n,t},
\end{aligned}
\tag{23}
$$

for the ten industries listed in Section 4, and

$$
\begin{aligned}
r_{i,t} &= \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L}\gamma_{ii,l}r_{i,t-l} + \sum_{l=1}^{L}\gamma_{ih,l}r_{h,t-l} + \sum_{l=0}^{L}\lambda_{ii,l}\omega_{i,t-l} + \sum_{l=0}^{L}\lambda_{ih,l}\omega_{h,t-l} \\
&\quad + \sum_{l=0}^{L}\lambda_{ii,l}^x X_t \omega_{i,t-l} + \sum_{l=0}^{L}\lambda_{ih,l}^x X_t \omega_{h,t-l} \\
&\quad + \sum_{v=1}^{6}\left(\sum_{l=0}^{L}\lambda_{ii,l}^v D_t^v \omega_{i,t-l}\right) + \sum_{v=1}^{5}\left(\sum_{l=0}^{L}\lambda_{ih,l}^v D_t^v \omega_{h,t-l}\right) + \varepsilon_{i,t},
\end{aligned}
\tag{24}
$$

among randomly selected pairs of stocks, where X_t is either $SDLTEPS_m$, $SDMMS_m$, or ANA_m, while X_t^v is $EPSVOL_q$, $EURVOL_t$, $RISKAV_t$, $r_{RF,t}$, and d_t^{CR}, respectively. As in Section 4.1, we compute (but again do not report here) the resulting differences between absolute OLS estimates of direct and cross-price impact from Eqs. (23) and (24) when the corresponding information variable X_t is historically *high* $(X_{t,70^{th}})$ and those when X_t is historically *low* $(X_{t,30^{th}})$. We find these differences to be generally consistent in sign, magnitude, and (economic and statistical) significance with those in Tables 5 and 6 among both industry portfolios and randomly selected pairs of stocks, i.e., to lead to qualitatively similar inference. Hence, we conclude that the evidence so far presented in support of our theory is robust to allowing for direct and cross-price impact to respond to fluctuations in fundamental volatility, risk aversion, or financial constraints.

Lastly, we also consider the empirical relevance of *price observability* for equilibrium cross-asset liquidity stemming from the strategic trading activity of risk neutral informed investors, as suggested by Bernhardt and Taub (2008). This channel is complementary to the transmission mechanism described in our model, based instead on order flow observability. To that purpose, we exploit an important feature of the NYSE, namely the fact that NYSE dealers (*specialists*) specialize in nonintersecting subsets of the traded stocks (e.g., Corwin, 1999; Hasbrouck, 2007). To the extent that cross-order flow observability is likely to be the highest for stocks dealt by the same specialist, ceteris paribus we would expect average cross-price impact to be higher across those stocks than across stocks dealt by different specialists. To test for this argument, we use specialist information on NYSE-listed stocks from the NYSE Post and Panel File (e.g.,

Coughenour and Deli, 2002). This information — available to us exclusively between November 2001 and June 2004 — is accessible to all market participants and allows to identify the specialists dealing multiple NYSE stocks by matching those stocks' Post and Panel locations on the NYSE trading floor.[26] We then estimate and compare cumulative direct and pairwise absolute cross-price impact (Eq. (17)) accompanying a one standard deviation shock to direct and cross-stock net order flow for two sets of stock pairs (with a common history of all quarterly earnings within our full sample) over the pairs' longest common trading history in TAQ within the subperiod 11/2001-6/2004: *i)* all pairs of stocks always dealt by the *same* specialist and specialist firm during that interval (80 pairs) and *ii)* the same number of randomly selected pairs of stocks always dealt by a *different* specialist and specialist firm during that interval.

We report averages of these estimates for each quintile of those stocks' absolute earnings correlations in Panels A (random NYSE-only stock pairs dealt by the same specialist) and B (random NYSE-only stock pairs dealt by different specialists) of Table 9. These estimates highlight two interesting results. First, when compared to those reported in Table 4 for random pairs of NYSE and NASDAQ stocks, Table 9 suggests that despite the smaller sample size, our inference about the economic and statistical significance of direct and cross-price impact among U.S. stocks is insensitive to whether those stocks are traded at the hybrid (open outcry, dealer, and electronic limit order book) NYSE market or at the (primarily) dealer NASDAQ market. Second, average estimated direct and especially absolute cross-price impact among pairs of stocks dealt by the same specialists (i.e., with the highest cross-order flow observability, in Panel A) are as often statistically significant but *greater* in magnitude than the corresponding estimates among pairs stocks dealt by different specialists (i.e., with more limited cross-order flow observability, in Panel B) — as implied by our model — both over the entire sample and within most quintiles of firm pairs sorted according to their absolute earnings correlations ($|\rho_{i,h}|$).[27] These differences are also economically significant. For example, Table 9 shows that a one standard deviation shock to a

---

[26]E.g., the current map of the NYSE trading floor is available at http://marketrac.nyse.com/mt/index.html. Order flow observability is also likely to be higher for NYSE stocks dealt by specialists employed by the same *specialist firms* (albeit lower than for stocks dealt by the same specialist), yet so are those firms' efforts at managing their aggregate stock inventory (e.g., Coughenour and Saad, 2004). Consistently, the inference that follows is qualitatively similar but weaker when grouping stocks according to their specialist firms.

[27]Importantly for this comparison, those average absolute pairwise EPS correlations are instead nearly identical in both sets of firm pairs. Further unreported analysis indicates that, as in Coughenour and Saad (2004), the NYSE specialists' stock portfolios in our sample are not concentrated across such stock characteristics as industry or market capitalization. Consistently, Corwin (2004) observes that NYSE stocks are allocated among specialist firms primarily according to those firms' relative position in the *queue* — i.e., the time since each received a prior allocation — rather than according to any particular stock characteristic.

stock's order imbalance moves that (another) stock's daily returns by an average of 7.1 (2.3) basis points more — i.e., by 23% (11%) more — if both stocks are dealt by the same specialist than if they are dealt by different specialists, and by as much as 15.7 (4.3) basis points more — i.e., by as much as 46% (20%) more — within the fourth (fifth) quintile of firm pairs (whose quarterly EPS correlations are most often statistically significant). This evidence indicates that, consistent with our model, order-flow observability has a first-order effect on direct and cross-asset liquidity in the U.S. stock market.

4.3 Public News and Cross-Trading

We have so far examined the equilibrium implications of the cross-trading activity of speculators endowed with private information about the terminal payoffs of the traded assets, v, for the process of price co-formation in the U.S. stock market. Yet, public news about those assets' fundamentals is often released to all market participants in the U.S. financial markets. For instance, on average at least one of the 18 macroeconomic news items described in Section 3.2 is released on 60% of the business days in every month between January 1993 and June 2004. In Section 2.2 we show that access to public, trade-free marketwide information, S_p, for all market participants in our economy reduces both direct and cross-price impact for all traded securities since it attenuates adverse selection risk for otherwise uninformed MMs — the more so the better is the public signal's quality (Corollary 2) and the more severe is that risk prior to its release (e.g., in the presence of fewer or more heterogeneous speculators, Remark 2).

In this section, we assess the empirical relevance of these considerations by employing our database of macroeconomic news releases mentioned above. We do so by amending the regression models of Eqs. (16) and (17) to allow for the release (and quality) of public information, as well as for its interaction with the marketwide information environment (number of speculators and information heterogeneity), to affect direct and cross-asset liquidity in a parsimonious way, as follows:

$$
\begin{aligned}
r_{n,t} &= \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10} \sum_{l=1}^{L} \gamma_{nj,l} r_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l} \left(1 - D_t^p\right) \omega_{j,t-l} \\
&\quad + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l}^p D_t^p \omega_{j,t-l} + \varepsilon_{n,t},
\end{aligned}
\tag{25}
$$

and

$$
\begin{aligned}
r_{i,t} &= \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l} \left(1 - D_t^p\right) \omega_{i,t-l} \\
&\quad + \sum_{l=0}^{L} \lambda_{ih,l} \left(1 - D_t^p\right) \omega_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l}^p D_t^p \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}^p D_t^p \omega_{h,t-l} + \varepsilon_{i,t}. \tag{26}
\end{aligned}
$$

In Eqs. (25) and (26), D_t^p is a dummy variable equal to one if any of the 18 macroeconomic news in our sample is released on day t and equal to zero otherwise (over our full sample period January 1993-June 2004, again from MMS).[28] If Corollary 2 is correct, we expect direct and absolute cross-price impact to decline in proximity of those news releases, i.e., the differences $\left| \sum_{l=0}^{L} \lambda_{nj,l}^p \right| - \left| \sum_{l=0}^{L} \lambda_{nj,l} \right|$, $\sigma\left(\omega_{i,t}\right)\left(\sum_{l=0}^{L} \lambda_{ii,l}^p - \sum_{l=0}^{L} \lambda_{ii,l}\right)$, and $\sigma\left(\omega_{h,t}\right)\left(\left| \sum_{l=0}^{L} \lambda_{ih,l}^p \right| - \left| \sum_{l=0}^{L} \lambda_{ih,l} \right|\right)$ to be *negative*. We report estimates for these differences for our industry portfolios and randomly selected stock pairs in Table 10 and 11, respectively.

The evidence in those tables is only weakly supportive of Corollary 2. Direct and absolute cross-industry permanent price impact are most often statistically unaffected by the release of public news, although the latter is much more likely to decline when it is. In those circumstances, these effects are economically significant as well: e.g., Table 10 indicates that a one standard deviation shock to cross-industry order flow during non-announcement days affects cross-industry returns by an average of 20.3 basis points more than during announcement days; a one standard deviation shock to Shops stocks during non-announcement days lowers daily Energy stock returns on average by 36.2 basis points more than during announcement days. Similarly weak support for Corollary 2 comes from Table 11, in which direct and absolute cross-stock price impact among randomly selected stock pairs are generally lower in correspondence with the release of U.S. macroeconomic news, as postulated by our model, yet both are rarely statistically significantly so. In those cases, however, the estimated effect is economically significant — e.g., $\sigma\left(\omega_{h,t}\right)\left(\left| \sum_{l=0}^{L} \lambda_{ih,l}^p \right| - \left| \sum_{l=0}^{L} \lambda_{ih,l} \right|\right)^* = -21.1$ bps within the highest earnings correlation quintile of firms.

As noted by Pasquariello and Vega (2008), the literature suggests several alternative mechanisms that may mitigate the improvement in market liquidity accompanying the release of public signals as postulated by our model. Chowdhry and Nanda (1991) argue that speculators may divert their trading activity to the most liquid venues to maximize their expected profits. In such a setting, high-quality public information, by devaluing their private signals, may induce those speculators to migrate to the most liquid assets, thus deteriorating other assets' equilibrium liquidity. According to Kim and Verrecchia (1994), the release of public signals may worsen market liquidity when private information is costly unless if the precision of those signals is high and private information is less heterogenous. Consistently, we show (Corollaries 2 and 3) that both

---

[28]The inference that follows is qualitatively similar when employing all of the 25 macroeconomic announcements in the original MMS database, i.e., including those for which MMS does not report the dispersion of analyst forecasts: Capacity Utilization, Personal Income, Consumer Credit, Personal Consumption Expenditures, Business Inventory, Government Budget, and Target Fed Funds Rate.

factors affect the relation between the availability of public signals and direct and cross-asset liquidity.

We intend to assess the relevance of these considerations for the evidence in Tables 10 and 11. To that purpose, we measure the quality of released public information as the absolute difference between initial macroeconomic announcements and their last informative revision (i.e., not due to definitional changes), as in Pasquariello and Vega (2007, 2008). These revisions, from the Federal Reserve Bank of Philadelphia Real Time Data Set (RTDS), are available to us only for Capacity Utilization, Industrial Production, and Nonfarm Payroll Employment (arguably the most important of the announcements in our sample, e.g., Andersen and Bollerslev, 1998; Andersen et al., 2007; Brenner et al., 2008).[29] Intuitively, this approach is motivated by the observation that the final published informative revision of a macroeconomic variable constitutes the most accurate measure for that variable, and that those differences can be interpreted as noise since they are predictable based on past information (e.g., Mork, 1987; Aruoba, 2004; Faust et al., 2005). Thus, we amend Eqs. (25) and (26) to include the cross-products of direct and cross-order imbalances with those revisions in macroeconomic announcement days as follows:

$$
\begin{aligned}
r_{n,t} &= \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10} \sum_{l=1}^{L} \gamma_{nj,l} r_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l} \left(1 - D_t^p\right) \omega_{j,t-l} \\
&+ \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l}^p D_t^p \omega_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l}^{px} D_t^p X_t^p \omega_{j,t-l} + \varepsilon_{n,t},
\end{aligned}
\tag{27}
$$

and

$$
\begin{aligned}
r_{i,t} &= \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l} \left(1 - D_t^p\right) \omega_{i,t-l} \\
&+ \sum_{l=0}^{L} \lambda_{ih,l} \left(1 - D_t^p\right) \omega_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l}^p D_t^p \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}^p D_t^p \omega_{h,t-l} \\
&+ \sum_{l=0}^{L} \lambda_{ii,l}^{px} D_t^p X_t^p \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}^{px} D_t^p X_t^p \omega_{h,t-l} + \varepsilon_{i,t},
\end{aligned}
\tag{28}
$$

where D_t^p is a dummy variable equal to one if either of the public news p in the RTDS database is released on day t and equal to zero otherwise, and $X_t = ABSREV_t^p$ is the corresponding absolute revision. We estimate Eqs. (27) and (28) for Capacity Utilization, Industrial Production, and Nonfarm Payroll Employment separately and, as in Section 4.1, compute the differences between OLS estimates of direct and absolute cross-price impact in days when the released public news is of historically *low* quality — i.e., for $ABSREV_t^p$ at the top 70th percentile of its empirical distribution, $ABSREV_{t,70^{th}}^p$ — and those in days when the released public news is of historically

---

[29]For a more detailed description of the RTDS dataset and its properties, see Croushore and Stark (2001). Pasquariello and Vega (2007, 2008) find that those macroeconomic news releases improve the liquidity of the U.S. Treasury bond market the most when of the highest such quality.

high quality — i.e., for $ABSREV_t^p$ at the bottom 30th percentile of its empirical distribution, $ABSREV_{t,70^{th}}^p$. For economy of space, we report these estimates only for Nonfarm Payroll Employment, in Tables 12 and 13, respectively. Inference from the release of Capacity Utilization and Industrial Production news is qualitatively and quantitatively similar.[30]

Corollary 2 implies that the release of public signals of better quality leads to greater direct and cross-asset liquidity. The evidence from the estimation of Eqs. (27) and (28) provides some support for these implications of our model, yet only among industry portfolios, i.e., where the information content and quality of macroeconomic news are more likely to matter. In particular, the estimated differences $\left| \sum_{l=0}^L \lambda_{nj,l}^p + \sum_{l=0}^L \lambda_{nj,l}^{px} X_{t,70^{th}}^p \right| - \left| \sum_{l=0}^L \lambda_{nj,l}^p + \sum_{l=0}^L \lambda_{nj,l}^{px} X_{t,30^{th}}^p \right|$ in Table 12 are generally *positive* and statistically significant, i.e., direct and absolute cross-price impact among industry portfolios of U.S. stocks are generally *higher* in Nonfarm Payroll Employment announcement days when the corresponding $ABSREV_t^p$ is historically high than when $ABSREV_t^p$ is historically low. In those circumstances, the estimated improvement in direct and cross-industry liquidity is economically significant. For instance, the absolute cross-industry price impact of a one standard deviation shock to cross-industry order flow on days when high-quality Nonfarm Payroll Employment numbers are released is on average 34 basis points greater than when the quality of that announcement is low, and as high as 68 bps greater in response to order imbalances in Manufacturing stocks. Direct and cross-price impact among most random stock pairs (in Table 13) are instead either insensitive to, or even increase in correspondence with the release of public news of better quality. Only within the quintile of firm pairs with the highest absolute earnings correlations — i.e., whose estimates of direct and cross-stock liquidity display the greatest sensitivity to macroeconomic news arrivals in Table 11 — both $\sigma\left(\omega_{i,t}\right) \left(\sum_{l=0}^L \lambda_{ii,l}^{px} X_{t,70^{th}}^p - \sum_{l=0}^L \lambda_{ii,l}^{px} X_{t,30^{th}}^p \right) > 0$ and $\sigma\left(\omega_{h,t}\right) \left(\left| \sum_{l=0}^L \lambda_{ih,l}^p + \sum_{l=0}^L \lambda_{ih,l}^{px} X_{t,70^{th}}^p \right| - \left| \sum_{l=0}^L \lambda_{ih,l}^p + \sum_{l=0}^L \lambda_{ih,l}^{px} X_{t,30^{th}}^p \right| \right) > 0$, consistent with Corollary 2.

In light of this evidence, we conclude that as postulated by our theory, the availability of public signals of the terminal payoffs of all stocks improves (albeit not always importantly so) direct, cross-industry, and cross-stock liquidity in the U.S. equity markets by mitigating the adverse selection risk stemming from the strategic direct and cross-trading activity of informed market participants.

---

[30]Since revision data is available only for a subset of the 25 macroeconomic signals in the MMS database released over our sample period Eqs. (27) and (28) control explicitly for direct and cross-price impact and trading activity in all other non-announcement or other-announcement days $(1 - D_t^p)$.

5   Conclusions

This study provides a theory and empirical evidence of the informational role of trading for the process of price co-formation in the U.S. equity markets.

To motivate our empirical analysis, we develop a parsimonious model of strategic, informed cross-trading in a multi-asset economy in the presence of two realistic market frictions — imperfect competition among speculators and marketwide information heterogeneity — and a common public signal. The ensuing setting — based on Kyle (1985), Caballé and Krishnan (1994), and Pasquariello (2007) — allows us to precisely characterize the equilibrium properties of both direct and cross-price impact — the permanent impact of informed trades in one asset on both the price of that asset and the prices of other (either related or fundamentally unrelated) assets — when extant channels of trade and price *co*-formation in the literature (inventory management, correlated information, portfolio rebalancing, correlated liquidity, and private observability) are ruled out by construction. We show that in those circumstances cross-asset liquidity is the equilibrium outcome of strategic trading activity of risk neutral speculators across many assets to mask their information advantage about some other assets.

We find strong support for such cross-asset informational effects in a comprehensive sample of the trading activity in NYSE and NASDAQ stocks between 1993 and 2004. In particular, we report robust evidence that order flow in one stock or industry has a significant and persistent impact on daily returns of other stocks or industries. Our empirical analysis also suggests that direct and absolute cross-price impact are *i)* smaller when speculators are more numerous; *ii)* greater when marketwide dispersion of beliefs is higher; *iii)* greater among stocks dealt by the same specialist; and *iv)* smaller when macroeconomic news of good quality is released, consistent with our stylized model.

Overall, these novel findings indicate that cross-asset liquidity is economically and statistically significant in the U.S. stock market as well as crucially related to the informational role of strategic cross-trading. We believe this is an important contribution to the literature, one that bears important implications for future research on the process of price formation in financial markets.

6   Appendix A

Proof of Proposition 1.   The basic economy of Section 2.1 nests in the more general setting of Pasquariello (2007). In addition, the distributional assumptions of Section 2.1 imply that

$\Sigma_\delta = \rho \Sigma_v$ and $\Sigma_c = \rho^2 \Sigma_v = \rho \Sigma_\delta$. Hence, the symmetric linear equilibrium of Proposition 1 follows from Proposition 1 and Remark 1 in Pasquariello (2007). Uniqueness of that equilibrium then ensues from the assumption that $\Sigma_z = \sigma_z^2 I$ (Proposition 3.2 in Caballé and Krishnan, 1994). ∎

Proof of Remark 1. The statement of the remark follows from the definition of Λ in Proposition 1 (Eq. (4)) and the assumption that Σ_v is SPD. Specifically, the latter implies that $\Sigma_v^{1/2} = C \Delta C$ where Δ is a diagonal matrix whose diagonal terms are given by the square roots of the characteristic roots of Σ_v ($\lambda_n > 0$) and C is a matrix whose columns are made of the corresponding orthogonal characteristic vectors c_n, i.e., such that $\Sigma_v C = C \Delta$ (e.g., Greene, 1997, pp. 36-43). It then follows that $\Sigma_v^{1/2}(l,j) = \sum_{n=1}^{N} c_{nl} c_{jn} \sqrt{\lambda_n}$ will be different from zero if so is $\Sigma_v(l,j)$, and may be so although $\Sigma_v(l,j) = 0$. ∎

Proof of Corollary 1. Direct and cross-asset liquidity are increasing in the number of speculators, since $\frac{\partial |\Lambda(n,j)|}{\partial M} = \frac{\rho[2-(M+1)\rho]}{2\sqrt{M}\rho[2+(M-1)\rho]^2 \sigma_z} \left| \Sigma_v^{1/2}(n,j) \right| < 0$ under most parametrizations (i.e., except in the "small" region of $\{M, \rho\}$ where M is a "small" integer $\leq \frac{2-\rho}{\rho}$ and the speculators' private signals of v are "reasonably" precise). Moreover, $\lim_{M \to \infty} |\Lambda(n,j)| = 0$. The second part of the statement follows from the fact that $\frac{\partial |\Lambda(n,j)|}{\partial \rho} = \frac{M[2-(M-1)\rho]}{2\sqrt{M}\rho[2+(M-1)\rho]^2 \sigma_z} \left| \Sigma_v^{1/2}(n,j) \right| \geq 0$ if $\rho \leq \frac{2}{M-1}$, i.e., in the presence of "few" speculators ("small" M), and negative otherwise. ∎

Proof of Proposition 2. The amended economy of Section 2.2 nests in the setting of Section 2.1, since $\Sigma_\delta^* = \rho^* \Sigma_v^*$ and $\Sigma_c^* = \rho^{*2} \Sigma_v^* = \rho^* \Sigma_\delta^*$. Hence, existence and uniqueness of the symmetric linear equilibrium of Proposition 2 follow from Proposition 1 and Remark 1 in Pasquariello (2007) and Proposition 3.2 in Caballé and Krishnan (1994), respectively. ∎

Proof of Corollary 2. The availability of a public signal vector of v increases both direct and absolute cross-asset liquidity (i.e., lowers any nonzero $\Lambda(n,n)$ and $|\Lambda(n,j)|$) since the expression for Λ_p in Eq. (6) can be written as $\Lambda_p = \frac{(1-\psi_p)[2+(M-1)\rho]}{\sqrt{1-\rho\psi_p}[2+(M-1)\rho^*]} \Lambda = \phi_p \Lambda$ where $\phi_p < 1$ for any $\rho \in (0,1)$ and $\psi_p \in (0,1)$. In addition, it can be shown that $\frac{\partial \phi_p}{\partial \psi_p} < 0$ over the range of feasible values for ρ and ψ_p. ∎

Proof of Remark 2. The improvement in direct and cross-asset liquidity due to the availability of a public signal S_p is decreasing in M since $\Lambda_p = \phi_p \Lambda$ (where $\phi_p < 1$ for $\rho \in (0,1)$ and $\psi_p \in (0,1)$, see the proof of Corollary 2) and it can be shown that $\frac{\partial \phi_p}{\partial M} > 0$ for any $\rho \in (0,1)$ and $\psi_p \in (0,1)$. The second part of the statement follows from the observation that $\frac{\partial |\Lambda_p(n,j)|}{\partial \rho} = $

$\frac{M\sqrt{1-\psi_p}\left\{2\rho^2(1-\psi_p^2)-2\rho(1-\psi_p)\left[\frac{(M-1)(1-\psi_p)}{(1-\rho\psi_p)^2}\right]+(1-\psi_p)[2+(M-1)\rho^*]\right\}}{2\sqrt{M}\rho^*[2+(M-1)\rho^*]^2(1-\rho\psi_p)^2 \sigma_z} \left| \Sigma_v^{1/2}(n,j) \right| \geq 0$ if $\rho \leq \frac{2}{(M-1)-\psi_p(M-3)}$,

i.e., in the presence of "few" speculators ("small" M), and negative otherwise. We show a similar result for $|\Lambda(n,j)|$ in the proof of Corollary 1, yet $\frac{2}{(M-1)-\psi_p(M-3)} \geq \frac{2}{M-1}$, the threshold ρ at which $\frac{\partial |\Lambda(n,j)|}{\partial \rho} = 0$, for any $\psi_p \in (0,1)$. This implies that $|\Lambda(n,j)| - |\Lambda_p(n,j)| > 0$ is first increasing then decreasing in ρ when M is "small", while first decreasing then increasing when M is "large." ∎

7 Appendix B

$$\Sigma_v = \left[\begin{array}{ccc} 2 & 0.5 & 0 \\ 0.5 & 1.5 & 0.5 \\ 0 & 0.5 & 2 \end{array} \right]. \tag{B-1}$$

References

Acharya, V., and Pedersen, L., 2005, Asset Pricing with Liquidity Risk, *Journal of Financial Economics*, 77, 385-410.

Admati, A., 1985, A Noisy Rational Expectations Equilibrium for Multi-Asset Securities Markets, *Econometrica*, 53, 629-657.

Admati, A., and Pfleiderer, P., 1988, A Theory of Intraday Trading Patterns: Volume and Price Variability, *Review of Financial Studies*, 1, 3-40.

Amihud, Y., 2002, Illiquidity and Stock Returns: Cross-Section and Time-Series Effects, *Journal of Financial Markets*, 5, 31-56.

Amihud, Y., and Mendelson, H., 1980, Dealership Market: Market Making with Inventory, *Journal of Financial Economics*, 17, 223-249.

Amihud, Y., and Mendelson, H., 1986, Asset Pricing and the Bid-Ask Spread, *Journal of Financial Economics*, 8, 31-53.

Andersen, T., and Bollerslev, T., 1998, Deutsche Mark-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies, *Journal of Finance*, 53, 219-265.

Andersen, T., T. Bollerslev, F. Diebold, and Vega, C., 2003, Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange, *American Economic Review*, 93, 38-62.

Andrade, S., Chang, C., and Seasholes, M., 2008, Trading Imbalances, Predictable Reversals, and Cross-Stock Effects, *Journal of Financial Economics*, 88, 406-423.

Aruoba, B., 2004, Empirical Properties of Macroeconomic Revisions, Working Paper, University of Pennsylvania.

Back, K., Cao, H., and Willard, G., 2000, Imperfect Competition Among Informed Traders, *Journal of Finance*, 55, 2117-2155.

Barberis, N., Shleifer, A., and Wurgler, J., 2005, Comovement, *Journal of Financial Economics*, 75, 283-317.

Bernhardt, D., and Taub, B., 2008, Cross-Asset Speculation in Stock Markets, *Journal of Finance*, forthcoming.

Berry, T., and Howe, K., 1994, Public Information Arrival, *Journal of Finance*, 49, 1331-1346.

Blume, M., and Goldstein, M., 1997, Quotes, Order Flow, and Price Discovery, *Journal of Finance*, 52, 221-244.

Bessembinder, H., 1999, Trade Execution Costs on NASDAQ and NYSE: A Post Reform Comparison, *Journal of Financial and Quantitative Analysis*, 34, 387-407.

Bessembinder, H., 2003, Issues in Assessing Trade Execution Costs, *Journal of Financial Markets*, 6, 233-257.

Boehmer, E., and Wu, J., 2008, Order Flow and Prices, Working Paper, Texas A&M University.

Bollerslev, T., and Zhou, H., 2007, Expected Stock Returns and Variance Risk Premia, Finance and Economics Discussion Series 2007-11, Board of Governors of the Federal Reserve System.

Bordo, M., Dueker, M., and Wheelock, D., 2008, Inflation, Monetary Policy, and Stock Market Conditions, NBER Working Paper No. 14019.

Boyer, B., 2006, Style Investing and the Book-to-Market Factor, Working Paper, Brigham Young University.

Brandt, M., and Kavajecz, K., 2004, Price Discovery in the U.S. Treasury Market: The Impact of Order Flow and Liquidity on the Yield Curve, *Journal of Finance*, 59, 2623-2654.

Brennan, M., Jegadeesh, N., and Swaminathan, B., 1993, Investment Analysis and the Adjustment of Stock Prices to Common Information, *Review of Financial Studies*, 4, 799-824.

Brennan, M., and Subrahmanyam, A., 1995, Investment Analysis and Price Formation in Securities Markets, *Journal of Financial Economics*, 38, 361-381.

Brennan, M., and Subrahmanyam, A., 1996, Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns, *Journal of Financial Economics*, 41, 441-464.

Brenner, M., Pasquariello, P., and Subrahmanyam, M., 2008, On the Volatility and Comovement of U.S. Financial Markets Around Macroeconomic News Announcements, *Journal of Financial and Quantitative Analysis*, forthcoming.

Caballé, J., and Krishnan, M., 1994, Imperfect Competition in a Multi-Security Market with Risk Neutrality, *Econometrica*, 62, 695-704.

Calvo, G., 1999, Contagion in Emerging Markets: When Wall Street is a Carrier, Working Paper, University of Maryland.

Campbell, J., Lettau, M., Malkiel, B., and Xu, Y., 2001, Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk, *Journal of Finance*, 56, 1-43.

Chan, K., 1993, Imperfect Information and Cross-Autocorrelation among Stock Prices, *Journal of Finance*, 48, 1211-1230.

Chordia, T., Huh, S., and Subrahmanyam, A., 2007, The Cross-Section of Expected Trading Activity, *Review of Financial Studies*, 20, 709-740.

Chordia, T., Roll, R., and Subrahmanyam, A., 2000, Commonality in Liquidity, *Journal of Financial Economics*, 56, 3-28.

Chordia, T., Sarkar, A., and Subrahmanyam, A., 2005, An Empirical Analysis of Stock and Bond Market Liquidity, *Review of Financial Studies*, 18, 85-130.

Chordia, T., Sarkar, A., and Subrahmanyam, A., 2008, The Microstructure of Cross-Autocorrelations, Working Paper, UCLA.

Chordia, T., and Swaminathan, B., 2000, Trading Volume and Cross-Autocorrelation in Stock Returns, *Journal of Finance*, 55, 913-935.

Chowdhry, B., and Nanda, V., 1991, Multimarket Trading and Market Liquidity, *Review of Financial Studies*, 4, 483-511.

Cohen, L., and Frazzini, A., 2006, Economic Links and Predictable Returns, *Journal of Finance*, forthcoming.

Comerton-Forde, C., Hendershott, T., Jones, C., Moulton, P., and Seasholes, M., 2008, Time Variation in Liquidity: The Role of Market Maker Inventories and Revenues *Journal of Finance*, forthcoming.

Constantinides, G., 1986, Capital Market Equilibrium with Transaction Costs, *Journal of Political Economy*, 94, 842-862.

Corwin, S., 1999, Differences in Trading Behavior across NYSE Specialist Firms, *Journal of Finance*, 54, 721-745.

Corwin, S., 2004, Specialist Performance and New Listing Allocations on the NYSE: An Empirical Analysis, *Journal of Financial Markets*, 7, 27-51.

Coughenour, J. and Deli, D., 2002, Liquidity Provision and the Organizational Form of NYSE Specialist Firms, *Journal of Finance*, 57, 841-869.

Coughenour, J. and Saad, M., 2004, Common Market Makers and Commonality in Liquidity, *Journal of Financial Economics*, 73, 37-69.

Croushore, D., and Stark, T., 2001, A Real-Time Data Set of Macroeconomists, *Journal of Econometrics*, 105, 111-130.

DeLong, B., Shleifer, A., Summers, L., and Waldmann, R., 1990, Noise Trader Risk in Financial Markets, *Journal of Political Economy*, 98, 703-738.

Diamond, D., and Verrecchia, R., 1981, Information Aggregation in a Noisy Rational Expectations Economy, *Journal of Financial Economics*, 9, 221-235.

Diamond, D., and Verrecchia, R., 1991, Disclosure, Liquidity, and the Cost of Capital, *Journal of Finance*, 46, 1325-1359.

Diether, K., Malloy, C., and Scherbina, A., 2002, Differences of Opinion and the Cross Section of Stock Returns, *Journal of Finance*, 57, 2113-2141.

Duffie, D., Garleanu, N., and Pedersen, L., 2005. Over-The-Counter Markets, *Econometrica*, 73, 1815-1847.

Duffie, D., Garleanu, N., and Pedersen, L., 2006, Valuation in Over-The-Counter Markets, *Review of Financial Studies*, forthcoming.

Evans, M., and Lyons, R., 2002, Order Flow and Exchange Rate Dynamics, *Journal of Political Economy*, 110, 170-180.

Evans, M., and Lyons, R., 2007, Exchange Rate Fundamentals and Order Flow, NBER Working Paper No. 13151.

Evans, M., and Lyons, R., 2008, How is Macro News Transmitted to Exchange Rates?, *Journal of Financial Economics*, 88, 26-50.

Fama, E., and French, K., 1988, Permanent and Temporary Components of Stock Prices, *Journal of Political Economy*, 96, 246-273.

Faust, J., Rogers, J., and Wright, J., 2005, News and Noise in G-7 GDP Announcements, *Journal of Money, Credit and Banking*, 37, 403-417.

Fleming, J., Kirby, C., and Ostdiek, B., 1998. Information and Volatility Linkages in the Stock, Bond, and Money Markets, *Journal of Financial Economics*, 49, 111-137.

Fleming, M., and Remolona, E., 1997, What Moves the Bond Market?, *Economic Policy Review*, Federal Reserve Bank of New York, 3, 31-50.

Foster, F., and Viswanathan, S., 1990, A Theory of the Interday Variations in Volume, Variance, and Trading Costs in Securities Markets, *Review of Financial Studies*, 3, 593-624.

Foster, F., and Viswanathan, S., 1993, The Effect of Public Information and Competition on Trading Volume and Price Volatility, *Review of Financial Studies*, 6, 23-56.

Foster, F., and Viswanathan, S., 1996, Strategic Trading When Agents Forecast the Forecasts of Others, *Journal of Finance*, 51, 1437-1478.

Garman, M., 1976, Market Microstructure, *Journal of Financial Economics*, 3, 257-275.

Green, C., 2004, Economic News and the Impact of Trading on Bond Prices, *Journal of Finance*, 59, 1201-1233.

Green, C., and Hwang, B., 2008, Price-Based Return Comovement, *Journal of Financial Economics*, forthcoming.

Greene, W., 1997, *Econometric Analysis*, Prentice Hall.

Greenwood, R., 2008, Excess Comovement of Stock Returns: Evidence from Cross-Sectional Variation in Nikkei 225 Weights, *Review of Financial Studies*, 21, 1153-1186.

Grossman, S., and Miller, M., 1988, Liquidity and Market Structure, *Journal of Finance*, 43, 617-637.

Hartford, J., and Kaul, A., 2005, Correlated Order Flow: Pervasiveness, Sources, and Pricing Effects, *Journal of Financial and Quantitative Analysis*, 40, 29-55.

Hasbrouck, J., 1988, Trades, Quotes, Inventories, and Information, *Journal of Financial Economics*, 22, 229-252.

Hasbrouck, J., 1991, Measuring the Information Content of Stock Trades, *Journal of Finance*, 46, 179-207.

Hasbrouck, J., 2007, *Empirical Market Microstructure*, Oxford University Press.

Hasbrouck, J., and Seppi, D., 2001, Common Factors in Prices, Order Flows, and Liquidity, *Journal of Financial economics*, 59, 383-411.

Holden, G., and Subrahmanyam, A., 1992, Long-Lived Private Information and Imperfect Competition. *Journal of Finance*, 47, 247-270.

Huang, M., 2003, Liquidity Shocks and Equilibrium Liquidity Premia, *Journal of Economic Theory*, 109, 104-129.

Jones, C., Kaul, G., and Lipson, M., 1994, Transactions, Volume, and Volatility, *Review of Financial Studies*, 7, 631-651.

Kallberg, J., and Pasquariello, P., 2008, Time-Series and Cross-Sectional Excess Comovement in Stock Indexes, *Journal of Empirical Finance*, 15, 481-502.

Karolyi, G., and Stulz, R., 1996, Why Do Markets Move Together? An Investigation of U.S.-Japan Stock Return Comovements, *Journal of Finance*, 51, 951-986.

Kim, O., and Verrecchia, R., 1994, Market Liquidity and Volume around Earnings Announcements, *Journal of Accounting and Economics*, 17, 41-67.

Kim, O., and Verrecchia, R., 1997, Pre-Announcement and Event-Period Private Information, *Journal of Accounting and Economics*, 24, 395-419.

King, M., and Wadhwani, S., 1990, Transmission of Volatility between Stock Markets, *Review of Financial Studies*, 3, 5-33.

Kodres, L., and Pritsker, M., 2002, A Rational Expectations Model of Financial Contagion, *Journal of Finance*, 57, 769-799.

Kumar, A., and Lee, C., 2006, Retail Investor Sentiment and Return Comovements, *Journal of Finance*, 61, 2451-2486.

Kyle, A., 1985, Continuous Auctions and Insider Trading, *Econometrica*, 53, 1315-1335.

Kyle, A., and Xiong, W., 2001, Contagion as a Wealth Effect, *Journal of Finance*, 56, 1401-1440.

Lang, M., and Lundholm, R., 1996, Corporate Disclosure Policy and Analyst Behavior, *Accounting Review*, 71, pp. 467-492.

Lee, C., and Ready, M., 1991, Inferring Trade Direction from Intraday Data, *Journal of Finance*, 46, 733-746.

Lo, A., and MacKinlay, C., 1990, When Are Contrarian Profits Due to Stock Market Overreaction?, *Review of Financial Studies*, 3, 175-205.

Maddala, G., 1987, *Econometrics*, McGraw-Hill.

McQueen, G., Pinegar, M., and Thorley, S., 1996, Delayed Reaction to Good News and the Cross-Autocorrelation of Portfolio Returns, *Journal of Finance*, 51, 889-919.

Moeller, S., Schlingemann, F., and Stulz, R., 2006, Do Acquirers with More Uncertain Growth Prospects Gain Less from Acquisitions?, *Review of Financial Studies*, forthcoming.

Mork, K., 1987, Ain't Behavin': Forecast Errors and Measurement Errors in Early GNP Estimates, *Journal of Business and Economic Statistics*, 5, 165-175.

Nyberg, P., and Wilhelmsson, A., 2008, Volatility Risk Premium, Risk Aversion, and the Cross-Section of Stock Returns, Working Paper, Swedish School of Economics and Business Administration.

Pasquariello, P., 2007, Imperfect Competition, Information Heterogeneity, and Financial Contagion, *Review of Financial Studies*, 20, 391-426.

Pasquariello, P., and Vega, C., 2007, Informed and Strategic Order Flow in the Bond Markets, *Review of Financial Studies*, 20, 1975-2019.

Pasquariello, P., and Vega, C., 2008, The On-The-Run *Liquidity* Phenomenon, *Journal of Financial Economics*, forthcoming.

Pastor, L., and Stambaugh, R., 2003, Liquidity Risk and Expected Stock Returns, *Journal of Political Economy*, 113, 642-685.

Pindyck, R., and Rotemberg, J., 1990, The Excess Co-Movement of Commodity Prices, *Economic Journal*, 100, 1173-1189.

Pindyck, R., and Rotemberg, J., 1993, The Comovement of Stock Prices, *Quarterly Journal of Economics*, 108, 1073-1104.

Qu, S., Starks, L., and Yan, H., 2004, Risk, Dispersion of Analyst Forecasts, and Stock Returns, Working Paper, University of Texas.

Sadka, R., and Scherbina, A., 2007, Analyst Disagreement, Mispricing, and Liquidity, *Journal of Finance*, 62, 2367-2403.

Shiller, R., 1989, Comovements in Stock Prices and Comovements in Dividends, *Journal of Finance*, 44, 719-729.

Shleifer, A., 2000, *Inefficient Markets*, Oxford University Press Inc., New York.

Vayanos, D., 1998, Transaction Costs and Asset Prices: A Dynamic Equilibrium Model, *Review of Financial Studies*, 11, 1-58.

Vayanos, D., and Vila, J.-L., 1999, Equilibrium Interest Rate and Liquidity Premium with Transaction Costs, *Economic Theory*, 13, 509-539.

Vega, C., and Wu, J., 2008, Stock Market Microstructure Measures of Information Asymmetry are Related to Marketwide Information, Working Paper, Board of Governors of the Federal Reserve System.

Veldkamp, L., 2006, Information Markets and the Comovement of Asset Prices, *Review of Economic Studies*, 73, 823-845.

Yu, J., 2008, Commonality in Disagreement and Asset Pricing, Working Paper, Columbia University.

Yuan, K., 2005, Asymmetric Price Movements and Borrowing Constraints: A REE Model of Crisis, Contagion, and Confusion, *Journal of Finance*, 60, 379-411.

Table 1. Summary Statistics

This table reports summary statistics for daily, industry-level, equal-weighted returns $r_{n,t}$ (Eq. (12)) and net order flow $\omega_{n,t}$ (Eq. (13)) in our merged TAQ/CRSP/COMPUSTAT dataset between 1/1993 and 6/2004 (2,889 observations). Both variables are in percentage, i.e., are multiplied by 100. A "*", "**", or "***" indicate significance of the mean at the 10%, 5%, or 1% level, respectively.

	$r_{n,t}$ (%)					$\omega_{n,t}$ (%)				
	Mean	Median	Stdev	1%	99%	Mean	Median	Stdev	1%	99%
Durables	-0.004	0.029	1.201	-3.136	3.123	3.182***	3.112	8.287	-15.670	21.802
Nondurables	-0.004	0.038	0.839	-2.279	2.130	2.241***	2.148	7.583	-15.467	19.509
Manufacturing	-0.004	0.029	1.105	-2.904	3.071	3.530***	3.498	7.094	-13.407	19.421
Energy	0.023	0.033	1.363	-3.662	3.580	3.867***	4.228	8.687	-17.149	22.057
HighTech	-0.032	0.118	1.936	-5.330	5.052	-0.287**	0.092	7.318	-17.019	15.619
Telecom	-0.044	0.048	1.648	-4.851	4.360	0.685***	0.717	7.127	-15.697	16.607
Shops	-0.025	0.027	1.134	-2.902	2.924	2.273***	2.275	6.810	-13.499	17.890
Health	-0.003	0.065	1.294	-3.586	3.232	0.355***	0.322	6.876	-15.175	16.824
Utilities	-0.005	0.022	0.970	-2.933	2.484	-1.179***	-0.494	11.452	-27.582	21.391
Other	-0.007	0.063	1.114	-3.126	2.940	2.629***	2.636	6.974	-13.485	18.611

Table 2. Industry-Level Earnings Correlations

This table reports estimated Pearson correlations $\rho_{n,j}$ (over the sample period 1/1993-6/2004, 46 observations) among industry-level, equal-weighted averages of quarterly earnings (EPS basic, excluding extraordinary items) of the corresponding firms, $EPS_{n,q}$, as defined in Eq. (14), for each of the ten industries in our sample (Durables, Nondurables, Manufacturing, Energy, HighTech, Telecom, Shops, Health, Utilities, and Other). A "*", "**", or "***" indicate significance at the 10%, 5%, or 1% level, respectively.

	Durables	Nondurables	Manufacturing	Energy	HighTech	Telecom	Shops	Health	Utilities	Other
Durables	1									
Nondurables	0.026	1								
Manufacturing	0.577***	0.246*	1							
Energy	-0.079	0.049	0.009	1						
HighTech	0.395***	0.104	0.784***	-0.185	1					
Telecom	0.128	0.152	0.505***	-0.118	0.763***	1				
Shops	0.045	-0.021	0.304**	-0.240	0.583***	0.580***	1			
Health	0.235	0.107	0.537***	-0.012	0.658***	0.742***	0.565***	1		
Utilities	-0.035	0.059	0.266	0.394***	0.139	0.021	-0.047	0.188	1	
Other	0.423***	0.219	0.731***	-0.013	0.782***	0.660***	0.423***	0.593***	0.090	1

Table 3. Direct and Cross-Industry Price Impact

This table reports estimates of direct and cross-industry permanent price impact ($\sum_{l=0}^{L} \lambda_{nn,l}$, on the diagonal, and $\sum_{l=0}^{L} \lambda_{nj,l}$, off the diagonal, respectively) from the following regression model (Eq. (16)):

$$r_{n,t} = \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10} \sum_{l=1}^{L} \gamma_{nj,l} r_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l} \omega_{j,t-l} + \varepsilon_{n,t},$$

where $r_{n,t-l}$ is the equal weighted average of daily stock returns in industry portfolio n on day $t-l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{n,t-l}$ is the aggregate, equal-weighted, industry-level net order flow (net scaled number of transactions) on day $t-l$, and $L = 3$. We estimate Eq. (16) by OLS over the sample period 1/1993-6/2004 (2,889 observations) and assess the statistical significance of the estimated coefficients with Newey-West standard errors. Coefficient estimates are multiplied by 100. A " * ", " ** ", or " *** " indicate significance at the 10%, 5%, or 1% level, respectively. R_a^2 is the adjusted R^2.

	Durables	Nondurables	Manufacturing	Energy	HighTech	Telecom	Shops	Health	Utilities	Other	R_a^2
Durables	2.166***	-2.306***	1.769***	0.392	-1.610***	0.276	-0.029	0.852*	-0.149	-0.729	69%
Nondurables	0.339	2.556***	-0.369	-0.439*	-0.436	-0.297	0.313	-0.318	-0.388**	-0.684	67%
Manufacturing	1.047***	-0.833*	3.324***	-0.102	-0.868**	0.154	-0.140	-0.520**	-0.343**	-1.533***	84%
Energy	0.734	-2.749***	-1.447*	5.985***	3.773***	0.109	-1.542	-0.625	0.451	-2.296*	54%
HighTech	0.153	1.252***	-0.696	0.115	3.053***	-0.293	0.001	-1.028***	0.178	-2.478***	93%
Telecom	-0.982**	1.327**	1.250*	-1.042***	-1.170**	3.855***	-0.545	-0.458	-0.560**	-2.109**	81%
Shops	-0.625*	-0.157	-0.038	-0.141	-1.984***	-0.386	5.803***	0.159	-0.567***	-1.520***	84%
Health	0.512	-2.611***	-0.242	-0.021	-2.581***	0.530	0.040	4.670***	0.110	0.046	81%
Utilities	-0.405	-1.531**	0.050	0.617*	1.622***	-1.393***	-1.222	-0.197	1.549***	0.563	43%
Other	-0.029	-0.245	-0.666**	-0.478***	-1.462***	-0.572**	-1.379***	-0.286	-0.090	4.390***	90%

44

Table 4. Direct and Absolute Cross-Price Impact for Random Stock Pairs

This table reports estimates of average direct and absolute cross-stock permanent price impact ($\sum_{l=0}^{L} \lambda_{ii,l}$ and $\left|\sum_{l=0}^{L} \lambda_{ih,l}\right|$, respectively) accompanying a one standard deviation shock to the corresponding order flow ($\sigma(\omega_{i,t})$ and $\sigma(\omega_{h,t})$, respectively) in basis points (i.e., multiplied by 10,000) from the following regression model (Eq. (17)):

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l}\omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}\omega_{h,t-l} + \varepsilon_{i,t},$$

where $r_{i,t-l}$ is the equal weighted average of the daily returns of randomly selected stock i on day $t - l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{i,t-l}$ is the daily aggregate net order flow (net scaled number of transactions) in firm i on day $t - l$, and $L = 3$. We estimate Eq. (17) by OLS for two thousand randomly selected pairs of stocks with a common history of all quarterly earnings over each pair's longest common trading history in TAQ within the sample period 1/1993-6/2004. We then compute averages of the estimated coefficients within each quintile of firm pairs sorted according to their absolute earnings correlations ($\left|\rho_{i,h}\right|$) from the lowest to the highest, as well as over the fraction of the pairs (denoted as %*) for which those estimates are statistically significant at the 10% level (with Newey-West standard errors, marked as "*").

			Quintiles of $\left|\rho_{i,h}\right|$			
	Total	Low	2	3	4	High
$\left|\rho_{i,h}\right|$	0.245	0.040	0.124	0.214	0.330	0.518
%*	43%	0%	0%	17%	100%	100%
$\left|\rho_{i,h}\right|^*$	0.411	n.a.	n.a.	0.255	0.330	0.518
$\sigma(\omega_{i,t})\sum_{l=0}^{L} \lambda_{ii,l}$	40.38	38.83	38.59	41.69	42.19	40.61
%*	93%	93%	93%	94%	92%	92%
$\sigma(\omega_{i,t})\sum_{l=0}^{L} \lambda_{ii,l}^*$	43.03	41.15	40.95	43.82	45.38	43.89
$\sigma(\omega_{h,t})\left|\sum_{l=0}^{L} \lambda_{ih,l}\right|$	6.74	6.85	6.19	6.96	7.10	6.57
%*	14%	16%	13%	16%	16%	12%
$\sigma(\omega_{h,t})\left|\sum_{l=0}^{L} \lambda_{ih,l}\right|^*$	14.80	13.75	13.11	16.28	15.79	14.89

Table 5. Marketwide Information Heterogeneity: Direct and Cross-Industry Price Impact

This table reports the differences between estimates of direct and absolute cross-industry permanent price impact in days characterized by historically *high* and *low* information heterogeneity from the following regression model (Eq. (18)):

$$r_{n,t} = \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10} \sum_{l=1}^{L} \gamma_{nj,l} r_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l} \omega_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l}^{x} X_t \omega_{j,t-l} + \varepsilon_{n,t},$$

where $r_{n,t-l}$ is the equal weighted average of daily stock returns in industry portfolio n on day $t-l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{n,t-l}$ is the aggregate, equal-weighted, industry-level net order flow (net scaled number of transactions) on day $t-l$, X_t is either $SDLTEPS_m$ (the equal-weighted average of individual stock forecast standard deviations, Eq. (10), in Panel A) or $SDMMS_m$ (the simple average of the standardized dispersion of analyst forecasts of 18 macroeconomic variables, Eq. (11), in Panel B), and $L = 3$. We compute these differences as $\left| \sum_{l=0}^{L} \lambda_{nj,l} + \sum_{l=0}^{L} \lambda_{nj,l}^{x} X_{t,70^{th}} \right| - \left| \sum_{l=0}^{L} \lambda_{nj,l} + \sum_{l=0}^{L} \lambda_{nj,l}^{x} X_{t,30^{th}} \right|$, where $X_{t,70^{th}}$ and $X_{t,30^{th}}$ are the top 70th and bottom 30th percentile of the empirical distribution of X_t. We estimate Eq. (18) by OLS over the sample period 1/1993-6/2004 (2,889 observations) in Panel A, and over the subperiod 1/1993-12/2000 (2,017 observations) for Panel B, and assess the statistical significance of the estimated coefficients with Newey-West standard errors. Coefficient estimates are multiplied by 100. A "*", "**", or "***" indicate significance at the 10%, 5%, or 1% level, respectively. A "∘" indicates that neither sum is statistically significant but their difference is, or that only one sum is statistically significant and the difference is also significant but with the opposite sign. R_a^2 is the adjusted R^2.

46

Table 5. (*Continued*)

Panel A: $X_t = SDLTEPS_m$

	Durables	Nondurables	Manufacturing	Energy	HighTech	Telecom	Shops	Health	Utilities	Other	R_a^2
Durables	1.680*	0.754	1.001	-0.189	0.634	2.088**	1.777	0.446	0.246	1.901	70%
Nondurables	0.365	1.453*	1.309	-0.077	0.844	1.216**	0.414	-0.499	-0.042	2.804***	70%
Manufacturing	-1.276**	0.354	3.411***	0.287	-0.489	1.845***	0.511	0.799	-0.098	0.210	85%
Energy	0.063	-1.983	2.155	2.809***	0.376	-0.195	0.563	1.928	0.921	2.899	62%
HighTech	0.710	-0.190	0.568	0.104	-0.492	1.071*○	1.708*○	1.264**	0.350	-2.706**	94%
Telecom	-0.548	2.009	-2.739**	-0.249	-1.980**	3.364***	0.996	1.701	0.301	0.332	83%
Shops	-0.501	1.446***○	-0.248	0.052	1.839***	1.570***	3.694***	-0.165	0.148	1.292	86%
Health	0.905	-0.295	-0.078	0.073	3.393***	1.063	2.166**○	3.096***	0.155	0.182	83%
Utilities	-0.052	2.092*	-0.852	0.320	2.604***	0.740	-1.046	0.386	2.470***	2.808*○	50%
Other	0.165	0.350	0.008	0.188	1.669***	-0.117	0.007	-0.769*○	0.066	0.659	91%

Panel B: $X_t = SDMMS_m$

	Durables	Nondurables	Manufacturing	Energy	HighTech	Telecom	Shops	Health	Utilities	Other	R_a^2
Durables	0.084	0.044	0.473	-0.333	0.378	0.482	-0.167	0.178	0.188	0.189	61%
Nondurables	-0.018	0.410	-0.004	-0.373**○	-0.369	0.058	-0.438	-0.774**○	-0.037	0.712	64%
Manufacturing	-0.107	0.613	0.046	0.151	-0.029	-0.160	-0.783**	-0.038	0.013	0.114	79%
Energy	-0.340	-1.057	1.192*	-0.011	-0.957	-0.262	-0.851	-0.479	-0.338	-0.601	54%
HighTech	0.382	-0.192	-0.285	-0.081	0.587*	0.954***○	-0.329	-0.257	-0.020	-0.165	93%
Telecom	0.670*	1.097*	-0.615	0.228	0.352	0.897**	-0.710	-0.365	-0.182	0.435	81%
Shops	-0.228	-0.005	-0.458	-0.105	-0.178	-0.253	0.681	-0.202	0.337***	1.368***	84%
Health	0.158	-0.551	0.166	-0.497**	-0.050	0.451	0.533	-0.583	-0.210	-0.675	80%
Utilities	0.035	0.056	-0.039	0.314*○	0.407	0.298	-0.076	-0.226	0.503***	-0.407	49%
Other	0.148	-0.547*○	0.282	-0.010	0.577*	-0.177	0.365	-0.170	0.028	0.993***	89%

47

Table 6. Marketwide Information Heterogeneity: Direct and Absolute Cross-Price Impact for Random Stock Pairs

This table reports the differences between estimates of direct and absolute cross-stock permanent price impact in days characterized by historically *high* and *low* information heterogeneity, when accompanying a one standard deviation shock to the corresponding order flow ($\sigma(\omega_{i,t})$ and $\sigma(\omega_{h,t})$, respectively) in basis points (i.e., multiplied by 10,000), from the following regression model (Eq. (19)):

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l} \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l} \omega_{h,t-l}$$
$$+ \sum_{l=0}^{L} \lambda_{ii,l}^{x} X_t \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_t \omega_{h,t-l} + \varepsilon_{i,t},$$

where $r_{i,t-l}$ is the equal weighted average of the daily returns of randomly selected stock i on day $t - l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{i,t-l}$ is the daily aggregate net order flow (net scaled number of transactions) in firm i on day $t - l$, X_t is either $SDLTEPS_m$ (the equal-weighted average of individual stock forecast standard deviations, Eq. (10), in Panel A) or $SDMMS_m$ (the simple average of the standardized dispersion of analyst forecasts of 18 macroeconomic variables, Eq. (11), in Panel B), and $L = 3$. We compute these differences as $\sigma(\omega_{i,t}) \left(\sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,70^{th}} - \sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,30^{th}} \right)$, and $\sigma(\omega_{h,t}) \left(\left| \sum_{l=0}^{L} \lambda_{ih,l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,70^{th}} \right| - \left| \sum_{l=0}^{L} \lambda_{ih,l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,30^{th}} \right| \right)$, where $X_{t,70^{th}}$ and $X_{t,30^{th}}$ are the top 70th and bottom 30th percentile of the empirical distribution of X_t. We estimate Eq. (19) by OLS for two thousand randomly selected pairs of stocks with a common history of all quarterly earnings over each pair's longest common trading history in TAQ within the sample period 1/1993-6/2004 in Panel A and within the subperiod 1/1993-12/2000 in Panel B. We then compute averages of the estimated differences within each quintile of firm pairs sorted according to their absolute earnings correlations ($|\rho_{i,h}|$) from the lowest to the highest (in Table 4), as well as over the fraction of the pairs (denoted as %*) for which those estimates are statistically significant at the 10% level (with Newey-West standard errors, marked as " * ").

48

Table 6. (*Continued*)

			Quintiles of $\left|\rho_{i,h}\right|$			
	Total	Low	2	3	4	High
$\left|\rho_{i,h}\right|$	0.245	0.040	0.124	0.214	0.330	0.518

Panel A: $X_t = SDLTEPS_m$

	Total	Low	2	3	4	High
$\sigma\left(\omega_{i,t}\right)\left(\sum_{l=0}^{L}\lambda_{ii,l}^{x}X_{t,70^{th}} - \sum_{l=0}^{L}\lambda_{ii,l}^{x}X_{t,30^{th}}\right)$	20.83	20.57	18.74	21.11	22.93	20.80
$\%^*$	31%	29%	31%	32%	33%	31%
$\sigma\left(\omega_{i,t}\right)\left(\sum_{l=0}^{L}\lambda_{ii,l}^{x}X_{t,70^{th}} - \sum_{l=0}^{L}\lambda_{ii,l}^{x}X_{t,30^{th}}\right)^*$	35.67	35.61	34.90	36.76	37.06	33.91
$\sigma\left(\omega_{h,t}\right)\left(\left|\sum_{l=0}^{L}\lambda_{ih,l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_{t,70^{th}}\right| - \left|\sum_{l=0}^{L}\lambda_{ih,l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_{t,30^{th}}\right|\right)$	0.61	0.24	0.35	0.86	1.11	0.49
$\%^*$	18%	19%	17%	19%	18%	16%
$\sigma\left(\omega_{h,t}\right)\left(\left|\sum_{l=0}^{L}\lambda_{ih,l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_{t,70^{th}}\right| - \left|\sum_{l=0}^{L}\lambda_{ih,l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_{t,30^{th}}\right|\right)^*$	2.24	1.77	0.87	1.37	4.67	2.65

Panel B: $X_t = SDMMS_m$

	Total	Low	2	3	4	High
$\sigma\left(\omega_{i,t}\right)\left(\sum_{l=0}^{L}\lambda_{ii,l}^{x}X_{t,70^{th}} - \sum_{l=0}^{L}\lambda_{ii,l}^{x}X_{t,30^{th}}\right)$	11.77	13.03	10.92	11.67	12.70	10.53
$\%^*$	14%	15%	15%	15%	14%	11%
$\sigma\left(\omega_{i,t}\right)\left(\sum_{l=0}^{L}\lambda_{ii,l}^{x}X_{t,70^{th}} - \sum_{l=0}^{L}\lambda_{ii,l}^{x}X_{t,30^{th}}\right)^*$	17.70	14.79	16.56	20.81	14.76	23.01
$\sigma\left(\omega_{h,t}\right)\left(\left|\sum_{l=0}^{L}\lambda_{ih,l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_{t,70^{th}}\right| - \left|\sum_{l=0}^{L}\lambda_{ih,l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_{t,30^{th}}\right|\right)$	-0.34	-0.59	-0.13	0.24	-0.38	-0.84
$\%^*$	12%	11%	12%	14%	12%	12%
$\sigma\left(\omega_{h,t}\right)\left(\left|\sum_{l=0}^{L}\lambda_{ih,l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_{t,70^{th}}\right| - \left|\sum_{l=0}^{L}\lambda_{ih,l} + \sum_{l=0}^{L}\lambda_{ih,l}^{x}X_{t,30^{th}}\right|\right)^*$	0.55	-0.97	0.30	0.66	3.30	-0.63

49

Table 7. Marketwide Number of Speculators: Direct and Cross-Industry Price Impact

This table reports the differences between estimates of direct and absolute cross-industry permanent price impact in days characterized by a historically *high* and *low* number of speculators from the following regression model (Eq. (18)):

$$r_{n,t} = \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10} \sum_{l=1}^{L} \gamma_{nj,l} r_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l} \omega_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l}^{x} X_t \omega_{j,t-l} + \varepsilon_{n,t},$$

where $r_{n,t-l}$ is the equal weighted average of daily stock returns in industry portfolio n on day $t - l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{n,t-l}$ is the aggregate, equal-weighted, industry-level net order flow (net scaled number of transactions) on day $t - l$, X_t is ANA_m, the equal-weighted average of analyst coverage among the stocks in our sample (Eq. (20)), and $L = 3$. We compute these differences as $\left| \sum_{l=0}^{L} \lambda_{nj,l} + \sum_{l=0}^{L} \lambda_{nj,l}^{x} X_{t,70^{th}} \right| - \left| \sum_{l=0}^{L} \lambda_{nj,l} + \sum_{l=0}^{L} \lambda_{nj,l}^{x} X_{t,30^{th}} \right|$, where $X_{t,70^{th}}$ and $X_{t,30^{th}}$ are the top 70th and bottom 30th percentile of the empirical distribution of X_t. We estimate Eq. (18) by OLS over the sample period 1/1993-6/2004 (2,889 observations), and assess the statistical significance of the estimated coefficients with Newey-West standard errors. Coefficient estimates are multiplied by 100. A " * ", " ** ", " *** " indicate significance at the 10%, 5%, or 1% level, respectively. A "$_o$" indicates that neither sum is statistically significant but their difference is, or that only one sum is statistically significant and the difference is also significant but with the opposite sign. R_a^2 is the adjusted R^2.

	Durables	Nondurables	Manufacturing	Energy	HighTech	Telecom	Shops	Health	Utilities	Other	R_a^2
Durables	-0.349	-0.938	0.801	-0.224	-1.488**	-0.056	2.652**$_o$	-0.686	0.186	-1.973*$_o$	69%
Nondurables	0.337	-0.316	0.898	-0.050	-0.781*	0.191	0.577	0.120	-0.252	-0.434	69%
Manufacturing	-0.134	-0.088	1.480**	0.121	-0.498	0.042	0.942	-0.332	0.295	0.480	84%
Energy	-0.293	1.159	-0.531	-0.146	-0.028	1.441*$_o$	0.327	-0.441	1.249**	-0.147	55%
HighTech	0.185	-0.121	0.281	-0.149	-1.380***	-0.019	-0.971	-0.406	-0.540*	-0.109	93%
Telecom	0.211	-2.039***	1.155	-0.308	0.622	0.092	-1.172	-0.698	-0.627	0.878	82%
Shops	-0.687*	1.172**$_o$	1.489***	-0.100	-1.567***	0.123	-0.900	0.792*$_o$	0.136	-0.270	85%
Health	0.762	0.558	1.559**	0.010	-0.407	0.438	1.183*$_o$	-1.261**	-0.564*$_o$	0.003	82%
Utilities	1.638**	-1.008	1.967**	0.407	0.400	-0.503	1.644	-0.656	-1.259***	0.906	46%
Other	0.334	-0.789**$_o$	0.445	-0.212	-1.262***	0.022	0.053	0.215	-0.376**	-0.730	91%

50

Table 8. Marketwide Number of Speculators: Direct and Absolute Cross-Price Impact for Random Stock Pairs

This table reports the differences between estimates of direct and absolute cross-stock permanent price impact in days characterized by a historically *high* and *low* number of speculators, when accompanying a one standard deviation shock to the corresponding order flow ($\sigma(\omega_{i,t})$ and $\sigma(\omega_{h,t})$, respectively) in basis points (i.e., multiplied by 10,000), from the following regression model (Eq. (19)):

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l} \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l} \omega_{h,t-l}$$
$$+ \sum_{l=0}^{L} \lambda_{ii,l}^{x} X_m \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_m \omega_{h,t-l} + \varepsilon_{i,t},$$

where $r_{i,t-l}$ is the equal weighted average of the daily returns of randomly selected stock i on day $t-l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{i,t-l}$ is the daily aggregate net order flow (net scaled number of transactions) in firm i on day $t-l$, X_t is ANA_m, the equal-weighted average of analyst coverage among the stocks in our sample (Eq. (20)), and $L = 3$. We compute these differences as $\sigma(\omega_{i,t})\left(\sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,70^{th}} - \sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,30^{th}}\right)$, and $\sigma(\omega_{h,t})\left(\left|\sum_{l=0}^{L} \lambda_{ih,l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,70^{th}}\right| - \left|\sum_{l=0}^{L} \lambda_{ih,l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,30^{th}}\right|\right)$, where $X_{t,70^{th}}$ and $X_{t,30^{th}}$ are the top 70th and bottom 30th percentile of the empirical distribution of X_t. We estimate Eq. (19) by OLS for two thousand randomly selected pairs of stocks with a common history of all quarterly earnings over each pair's longest common trading history in TAQ within the sample period 1/1993-6/2004. We then compute averages of the estimated differences within each quintile of firm pairs sorted according to their absolute earnings correlations ($|\rho_{i,h}|$) from the lowest to the highest (in Table 4), as well as over the fraction of the pairs (denoted as %*) for which those estimates are statistically significant at the 10% level (with Newey-West standard errors, marked as " * ").

		Quintiles of $|\rho_{i,h}|$				
	Total	Low	2	3	4	High
$|\rho_{i,h}|$	0.245	0.040	0.124	0.214	0.330	0.518
$\sigma(\omega_{i,t})\left(\sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,70^{th}} - \sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,30^{th}}\right)$	-0.57	-0.81	-0.83	-0.93	-0.34	0.08
%*	18%	19%	19%	19%	19%	16%
$\sigma(\omega_{i,t})\left(\sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,70^{th}} - \sum_{l=0}^{L} \lambda_{ii,l}^{x} X_{t,30^{th}}\right)^{*}$	-3.11	-5.07	-3.89	-2.81	-2.92	-0.52
$\sigma(\omega_{h,t})\left(\left|\sum_{l=0}^{L} \lambda_{ih,l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,70^{th}}\right| - \left|\sum_{l=0}^{L} \lambda_{ih,l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,30^{th}}\right|\right)$	0.17	-0.13	0.34	0.43	-0.01	0.22
%*	12%	12%	14%	10%	11%	14%
$\sigma(\omega_{h,t})\left(\left|\sum_{l=0}^{L} \lambda_{ih,l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,70^{th}}\right| - \left|\sum_{l=0}^{L} \lambda_{ih,l} + \sum_{l=0}^{L} \lambda_{ih,l}^{x} X_{t,30^{th}}\right|\right)^{*}$	-0.10	-0.21	-0.03	-2.99	0.13	1.66

Table 9. Price Observability: Direct and Absolute Cross-Price Impact for Random Stock Pairs

This table reports estimates of average direct and cross-stock permanent price impact ($\sum_{l=0}^{L} \lambda_{ii,l}$ and $\sum_{l=0}^{L} \lambda_{hh,l}$, and $\left|\sum_{l=0}^{L} \lambda_{ih,l}\right|$ and $\left|\sum_{l=0}^{L} \lambda_{hi,l}\right|$, respectively) accompanying a one standard deviation shock to the corresponding order flow ($\sigma(\omega_{i,t})$ and $\sigma(\omega_{h,t})$, respectively) in basis points (i.e., multiplied by 10,000) from the following regression model (Eq. (17)):

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l} \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l} \omega_{h,t-l} + \varepsilon_{i,t},$$

where $r_{i,t-l}$ is the equal weighted average of the daily returns of randomly selected stock i on day $t-l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{i,t-l}$ is the daily aggregate net order flow (net scaled number of transactions) in firm i on day $t-l$, and $L=3$. We estimate Eq. (17) by OLS for eighty randomly selected pairs of NYSE stocks with a common history of all quarterly earnings over each pair's longest common trading history in TAQ within the sample period 11/2001-6/2004, as well as dealt by either the same (Panel A) or different specialist and specialist firm (Panel B). We then compute averages of the estimated coefficients within each quintile of firm pairs sorted according to their absolute earnings correlations ($|\rho_{i,h}|$) from the lowest to the highest, as well as over the fraction of the pairs (denoted as %*) for which those estimates are statistically significant at the 10% level (with Newey-West standard errors, marked as " * ").

	Panel A: Same specialist						Panel B: Different specialist					
		Quintiles of $\lvert\rho_{i,h}\rvert$						Quintiles of $\lvert\rho_{i,h}\rvert$				
	total	low	2	3	4	high	total	low	2	3	4	high
$\lvert\rho_{i,h}\rvert$	0.236	0.029	0.112	0.193	0.318	0.526	0.256	0.035	0.106	0.229	0.382	0.529
%*	35%	0%	0%	13%	63%	100%	53%	0%	0%	63%	100%	100%
$\lvert\rho_{i,h}\rvert^{*}$	0.425	n.a.	n.a.	0.219	0.306	0.526	0.409	n.a.	n.a.	0.260	0.382	0.529
$\sigma(\omega_{i,t}) \sum_{l=0}^{L} \lambda_{ii,l}$	26.23	19.35	23.53	20.87	41.31	26.09	24.48	24.68	29.13	22.54	26.84	19.22
%*	56%	50%	56%	50%	75%	50%	70%	81%	75%	75%	75%	44%
$\sigma(\omega_{i,t}) \sum_{l=0}^{L} \lambda_{ii,l}^{*}$	37.76	27.52	38.08	28.48	49.80	38.86	30.70	28.23	35.61	26.96	34.12	27.45
$\sigma(\omega_{h,t}) \left\lvert\sum_{l=0}^{L} \lambda_{ih,l}\right\rvert$	13.08	15.25	14.09	11.15	10.63	14.31	10.80	12.02	9.87	10.18	11.12	10.78
%*	19%	25%	25%	25%	6%	13%	19%	19%	19%	31%	6%	19%
$\sigma(\omega_{h,t}) \left\lvert\sum_{l=0}^{L} \lambda_{ih,l}\right\rvert^{*}$	24.00	29.82	22.87	19.26	19.70	26.25	21.71	26.90	25.55	16.62	19.43	21.91

Table 10. Macroeconomic News: Direct and Cross-Industry Price Impact

This table reports the differences between estimates of direct and absolute cross-industry permanent price impact in announcement and non-announcement days from the following regression model (Eq. (25)):

$$r_{n,t} = \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10}\sum_{l=1}^{L}\gamma_{nj,l}r_{j,t-l} + \sum_{j=1}^{10}\sum_{l=0}^{L}\lambda_{nj,l}(1-D_t^p)\,\omega_{j,t-l} + \sum_{j=1}^{10}\sum_{l=0}^{L}\lambda_{nj,l}^p D_t^p \omega_{j,t-l} + \varepsilon_{n,t},$$

where $r_{n,t-l}$ is the equal weighted average of daily stock returns in industry portfolio n on day $t-l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{n,t-l}$ is the aggregate, equal-weighted, industry-level net order flow (net scaled number of transactions) on day $t-l$, D_t^p is a dummy variable equal to one if any of the 18 macroeconomic news in our sample is released on day t and equal to zero otherwise, and $L=3$. We estimate Eq. (25) by OLS over the sample period 1/1993-6/2004 (2,889 observations), and assess the statistical significance of the resulting differences of estimated coefficients with Newey-West standard errors. Coefficient estimates are multiplied by 100. A " * ", " ** ", or " *** " indicate significance at the 10%, 5%, or 1% level, respectively. A "○" indicates that neither sum is statistically significant but their difference is, or that only one sum is statistically significant and the difference is also significant but with the opposite sign. R_a^2 is the adjusted R^2.

	Durables	Nondurables	Manufacturing	Energy	HighTech	Telecom	Shops	Health	Utilities	Other	R_a^2
Durables	0.122	-0.809	-2.327	-0.186	-0.473	-3.125**	-1.259	-0.715	-1.874***	1.236	69%
Nondurables	-0.615	1.204	1.697	0.011	0.601	0.925	-4.180***	0.224	0.740*	0.861	67%
Manufacturing	0.008	-0.794	-0.101	-0.001	1.288	-1.486*○	-1.843	0.398	0.616	2.270*	84%
Energy	-0.272	-1.997	3.355	-0.900	-3.187	-0.025	-5.319**	0.643	-1.528*	-1.425	54%
HighTech	0.155	-0.112	-0.667	0.086	0.554	0.887	-3.574**	-1.044	0.648	-0.829	93%
Telecom	-0.114	0.440	1.040	1.174	-0.118	0.411	-2.398	2.228	-0.108	0.361	81%
Shops	0.638	0.468	-0.488	0.058	0.412	-0.010	3.357**	-0.414	-0.283	2.039	84%
Health	-0.794	1.844	-1.834	-1.190*○	0.696	-1.032	-2.337	1.569	1.318***	0.159	81%
Utilities	-0.012	-1.220	-2.544*○	-1.571**	0.334	-0.270	-2.541	0.574	0.271	-0.412	43%
Other	-0.339	-0.713	0.433	0.110	-0.632	-0.599	-0.055	0.931	0.380	0.145	90%

53

Table 11. Macroeconomic News: Direct and Absolute Cross-Price Impact for Random Stock Pairs

This table reports the differences between estimates of direct and absolute cross-stock permanent price impact in non-announcement and non-announcement days accompanying a one standard deviation shock to the corresponding order flow ($\sigma\left(\omega_{i,t}\right)$ and $\sigma\left(\omega_{h,t}\right)$, respectively) in basis points (i.e., multiplied by 10,000) from the following regression model (Eq. (26)):

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} + \sum_{l=1}^{L} \lambda_{ii,l} \left(1 - D_t^p\right) \omega_{i,t-l}$$
$$+ \sum_{l=0}^{L} \lambda_{ih,l} \left(1 - D_t^p\right) \omega_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l}^p D_t^p \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}^p D_t^p \omega_{h,t-l} + \varepsilon_{i,t},$$

where $r_{i,t-l}$ is the equal weighted average of the daily returns of randomly selected stock i on day $t - l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{i,t-l}$ is the daily aggregate net order flow (net scaled number of transactions) in firm i on day $t - l$, D_t^p is a dummy variable equal to one if any of the 18 macroeconomic news in our sample is released on day t and equal to zero otherwise, and $L = 3$. We compute these differences as $\sigma\left(\omega_{i,t}\right) \left(\sum_{l=0}^{L} \lambda_{ii,l}^p - \sum_{l=0}^{L} \lambda_{ii,l}\right)$ and $\sigma\left(\omega_{h,t}\right) \left(\left|\sum_{l=0}^{L} \lambda_{ih,l}^p\right| - \left|\sum_{l=0}^{L} \lambda_{ih,l}\right|\right)$, respectively. We estimate Eq. (19) by OLS for two thousand randomly selected pairs of stocks with a common history of all quarterly earnings over each pair's longest common trading history in TAQ within the sample period 1/1993-6/2004. We then compute averages of the estimated differences within each quintile of firm pairs sorted according to their absolute earnings correlations ($\left|\rho_{i,h}\right|$) from the lowest to the highest, as well as over the fraction of the pairs (denoted as %*) for which those estimates are statistically significant at the 10% level (with Newey-West standard errors, marked as " * ").

		Quintiles of $\left|\rho_{i,h}\right|$				
	total	Low	2	3	4	High
$\left|\rho_{i,h}\right|$	0.245	0.040	0.124	0.214	0.330	0.518
$\sigma\left(\omega_{i,t}\right) \left(\sum_{l=0}^{L} \lambda_{ii,l}^p - \sum_{l=0}^{L} \lambda_{ii,l}\right)$	-1.88	-1.46	-1.63	-2.25	-1.58	-2.46
%*	8%	8%	9%	7%	5%	12%
$\sigma\left(\omega_{i,t}\right) \left(\sum_{l=0}^{L} \lambda_{ii,l}^p - \sum_{l=0}^{L} \lambda_{ii,l}\right)^*$	-0.33	1.76	4.83	-13.63	4.00	0.50
$\sigma\left(\omega_{h,t}\right) \left(\left|\sum_{l=0}^{L} \lambda_{ih,l}^p\right| - \left|\sum_{l=0}^{L} \lambda_{ih,l}\right|\right)$	-3.56	-2.28	-3.29	-3.91	-3.21	-5.11
%*	10%	9%	9%	9%	9%	13%
$\sigma\left(\omega_{h,t}\right) \left(\left|\sum_{l=0}^{L} \lambda_{ih,l}^p\right| - \left|\sum_{l=0}^{L} \lambda_{ih,l}\right|\right)^*$	-13.54	-6.80	-12.27	-11.66	-12.90	-21.06

Table 12. Macroeconomic News: Further Estimates of Direct and Cross-Industry Price Impact

This table reports estimates of direct and absolute cross-industry permanent price impact from the following regression model (Eq. (27)):

$$r_{n,t} = \alpha_n + \beta_n r_{M,t} + \sum_{j=1}^{10} \sum_{l=1}^{L} \gamma_{nj,l} r_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l} \left(1 - D_t^p\right) \omega_{j,t-l}$$
$$+ \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l}^p D_t^p \omega_{j,t-l} + \sum_{j=1}^{10} \sum_{l=0}^{L} \lambda_{nj,l}^{px} D_t^p X_t^p \omega_{j,t-l} + \varepsilon_{n,t},$$

where $r_{n,t-l}$ is the equal weighted average of daily stock returns in industry portfolio n on day $t-l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{n,t-l}$ is the aggregate, equal-weighted, industry-level net order flow (net scaled number of transactions) on day $t-l$, D_t^p is a dummy equal to one if Nonfarm Payroll Employment news is released on day t and equal to zero otherwise, $X_t^p = ABSREV_t^p$ is the corresponding absolute revision (from RTDS), and $L = 3$. We estimate Eq. (27) by OLS over the sample period 1/1993-06/2004 (2,889 observations), and assess the statistical significance of the resulting differences of estimated coefficients — $\left| \sum_{l=0}^{L} \lambda_{nj,l}^p + \sum_{l=0}^{L} \lambda_{nj,l}^{px} X_{t,70^{th}}^p \right| - \left| \sum_{l=0}^{L} \lambda_{nj,l}^p + \sum_{l=0}^{L} \lambda_{nj,l}^{px} X_{t,30^{th}}^p \right|$ — for each industry pair — with Newey-West standard errors. Coefficient estimates are multiplied by 100. A "∗", "∗∗", or "∗∗∗" indicate significance at the 10%, 5%, or 1% level, respectively. A "∘" indicates that neither sum is statistically significant but their difference is, or that only one sum is statistically significant and the difference is also significant but with the opposite sign. R_a^2 is the adjusted R^2.

	Durables	Nondurables	Manufacturing	Energy	HighTech	Telecom	Shops	Health	Utilities	Other	R_a^2
Durables	0.236	0.071	3.484	3.039*∘	-3.220	-4.915**∘	-10.509**	4.655	1.990	-14.503***∘	69%
Nondurables	1.295	0.146	3.971	1.142	-1.163	-2.605*∘	-3.253	-1.765	0.553	-4.134	67%
Manufacturing	-1.672	-0.033	-0.356	1.579*∘	-4.206***	1.364	5.585*∘	-6.128***	1.947**∘	-4.518*∘	84%
Energy	2.780*∘	0.539	11.838**	-0.217	4.352**∘	-2.628	-5.694	-8.979***∘	2.643*∘	0.446	54%
HighTech	-1.530	2.180	0.546	0.822	-0.878	-1.195	-5.599	3.476	2.388**∘	2.102	93%
Telecom	2.637**∘	-0.856	11.627***	0.328	3.709*∘	-1.162	0.827	5.800**∘	-1.503	5.217	81%
Shops	2.416**∘	1.803	5.352*∘	-0.657	0.879	1.022	-3.095	-0.717	0.867	-4.045	84%
Health	1.508	-1.916	1.222	-0.508	-2.404*	-0.428	-0.857	-3.353*∘	0.058	0.104	81%
Utilities	3.415*	-2.452	3.298	2.267	0.380	8.001***∘	-6.686	-5.360	0.419	-1.637	45%
Other	-0.692	2.230*∘	0.079	0.082	0.167	1.127	-1.009	-1.045	1.389**∘	-0.544	90%

Table 13. Macroeconomic News: Further Estimates of Direct and Absolute Cross-Price Impact for Random Stock Pairs

This table reports estimates of direct and absolute cross-stock permanent price impact accompanying a one standard deviation shock to the corresponding order flow ($\sigma(\omega_{i,t})$ and $\sigma(\omega_{h,t})$, respectively) in basis points (i.e., multiplied by 10,000) from the following regression model (Eq. (28)):

$$r_{i,t} = \alpha_i + \beta_i r_{M,t} + \sum_{l=1}^{L} \gamma_{ii,l} r_{i,t-l} + \sum_{l=1}^{L} \gamma_{ih,l} r_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l}(1-D_t^p)\,\omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}(1-D_t^p)\,\omega_{h,t-l}$$
$$+ \sum_{l=0}^{L} \lambda_{ii,l}^{p} D_t^p \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}^{p} D_t^p \omega_{h,t-l} + \sum_{l=0}^{L} \lambda_{ii,l}^{px} D_t^p X_t^p \omega_{i,t-l} + \sum_{l=0}^{L} \lambda_{ih,l}^{px} D_t^p X_t^p \omega_{h,t-l} + \varepsilon_{i,t},$$

where $r_{n,t-l}$ is the equal weighted average of daily stock returns in industry portfolio n on day $t-l$, $r_{M,t}$ is the equal-weighted daily return on all NYSE and NASDAQ stocks, $\omega_{n,t-l}$ is the aggregate, equal-weighted, industry-level net order flow (net scaled number of transactions) on day $t-l$, D_t^p is a dummy equal to one if Nonfarm Payroll Employment news is released on day t and equal to zero otherwise, $X_t^p = ABSREV_t^p$ is the corresponding absolute revision (from RTDS), and $L = 3$. We estimate Eq. (28) by OLS for two thousand randomly selected pairs of stocks with a common history of all quarterly earnings over each pair's longest common trading history in TAQ within the sample period 1/1993-6/2004. We then compute averages of the differences of the resulting estimated coefficients $-\sigma(\omega_{i,t})\left(\sum_{l=0}^{L}\lambda_{ii,l}^{px}X_{t,70^{th}}^{p} - \sum_{l=0}^{L}\lambda_{ii,l}^{px}X_{t,30^{th}}^{p}\right)$ and $\sigma(\omega_{h,t})\left(\left|\sum_{l=0}^{L}\lambda_{ih,l}^{p}+\sum_{l=0}^{L}\lambda_{ih,l}^{px}X_{t,70^{th}}^{p}\right| - \left|\sum_{l=0}^{L}\lambda_{ih,l}^{p}+\sum_{l=0}^{L}\lambda_{ih,l}^{px}X_{t,30^{th}}^{p}\right|\right)$ — within each quintile of firm pairs sorted according to their absolute earnings correlations $\left|\rho_{i,h}\right|$ from the lowest to the highest, as well as over the fraction of the pairs (denoted as %*) for which those estimates are statistically significant at the 10% level (with Newey-West standard errors, marked as "*").

	Total	Quintiles of $\left|\rho_{i,h}\right|$				
		Low	2	3	4	High
$\left|\rho_{i,h}\right|$	0.245	0.040	0.124	0.214	0.330	0.518
$\sigma(\omega_{i,t})\left(\sum_{l=0}^{L}\lambda_{ii,l}^{px}X_{t,70^{th}}^{p} - \sum_{l=0}^{L}\lambda_{ii,l}^{px}X_{t,30^{th}}^{p}\right)$	-3.54	-6.12	-3.24	-1.04	-4.15	-3.14
%*	13%	15%	15%	13%	12%	11%
$\sigma(\omega_{i,t})\left(\sum_{l=0}^{L}\lambda_{ii,l}^{px}X_{t,70^{th}}^{p} - \sum_{l=0}^{L}\lambda_{ii,l}^{px}X_{t,30^{th}}^{p}\right)^*$	-6.10	-16.07	-2.27	-0.17	-15.21	5.41
$\sigma(\omega_{h,t})\left(\left|\sum_{l=0}^{L}\lambda_{ih,l}^{p}+\sum_{l=0}^{L}\lambda_{ih,l}^{px}X_{t,70^{th}}^{p}\right| - \left|\sum_{l=0}^{L}\lambda_{ih,l}^{p}+\sum_{l=0}^{L}\lambda_{ih,l}^{px}X_{t,30^{th}}^{p}\right|\right)$	-3.07	-5.72	-5.31	0.53	-3.61	-1.26
%*	13%	16%	11%	12%	14%	11%
$\sigma(\omega_{h,t})\left(\left|\sum_{l=0}^{L}\lambda_{ih,l}^{p}+\sum_{l=0}^{L}\lambda_{ih,l}^{px}X_{t,70^{th}}^{p}\right| - \left|\sum_{l=0}^{L}\lambda_{ih,l}^{p}+\sum_{l=0}^{L}\lambda_{ih,l}^{px}X_{t,30^{th}}^{p}\right|\right)^*$	-11.27	-24.51	-18.89	-4.21	-11.40	8.54

## Figure 1. Three-Asset Economy: Measures of Liquidity

This figure plots measures of direct $(\Lambda\left(1,1\right))$ and cross-asset $(\Lambda\left(1,3\right))$ liquidity for the three-asset economy $(N=3)$ parametrized in Appendix B as a function of the degree of information heterogeneity among speculators (ρ) in the presence of few $(M=5)$ or many $(M=500)$ of them. Specifically, we plot $\Lambda\left(1,1\right)$ and $\Lambda\left(1,3\right)$ as a function of $\rho\in\left(0,1\right)$ when $M=5$ (Figures 1A and 1C, respectively) or 500 (Figures 1B and 1D, respectively) and either no public signal of v is available $(\Lambda$, continuous line) or a public signal of v $(S_p$ of precision $\psi_p=0.5)$ is released $(\Lambda_p$, dotted line).

a) $\Lambda\left(1,1\right)$ for $M=5$
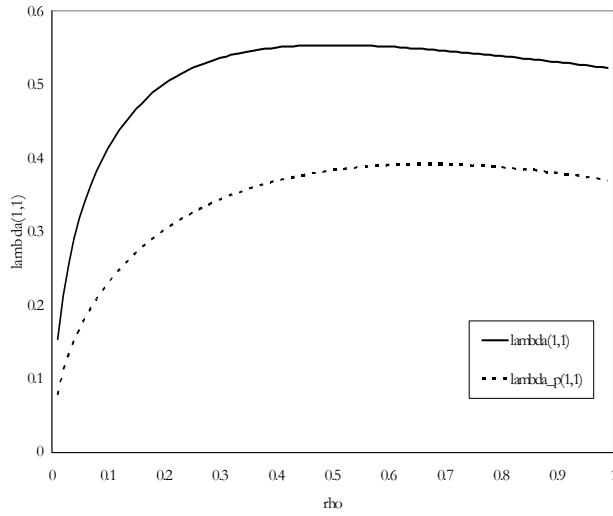
b) $\Lambda\left(1,1\right)$ for $M=500$
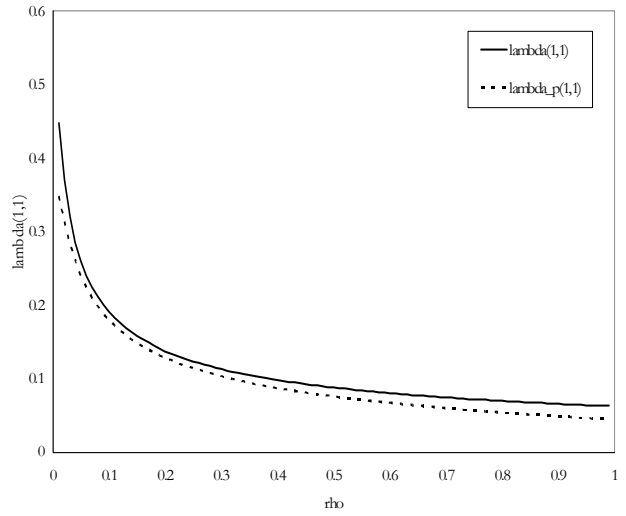
c) $\Lambda\left(1,3\right)$ for $M=5$

### d) $\Lambda\left(1,3\right)$ for $M=500$

57
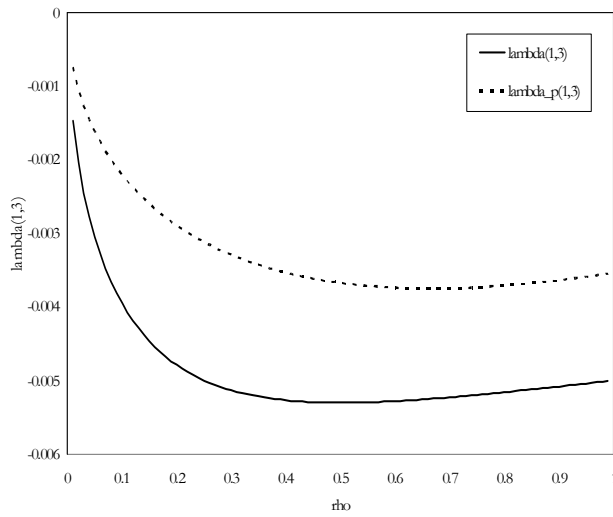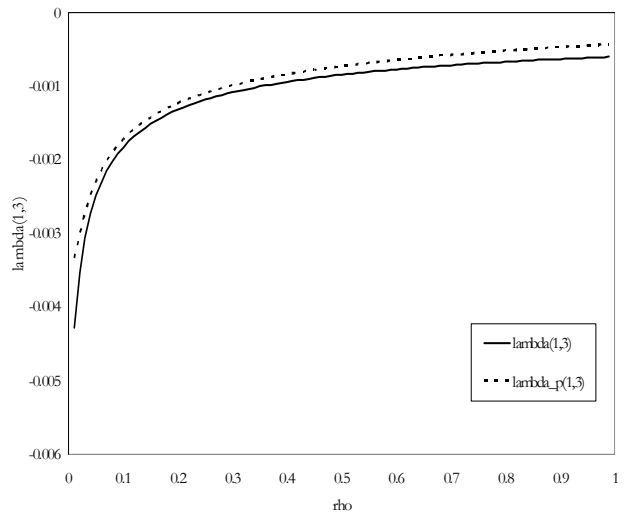
Figure 2. Plots of Marketwide Aggregates

Figure 2a plots $SDLTEPS_m$ (Eq. (10), continuous line), the equal-weighted average of firm-level standard deviations of analyst forecasts of long-term EPS growth (from I/B/E/S) in month m and $VWSDLTEPS_m$ (dashed line) the corresponding value-weighted average. Figure 2b plots $SDMMS_m$ (Eq. (11), continuous line), the simple average of standardized standard deviation of professional forecasts of 18 U.S. macroeconomic announcements (from MMS). Figure 2c plots ANA_m (Eq. (20), continuous line), the equal-weighted average of the number of analysts covering each of the firms in our sample in month m (from I/B/E/S). Figure 2d plots $EPSVOL_q$ (Eq. (21), continuous line), the equal-weighted average of firm-level earnings volatility in calendar quarter q (from COMPUSTAT). Figure 2e plots $EURVOL_m$ (continuous line), the monthly average of daily Eurodollar implied volatility (from Bloomberg). Figure 2f plots $RISKAV_m$ (Eq. (22), continuous line), the monthly difference between the end-of-month VIX index of implied volatility of S&P500 options with 30-day fixed maturity and the realized volatility of intraday S&P500 returns over that month (from Bollerslev and Zhou, 2007).
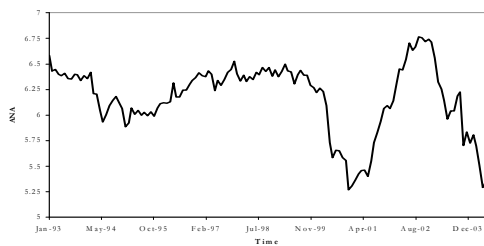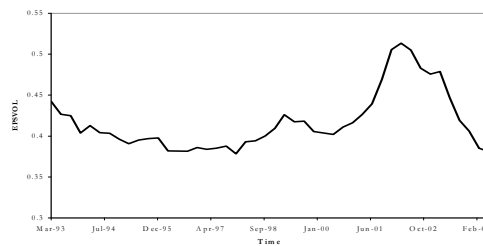
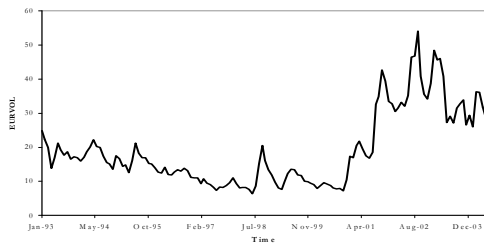a) $SDLTEPS_m$ & $VWSDLTEPS_m$
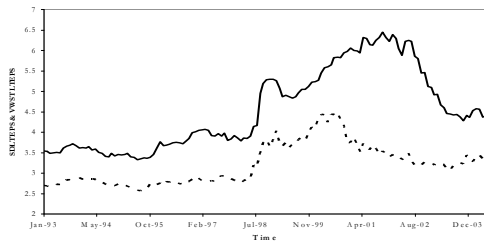
b) $SDMMS_m$
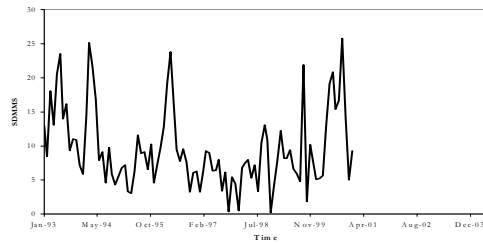
c) ANA_m

d) $EPSVOL_q$

e) $EURVOL_m$

f) $RISKAV_m$