Disagreement and Liquidity^{*}

Samuel Kruger[†]

January 28, 2019

Disagreement is an increasingly popular explanation for trade of informationally sensitive securities. Yet, there is limited research on the relation between disagreement and liquidity, particularly regarding how disagreement affects private information's impact on trading and liquidity. This paper proposes a model in which trading is entirely generated by disagreement stemming from overconfident interpretation of private signals. Contrary to traditional intuition, the model predicts that private information increases trading and enhances liquidity. A more general version of the model incorporates both disagreement and liquidity trading. The general model relates traditional intuition about private information destroying liquidity and trade to disagreement trading. The relation between private information and liquidity is not monotonic. Private information at first decreases liquidity but then enhances it, potentially explaining why private information seems to destroy liquidity in money markets but not in markets that are more informationally sensitive to start with.

JEL Codes: D82, G12, G14, G41

keywords: disagreement, liquidity, private information, overconfidence

^{*}I thank John Campbell, Josh Coval, Robin Greenwood, John Griffin, Sam Hanson, Travis Johnson, Richard Lowery, David Scharfstein, Andrei Shleifer, Alp Simsek, Erik Stafford, Jeremy Stein, Adi Sunderam, Anjan Thakor, and Luis Viceira, as well as seminar participants at Harvard and the University of Texas at Austin for helpful comments. Supplemental results can be found in an internet appendix at the author's website.

[†]The University of Texas at Austin, Department of Finance, McCombs School of Business, 2110 Speedway, B6600, Austin, TX 78712. Email: sam.kruger@mcombs.utexas.edu.

1 Introduction

How does private information affect liquidity when trading is primarily due to disagreement? Extensive trading of equities and other informationally sensitive securities is a puzzle for traditional models. Standard asset pricing models have no role for trading, and models that consider trading typically predict that private information decreases trading and destroys liquidity, defined as the ability to trade an asset without significantly changing its price. A growing literature posits that trade in equity markets is largely driven by disagreement, likely stemming from a combination of heterogeneous information and behavioral bias. While this explanation for trade is increasingly popular, there is little existing research on the relation between disagreement and liquidity and on how disagreement affects private information's impact on trading and liquidity. Holmstrom (2008) highlights the complex relationship between disagreement and liquidity by comparing money (liquidity) markets to stock markets: "Markets for liquidity are very different than stock markets. In the stock market, uncertainty and adverse selection fears are present all the time, but this does not prevent the markets from functioning.... Differences in beliefs often alleviate adverse selection. Stock markets thrive on differences in beliefs. Markets for liquidity are killed by them."

Since Akerlof (1970), economists have recognized that private information has the potential to destroy trade. To overcome private information, at least some trade must be motivated by something other than rationally processed information; otherwise, liquidity will dry up, markets will freeze, and the no-trade prediction of Milgrom and Stokey (1982) will prevail. Starting with the noisy rational expectations models of Grossman and Stiglitz (1980), Hellwig (1980), and Diamond and Verrecchia (1981) and including virtually all research on liquidity and market microstructure, this extra trading has been modeled as exogenous noise.

Traditional private information models make two basic predictions: (1) Significant noise trading is necessary to generate trading and liquidity; and (2) private information decreases trading and liquidity. Hong and Stein (2007) address the first prediction, noting that trading volume in traditional models is approximately pinned down by noise (non-informational) trading volume. For example, in Kyle's (1985) model exogenous noise trading represents half of total order flow variance, and in a moderate variance calibration of Diamond and Verrecchia's (1981) model, I find that liquidity trading is 95% of total trading volume. The New York Stock Exchange has daily volumes

in excess of \$30 billion. Liquidity trading of that magnitude seems implausible. One way out of this problem is to interpret noise trading more generally and explain most trading as uninformed noise from irrational traders. While theoretically possible, this is not particularly satisfying. If most trading is exogenously assumed, we aren't really left with a model of trading. Moreover, traditional models assume that noise trading is not just exogenous but also orthogonal to information. This is a critical assumption, and it is likely invalid if noise trading is driven by disagreement among market participants. For example, if disagreement comes from overconfident interpretation of private information, the same trade is at once informative and noise, making noise trading perfectly correlated with information.

Prediction (2) presents even more fundamental problems. In contrast to traditional intuition, trade increases with proxies for private information in the data. Moreover, Collin-Dufresne and Fos (2015) and Kacpercyk and Pagnotta (2018) find that markets are more liquid when informed traders are active.¹ Section 5 examines private information, turnover, and liquidity of stocks, corporate bonds, and stock options.² The analysis supports three stylized facts: (1) Trade and liquidity are positively correlated; (2) as private information increases relative to public information, trade increases and liquidity decreases; and (3) trade and liquidity are elevated following high past returns.³ Fact (1) supports the notion that trade and liquidity reinforce one another. Fact (2) contradicts the prediction of traditional models that private information destroys trading. Fact (3) shows that traditional models leave out something related to past returns.

Disagreement trading has become an increasingly popular explanation for trading volume and has the potential to explain observed patterns in private information, trade, and liquidity. To better understand these relations, I consider a model in which trading is primarily driven by disagreement. I.e., people trade because they have different beliefs about an asset's value. Counterparties

¹Collin-Dufresne and Fos (2015) study 13D filings by activist investors. Kacpercyk and Pagnotta (2018) study Securities Exchange Commission insider trading investigations. Both papers find that informed trade is associated with increased liquidity in contrast to predictions of traditional models. This pattern seems to be at least in part explained by endogenous trade timing and limit orders.

²I proxy for private information with analyst earnings forecast dispersion and also study periods around earnings announcements, which likely have elevated private information. For illiquidity, I use several measures of bid-ask spreads as well as Amihud's (2002) $illiq = \frac{|Return|}{\$Volume}$ measure.

³In equity markets, evidence for these facts has been shown before. In particular, Sadka and Scherbina (2007) show that liquidity decreases with analyst forecast dispersion; Lee, Mucklow, and Ready (1993) show that liquidity decreases around earnings announcements; Frazzini and Lamont (2007) show that turnover is elevated around earnings announcements; and Statman, Thorley, and Vorkink (2006) and Griffin, Nardari, and Stulz (2007) show that trading increases following high returns. I extend these findings to corporate bonds and stock options and show that trading increases with analyst dispersion and liquidity increases with past returns.

essentially make zero-sum bets about asset values, and they do so fully aware that other parties disagree with them. In the model, belief differences stem from overconfident interpretation of private information. The overconfidence bias creates trade and liquidity even when prices fully reveal the beliefs of other agents. Moreover, because disagreement stems from private signals, trading and liquidity increase with private information. Formally, the model considers disagreement trading among ex ante homogenous agents who simultaneously serve as informed traders, noise traders, and market makers. Agents are risk averse, receive endowments of a risky asset, observe private signals about the asset's value, and trade the asset with one another in an anonymous public market in which market price is visible to all agents. Agents are fully rational except for an overconfidence bias, which causes them to overestimate the precision of their own signals. I consider two versions of the model, a baseline model in which asset endowments are constant and a general model in which asset endowments are stochastic. The baseline model is an adaptation of Grossman (1976) with overconfidence. The general model is an adaptation of Diamond and Verrecchia (1981) with overconfidence.

The baseline model describes an environment in which trading is entirely driven by disagreement. Because supply is certain and there is no non-informational motive for trade, prices fully reveal aggregate information (as in Grossman, 1976). Nonetheless, overconfidence induces agents to disagree and trade, thereby generating liquidity (in contrast to Grossman, 1976). The model's main predictions are: (1) Overconfidence increases trading and liquidity; (2) Private information increases trading and liquidity; and (3) Public uncertainty decreases liquidity without affecting trading volume.

The prediction that private information increases trade and liquidity contrasts with traditional intuition. To understand this contrast, I consider a general model that adds non-informational liquidity trading, modeled with stochastic endowments, to the baseline model. When some trading is uninformed, increasing private information can destroy liquidity by increasing the share of informed trade relative to uninformed trade. Private information increases a trade's price impact by increasing the probability that the trade is informed. In the baseline model, all trade is informed so this channel is inoperative. More generally, the probability that a trade is informed is insensitive to the level of private information whenever uninformed trade is very large or very small relative to informed trade. In these settings, traditional intuition fails, and private information increases liquidity as in

the baseline model.

The next section reviews related literature. Section 3 introduces, solves, and derives comparative statics for the baseline model. Section 4 presents the general model, including comparative statics and numerical examples of how overconfidence and information structure affect trading and liquidity. Section 5 presents empirical analysis supporting the stylized facts introduced above. Section 6 assesses the model's empirical predictions. Section 7 concludes.

2 Literature review

Disagreement

Disagreement models posit that a combination of private information (or private interpretation of public information) and behavioral biases causes otherwise rational investors to disagree about asset values. Disagreement models are motivated by two shortcomings of standard asset pricing models with homogenous beliefs. First, standard models (at least in their simplest forms) are at odds with well-established asset pricing anomalies. Second, standard models have no role for trading and thus cannot explain the high turnover observed in many financial markets. Disagreement is an intuitively appealing rationale for trade and, depending on what drives the disagreement, also has the potential to explain pricing anomalies.

Hong and Stein (2007) provide a nice summary and taxonomy of disagreement models. The first ingredient for these models is some manner of private information. One possibility is that private information comes from gradual information flow as in the gradual dissemination of information to newswatchers over time in Hong and Stein (1999). Another possibility is that investors have limited attention and thus process only a subset of available information, possibly for entirely rational reasons related to the cost of attention (e.g., Peng and Xiong, 2006). It is also possible that investors see the same information but interpret it differently, possibly due to heterogeneous priors (e.g., Harris and Raviv, 1993; Kandel and Pearson, 1995).

The second ingredient for significant disagreement trading is a behavioral bias in information processing. Private information alone does not generate disagreement in standard models. Instead, market prices aggregate and fully reveal information (Grossman, 1976), which causes investors to agree about asset values and eliminates motives for trade (Milgrom and Stokey, 1982).

Introducing random asset supply or exogenous liquidity trading (e.g., Hellwig, 1980; Grossman and Stiglitz, 1980; Diamond and Verrecchia, 1981; Kyle, 1985) mitigates information revelation and creates some disagreement, but trading is still largely pinned down by assumptions about noise trading. For example, in Kyle’s (1985) model exogenous noise trading represents half of total order flow variance, and in a moderate variance calibration of Diamond and Verrecchia’s (1981) model, I find that liquidity trading is 95% of total trading volume. To yield more significant disagreement trading, investors must value their own information more highly than information extracted from market prices. Overconfidence is a convenient modeling device for achieving this result and is supported by substantial psychological evidence (see Odean, 1998; Daniel, Hirshleifer, and Subrahmanyam, 1998; and DeBondt and Thaler, 1995, for good discussions of overconfidence).

Scheinkman and Xiong (2003) model overconfidence arising from investors overestimating the precision of their own signals relative to the signals of other investors and show that it leads to disagreement, high trading volume, and speculative bubbles. Banerjee and Kremer (2010) model disagreement coming from different interpretations of a public signal, and Banerjee (2011) models disagreement from underestimating the precision of other agents’ signals. Eyster, Rabin, and Vayanos (2018) consider disagreement trading in a cursed equilibrium in which some traders entirely neglect the informational content of prices.⁴ Boot, Gopalan, and Thakor (2008) model disagreement arising from heterogeneous priors and show that disagreement about whether to undertake a potential project creates a governance challenge for public firms that may push them to go private. Kelley and Tetlock (2013) estimate a structural model of trading and conclude that overconfident disagreement is central to explaining trading volume.

Dynamic models of overconfidence posit that investors learn to be overconfident based on past experience. In particular, investors are subject to a self-attribution bias that causes them to overestimate how much their own skill was responsible for past successes. As a result, overconfidence is highest following positive returns. Daniel, Hirshleifer, and Subrahmanyam (1998) and Gervais and Odean (2001) model this phenomenon. Because investors hold the market in aggregate, self-attribution bias predicts that overconfidence should be high following high market returns. Stock-level overconfidence should be high following high individual stock returns because owners of the

⁴Eyster, Rabin, and Vayanos (2018) also consider an overconfidence model that is similar to the model proposed in this paper, but their focus is on trading, not liquidity.

stock just experienced high returns. Statman, Thorley, and Vorkink (2006) show that turnover is higher than normal following high market returns and high individual stock returns, consistent with high returns increasing overconfidence, which in turn increases trading intensity. Griffin, Nardari, and Stulz (2007) find similar evidence across countries.

Liquidity has not been a major focus of the disagreement literature and is not explicitly discussed in most disagreement papers. Exceptions include Odean (1998), Kyle, Obizhaeva, and Wang (2017), and Baker and Stein (2004). Odean (1998) adds overconfidence to the Kyle (1985) model of liquidity. When the informed insider receives a noisy signal and overestimates the precision of that signal, the insider’s overconfidence increases liquidity. Specifically, the market maker’s price function becomes less sensitive to order flow as the insider becomes more overconfident. Though they don’t explicitly focus on liquidity, the duopoly model of Kyle and Wang (1997) produces a similar result with two overconfident insiders. Kyle, Obizhaeva, and Wang (2017) show that overconfidence mitigates price impact and generates disagreement trading.⁵ Baker and Stein (2004) also model overconfidence as increasing liquidity. However, this relationship is by assumption based on the logic that the same behavioral biases that cause investors to overestimate the precision of their own signals will also cause them to underestimate the informational content of market prices. None of these papers address how overconfident disagreement affects the impact of private information on trading and liquidity.

Liquidity

The market microstructure literature aims to understand what causes illiquidity in financial markets. The backdrop is that in Walrasian equilibrium, beliefs are independent from prices and market prices perfectly reflect the demands of all agents. In real financial markets investors learn from prices, and prices can deviate from fundamental values creating costs to trading. Trading costs (illiquidity) come from two main sources. First, not all agents are active in markets at the same time. Thus, transactions must be facilitated by market makers. These market makers

⁵Kyle, Obizhaeva, and Wang’s (2017) model is similar to the model proposed in this paper. The main difference between their single period model and the baseline model in this paper is that investors in the Kyle, Obizhaeva, and Wang model have market power, whereas all agents in this paper are price takers. Despite the modeling similarities, Kyle, Obizhaeva, and Wang address different questions than this paper. Their focus is on how market power and overconfidence combine to give large investors an incentive to smooth trading whereas this paper analyzes how overconfident disagreement changes the impact of private information on trading and liquidity.

must cover whatever costs they incur by being constantly active in financial markets, and they must be compensated for the risk they take by holding assets while searching for counterparties. Market makers might also extract profits from strategic behavior. The second source of illiquidity is private information. Traders may worry that they are being exploited by counterparties with different information and therefore demand different prices. Biais, Glosten, and Spatt (2005) survey the microstructure literature. Vayanos and Wang (2012) propose a unified model encompassing multiple sources of illiquidity. Brunnermeier and Pedersen (2009) explore the relation between market liquidity (the ease with which an asset can be traded) and funding liquidity (the ease with which an asset can be used as collateral). Osambela (2015) studies the relation between funding liquidity and market liquidity in the presence of differences of opinion.

This paper focuses on information illiquidity because information is intimately tied to disagreement. The classic models of information liquidity are Kyle (1985) and Glosten and Milgrom (1985). In both models, market makers see order flow and are unsure whether the order flow comes from an informed insider or an uninformed liquidity trader. Market makers rationally infer the probability that order flow reflects information and adjust prices accordingly. Kyle (1985) describes this process in terms of the impact of order flow on price. Glosten and Milgrom (1985) highlight that private information naturally leads to bid-ask spreads. More recent microstructure research extends these frameworks to consider strategic behavior by market makers and market design issues. The microstructure liquidity literature generally does not model overconfidence. Odean (1998) and Kyle, Obizhaeva, and Wang (2017), discussed above, are notable exceptions.

The microstructure literature typically considers illiquidity in settings where market makers set prices and bid-ask spreads. However, market makers are not necessary for the concept of information illiquidity. For example, one can think about illiquidity in the rational expectations framework of Grossman (1976). All agents receive private signals, observe market prices, and form trading demands. Grossman's conclusion is that prices fully reveal the average signal. As a result, there is no disagreement and no trading. In effect, markets are infinitely illiquid. Hellwig (1980), Diamond and Verrecchia (1981), and Grossman and Stiglitz (1980) introduce uncertain asset supply so that prices are no longer fully revealing. As a result, agents disagree about asset prices and trade with one another. Prices finitely react to order flow. None of these models explicitly considers liquidity, but information illiquidity is just as present in them as in the market maker

microstructure models.

3 Baseline model

Setup

Agents have fixed endowments of a risky asset and receive private signals about the asset's value. The agents trade the asset in a public market in which all agents see the market-clearing price. Informed by their private signals and the observed market price, agents form beliefs about the asset's payoff and decide how much of it to buy or sell. The only departure from full rationality is that agents are overconfident about the precision of their own signals. All agents are identical other than their realizations of the private signal. Thus, the same agents simultaneously act as informed traders, behavioral traders, and market makers. The model is an adaptation of Grossman's (1976) rational expectations model in which the one significant change is that agents are overconfident about the precision of their signals. As in Grossman's model, the lack of supply variance or any other noise results in prices that are fully revealing. Nonetheless, overconfidence induces disagreement and trade.

The next section adds stochastic endowments. This adds liquidity trading to the model and makes prices only partially revealing. The baseline model described here is a limiting case of the more general model. I develop the baseline model separately both for expositional simplicity and because the baseline model corresponds to a setting in which trade is entirely generated by informed disagreement. Since the goal is to understand how liquidity interacts with disagreement trading, this is a natural place to start.

Assumptions

There are two assets, a risk-free asset in elastic supply that yields 1 unit of consumption and a risky asset in fixed supply that yields θ units of consumption and has a price of P (determined in equilibrium), where the risk-free asset is the numeraire. All agents share a common prior that θ is normally distributed with a mean of μ and precision of τ_p (i.e., $\theta \sim \mathcal{N}(\mu, \frac{1}{\tau_p})$). In addition to the common prior, there are private signals with precision τ_s about the asset's value. The signals

are $y_m = \theta + \varepsilon_m$, where $\varepsilon_m \stackrel{iid}{\sim} \mathcal{N}\left(0, \frac{1}{\tau_s}\right)$. τ_p and τ_s are both positive and finite. μ is finite, and it is most natural to think of it as positive.

The economy has N agents divided into M equal groups, each of size $\frac{N}{M}$. Agent i in group $m(i)$ sees signal $y_{m(i)}$. All agents have a known and certain endowment of $\frac{\mu_X}{N}$ units of the risky asset. Thus, each group has an aggregate endowment of $\frac{\mu_X}{M}$ and total asset supply is μ_X . M is greater than one and finite. More than one group is required to create disagreement and trade. A finite number of groups is required to prevent θ from being perfectly revealed by the aggregate of the signals. I consider the limiting case in which $N \rightarrow \infty$ to ensure that individual agents have negligible impact on the price of the risky asset.⁶ This limiting case is equivalent to a continuum of agents divided into M groups of equal mass.

All agents have constant absolute risk aversion preferences and risk tolerance of $\frac{\eta}{N}$ (i.e., $U_i(c_i) = -\exp\left\{-\frac{N}{\eta}c_i\right\}$). Thus, aggregate risk tolerance is η . Agents are overconfident about the precision of their own signals. They believe ε_m are independent and $\varepsilon_m \sim_i \mathcal{N}\left(0, \frac{1}{\psi_{i,m}\tau_s}\right)$, where $\psi_{i,m} = \psi$ if agent i is a member of group m and $\psi_{i,m} = 1$ otherwise. Unless otherwise noted, I assume ψ is greater than one and finite (i.e., agents are finitely overconfident). Agents know all model parameters including the overconfidence of other agents. Agents observe price but do not observe the signals of other agents.

Equilibrium

I consider an equilibrium in which the asset's price is a linear function of the average private signal, i.e.,

$$P = \alpha + \beta \bar{y} \tag{1}$$

where $\bar{y}$ is the average of the M private signals.

Because price is a 1:1 function of the average private signal, all agents effectively see the average signal, from which they can extract the average signal of agents in groups other than their own, $\bar{y}_{-m(i)}$. Specifically, $\bar{y}_{-m(i)} = \frac{M}{\beta(M-1)}(P - \alpha) - \frac{1}{M-1}y_{m(i)}$. Using Bayesian updating, agent i 's

⁶This is in contrast to the traditional market microstructure literature (e.g., Kyle, 1985), which models insiders as risk-neutral monopolists. Risk aversion and monopolistic behavior both have the effect of limiting asset demands. Incorporating both would unnecessarily complicate the model. Adding market power (as in Kyle, Obizhaeva, and Wang (2017)) increases the amount of overconfidence necessary to induce investors to trade, decreases trading volumes, and increases liquidity but does not alter the three propositions derived from the baseline model.

posterior beliefs as a function of $y_{m(i)}$ and P are:

$$E_i[\theta|y_{m(i)}, P] = \frac{\tau_p \mu + (\psi - 1) \tau_s y_{m(i)} + \frac{M \tau_s}{\beta} (P - \alpha)}{\tau_p + (M + \psi - 1) \tau_s} \quad (2a)$$

$$\text{Var}_i[\theta|y_{m(i)}, P] = (\tau_p + (M + \psi - 1) \tau_s)^{-1} \quad (2b)$$

Given his CARA utility, agent i 's asset demand is:

$$\begin{aligned} D_i &= \frac{E_i[\theta|y_{m(i)}, P] - P}{\frac{N}{\eta} \text{Var}_i[\theta|y_{m(i)}, P]} \\ &= \frac{\eta}{N} \left(\tau_p \mu - \frac{M \tau_s}{\beta} \alpha + (\psi - 1) \tau_s y_{m(i)} - \left(\tau_p + \left(M + \psi - 1 - \frac{M}{\beta} \right) \tau_s \right) P \right) \end{aligned} \quad (3)$$

The market clearing price must solve $\mu_X = \sum_i D_i$, which results in a price that is a linear function of $\bar{y}$. Equating its coefficients with the coefficients of (1) yields:

$$P = \frac{\tau_p \mu + (M + \psi - 1) \tau_s \bar{y}}{\tau_p + (M + \psi - 1) \tau_s} - \frac{\mu_X}{\eta (\tau_p + (M + \psi - 1) \tau_s)} \quad (4)$$

This result should not be surprising. $\frac{\tau_p \mu + (M + \psi - 1) \tau_s \bar{y}}{\tau_p + (M + \psi - 1) \tau_s}$ is the average posterior expectation of the agents and $\frac{\mu_X}{\eta (\tau_p + (M + \psi - 1) \tau_s)}$ is the risk premium required to hold asset supply μ_X with posterior variance $(\tau_p + (M + \psi - 1) \tau_s)^{-1}$ and aggregate risk tolerance η . Compared to the price that would prevail without overconfidence ($\psi = 1$), equation (4) shows that overconfidence biases price toward the private signals and decreases the required risk premium. For ψ that is small relative to M , both of these effects are modest.

Trading and liquidity

Each agent's directed trading volume is his net asset demand, $D_i - \frac{\mu_X}{N}$. Aggregate trading is $Vol = \frac{1}{2} \sum_m |Trade_m|$, where $Trade_m$ is the net asset demand of group m :

$$\begin{aligned} Trade_m &= \sum_{i:m(i)=m} \left[D_i - \frac{\mu_X}{N} \right] \\ &= \frac{\eta}{M} (\psi - 1) \tau_s (y_m - \bar{y}) \end{aligned} \quad (5)$$

$|Trade_m|$ is a folded mean-zero random variable so its expectation is $E|Trade_m| = \sqrt{\frac{2\text{Var}(Trade_m)}{\pi}}$.

Expected aggregate trading is:

$$\begin{aligned} E[Vol] &= \frac{1}{2} \sum_m E|Trade_m| \\ &= \frac{\eta(\psi - 1)}{2} \sqrt{\frac{2(1 - \frac{1}{M})\tau_s}{\pi}} \end{aligned} \quad (6)$$

From equation (6) it is clear that trading increases with ψ and τ_s and is unaffected by τ_p .

In the baseline model, there is no liquidity trading, but the concept of liquidity is still operative. Analogous to Kyle's (1985) lambda, I define illiquidity as the price impact of trade resulting from an exogenous shock. Intuitively, I am interested in how price would respond to an exogenous buy or sell trade. However, there is no exogenous trading in this model. Rather, trade is an endogenous response to underlying shocks received by agents. Illiquidity is the ratio of a shock's price impact to its impact on the shocked agents' trades (i.e., their net asset demand). Because agents only interact through their trades, this ratio exactly represents the trades' price impact. A shock affects an entire group of agents so the relevant net asset demand is the group's, which is given by $Trade_m$.⁷ Formally, illiquidity is:

$$\lambda \equiv \frac{\frac{dP}{dy_m}}{\frac{dTrade_m}{dy_m}} \quad (7)$$

Taking derivatives of $Trade_m$ (eq. 5) and P (eq. 4) with respect to y_m yields:

$$\begin{aligned} \lambda &= \frac{M(M + \psi - 1)}{\eta(M - 1)(\psi - 1)(\tau_p + (M + \psi - 1)\tau_s)} \\ &= \left\{ \frac{M}{(M - 1)\eta(\tau_p + (M + \psi - 1)\tau_s)} \right\} \\ &\quad + \left\{ \left[\frac{\tau_s}{\tau_p + (M + \psi - 1)\tau_s} \right] \left[\frac{\eta}{M} \left(\psi - \frac{(M + \psi - 1)}{M} \right) \tau_s \right]^{-1} \right\} \\ &= \{S\} + \{B\} \end{aligned} \quad (8)$$

⁷In the baseline model, the only shock is to a group's signal so this is the shock I use to define illiquidity. More generally, one could also consider the impact of group endowment shocks or even shocks to individual agents. In all cases, the resulting illiquidity is the same because in an anonymous market any trade must have the same price impact regardless of its source.

Equation (8) shows that a buy trade affects price in two ways. First, it decreases the net asset supply that must be held by the non-shocked agents. The supply effect (S) is the decline in the risk premium required by the non-shocked agents.⁸ Second, the trade increases non-shocked agents' posterior expectations. B captures this belief channel. The trade's impact on posterior expectations is the informational value of a y_m shock, $\left[\frac{\tau_s}{\tau_p + (M + \psi - 1)\tau_s} \right]$, divided by how aggressively the shocked agents trade on the shock, $\left[\frac{\eta}{M} \left(\psi - \frac{(M + \psi - 1)}{M} \right) \tau_s \right]$.⁹

Note that trading aggression is proportional to private signal precision and private signal value is the ratio of private signal precision to total precision. As a result, illiquidity depends only on aggregate information $(\tau_p + (M + \psi - 1)\tau_s)$, not its component parts. Public and private information both enhance liquidity. By contrast, overconfidence more than proportionally increases trading aggression without having much impact on signal value or aggregate information. Thus, overconfidence enhances liquidity, primarily through the belief channel.

The above equations and discussion support three propositions about how overconfidence and information structure affect trading and liquidity.

PROPOSITION 1. As overconfidence increases (ψ increases), trading and liquidity both increase.

PROPOSITION 2. As private information becomes more precise (τ_s increases), trading and liquidity both increase.

PROPOSITION 3. As public uncertainty increases (τ_p decreases), trading is unaffected and liquidity decreases.

Of the three propositions, proposition 2 is probably most surprising because it contrasts with traditional intuition that private information destroys liquidity and causes markets to break down. The general model develops this contrast in more detail and shows that private information enhances liquidity whenever the mix of informed and uninformed trade is insensitive to increases in private information. This is clearly the case in the baseline model because it only includes informed trade.

Propositions 2 and 3 concern different aspects of the information information structure that are often difficult disentangle empirically. Information asymmetry and heterogeneity can be high

⁸The non-shocked agents have posterior variance $(\tau_p + (M + \psi - 1)\tau_s)^{-1}$ and risk tolerance $\frac{(M-1)\eta}{M}$, resulting in a risk premium (price reduction) of $Supply \cdot \frac{M}{(M-1)\eta(\tau_p + (M + \psi - 1)\tau_s)}$.

⁹Taking the derivative of equation (5) with respect to y_m establishes that $\frac{dTrade_m}{dy_m} = \frac{\eta}{M} \left(\psi - \frac{(M + \psi - 1)}{M} \right) \tau_s$.

because private signals are precise or because there is little public information. In either case, private information is more valuable and beliefs rely more heavily on private signals. In combination, propositions 2 and 3 imply that as private information increases relative to public information, trading increases or remains constant with an ambiguous impact on liquidity.

4 General model

Setup and assumptions

The general model is identical to the baseline model except that endowments are uncertain and unknown to agents in other groups. Specifically, agent i in group m (i) is endowed with $\frac{M}{N}x_{m(i)}$ units of the risky asset, where $x_m \stackrel{iid}{\sim} \mathcal{N}\left(\frac{\mu_X}{M}, \frac{V}{M}\right)$ is the total endowment of group m . The resulting total asset supply is $X = \sum_m x_m \sim \mathcal{N}(\mu_X, V)$. Agents not in group m know the distribution of x_m , but do not observe x_m .

The model is an adaptation of Diamond and Verrecchia's (1981) noisy rational expectations model in which the one significant change is that agents are overconfident about the precision of their signals. The main distinguishing characteristic of Diamond and Verrecchia (1981) compared to other noisy rational expectations models (e.g., Hellwig, 1980) is that Diamond and Verrecchia consider endowment shocks to modeled agents whereas Hellwig and others consider direct, unobserved shocks to aggregate asset supply. The Diamond and Verrecchia approach is more natural because it makes liquidity trading an endogenous response to underlying endowment shocks instead of modeling liquidity trades as direct exogenous shocks to external asset supply. Diamond and Verrecchia's approach is also analytically convenient because it yields closed form solutions.

Results from the model are presented and discussed below. Derivations are in the Internet Appendix.

Equilibrium

I assume that price is a linear function of the average private signal and aggregate asset supply:

$$P = \alpha + \beta \bar{y} - \gamma (X - \mu_X) \quad (9)$$

Given the price function described by equation (9), agent i extracts a noisy signal ($A_{m(i)}$) for the average private signal of other groups ($\bar{y}_{-m(i)}$) from observing price, $y_{m(i)}$, and $x_{m(i)}$:

$$\begin{aligned} A_{m(i)} &= \frac{M}{\beta(M-1)}(P - \alpha) - \frac{1}{M-1}y_{m(i)} + \frac{\gamma M}{\beta(M-1)}\left(x_{m(i)} - \frac{\mu_X}{M}\right) \\ &= \bar{y}_{-m(i)} - \frac{\gamma M}{\beta}\left(\bar{x}_{-m(i)} - \frac{\mu_X}{M}\right) \end{aligned} \quad (10)$$

Note that $A_{m(i)}$ is independent of $y_{m(i)}$ and $x_{m(i)}$ and $A_{m(i)} \sim_i \mathcal{N}\left(\theta, \frac{1}{\tau_A}\right)$, where τ_A is the precision agent i attributes to $A_{m(i)}$:

$$\tau_A = \left(\frac{1}{(M-1)\tau_s} + \left(\frac{\gamma}{\beta}\right)^2 \frac{M}{M-1}V \right)^{-1} \quad (11)$$

Agent i forms posterior beliefs about the asset's payoff (θ) using Bayesian updating with signals $y_{m(i)}$ and $A_{m(i)}$. All agents use their posterior beliefs to determine their asset demands. Setting total asset demand equal to total asset supply results in a market-clearing price that is a linear function of the average private signal and aggregate asset supply just as assumed in equation (9). The resulting fixed point problem has the unique solution:

$$\alpha = \frac{(\eta^2\psi^2\tau_p\tau_s + MV\tau_p)\mu - \left(\eta\psi^2\tau_s + \frac{MV}{\eta}\right)\mu_X}{\eta^2\psi^2\tau_s(\tau_p + (M + \psi - 1)\tau_s) + (\tau_p + \psi\tau_s)MV} \quad (12a)$$

$$\beta = \frac{\psi\tau_s(\eta^2\psi(M + \psi - 1)\tau_s + MV)}{\eta^2\psi^2\tau_s(\tau_p + (M + \psi - 1)\tau_s) + (\tau_p + \psi\tau_s)MV} \quad (12b)$$

$$\gamma = \frac{\eta^2\psi(M + \psi - 1)\tau_s + MV}{\eta(\eta^2\psi^2\tau_s(\tau_p + (M + \psi - 1)\tau_s) + (\tau_p + \psi\tau_s)MV)} \quad (12c)$$

For many applications, the ratio of γ to β is an important quantity. I define this ratio as:

$$\Gamma \equiv \frac{\gamma}{\beta} = \frac{1}{\eta\psi\tau_s} \quad (13)$$

To see that the baseline model is a limiting case of the general model, note that as $V \rightarrow 0$, $X \xrightarrow{p} \mu_X$ and $P \xrightarrow{p} \frac{\tau_p\mu + (M + \psi - 1)\tau_s\bar{y}}{\tau_p + (M + \psi - 1)\tau_s} - \frac{\mu_X}{\eta(\tau_p + (M + \psi - 1)\tau_s)}$, the baseline model price.

Price informativeness

In the baseline model, price fully revealed all relevant information about the asset's value. When endowments are stochastic, this is no longer the case. How much less informative is price?

One measure of price informativeness, used by Diamond and Verrecchia (1981), is the posterior precision an agent achieves relative to what his posterior precision would be if he observed all signals. Under full information, each agent has a posterior precision of $\tau_p + (M + \psi - 1)\tau_s$.¹⁰ This is the same posterior precision achieved in the baseline model. In the general model, each agent knows the public prior and observes $y_{m(i)}$ and $A_{m(i)}$, resulting in posterior precision:

$$\tau_p + \psi\tau_s + \tau_A = \tau_p + \psi\tau_s + \left(\frac{1}{(M-1)\tau_s} + \Gamma^2 \frac{M}{M-1} V \right)^{-1} \quad (14)$$

More recently, Bai, Philippon, and Savov (2016) propose measuring price informativeness from the econometrician's point of view. Specifically, they measure price informativeness as the R^2 of a regression of price on future asset value. In the general model, $P = \alpha + \beta\bar{y} - \gamma(X - \mu_X)$ has a variance of $\beta^2 \left(\frac{1}{\tau_p} + \frac{1}{M\tau_s} \right) + \gamma^2 V$; asset value (θ) has a variance of $\frac{1}{\tau_p}$; and the covariance of price with asset value is $\frac{\beta}{\tau_p}$, resulting in:

$$R^2 = \frac{\tau_p^{-1}}{\tau_p^{-1} + (M\tau_s)^{-1} + \Gamma^2 V} \quad (15)$$

By contrast, if agents saw all private signals (as they effectively do in the baseline model), the R^2 of this regression would be $\frac{\tau_p^{-1}}{\tau_p^{-1} + (M\tau_s)^{-1}}$.

Under both measures of price informativeness, deviations from baseline full revelation price informativeness are functions of $\Gamma^2 V = \frac{V}{\eta^2 \psi^2 \tau_s^2}$. As noise (V) increases, price informativeness decreases. As risk tolerance (η), overconfidence (ψ), and private information (τ_s) grow, prices become more informative. Roughly speaking, price informativeness is determined by the relative levels of informed and liquidity trading. Liquidity trading is increasing in V . Informed trading is increasing in η , ψ , and τ_s .

¹⁰The public prior has precision τ_p ; an agent's own signal has subjective precision $\psi\tau_s$; and the $M-1$ signals of agents in other groups each have precision τ_s . Because the signals and prior are all independent, their precisions are additive.

Trading

As in the baseline model, group m 's directed trading volume is its net asset demand. The difference is that group m 's trading now depends on two random shocks instead of just one. Specifically, $Trade_m = \left\{ \frac{\eta}{M} \left(\psi \tau_s - \frac{\tau_A}{M-1} \right) (y_m - \bar{y}) \right\} - \left\{ \left(1 - \frac{\eta \gamma \tau_A}{(M-1)\beta} \right) (x_m - \bar{x}) \right\}$, where the $(y_m - \bar{y})$ term represents informed trading and the $(x_m - \bar{x})$ term represents liquidity trading. Informed trading is greater than in the baseline model.¹¹ Agents now have two motives for informed trade. First, overconfidence ($\psi > 1$) causes them to overweight their own signals as in the baseline model. Second, price no longer fully reveals the average signal, giving agents another reason to trade on their own signal. Liquidity trading is less than the endowment shocks themselves because endowment shocks are partially offset by demand changes. Agents realize that endowment shocks affect price and take this into account when determining their asset demands.¹² As a result, liquidity, asset riskiness, and risk tolerance all influence liquidity trading.

Expected trading volume increases in the variance of group m 's trading:¹³

$$\begin{aligned}
\text{Var}[Trade_m] &= \{\text{Var}[Informed_Trade_m]\} + \{\text{Var}[Liquidity_Trade_m]\} \\
&= \left\{ \frac{\eta^2 (M-1) \psi^2 \tau_s (\eta^2 (\psi-1) \psi \tau_s + MV)^2}{M^3 (\eta^2 \psi^2 \tau_s + MV)^2} \right\} \\
&\quad + \left\{ \frac{(M-1) (\eta^2 (\psi-1) \psi \tau_s + MV)^2 V}{M^2 (\eta^2 \psi^2 \tau_s + MV)^2} \right\} \\
&= \frac{(M-1) (\eta^2 (\psi-1) \psi \tau_s + MV)^2}{M^3 (\eta^2 \psi^2 \tau_s + MV)} \tag{16}
\end{aligned}$$

From equation (16), it is clear that public uncertainty (τ_p^{-1}) has no impact on trading volume, just as in the baseline model. Taking derivatives with respect to ψ , one can also see that overconfidence increases overall and informed trading (again consistent with the baseline model). Overconfidence can initially decrease liquidity trading if V is large, but as ψ increases it eventually decreases

¹¹To see that informed trade is greater than in the baseline model, note that $\tau_A < (M-1)\tau_s$. Thus, $\frac{\eta}{M} \left(\psi \tau_s - \frac{\tau_A}{M-1} \right) < \frac{\eta}{M} (\psi-1) \tau_s$, the trading coefficient on $(y_m - \bar{y})$ in the baseline model.

¹²Agents are atomistic and do not have any price impact by themselves (thus no monopoly pricing motive is present), but endowment shocks are shared by a positive mass of agents. The group's endowment shock drives a wedge between price and group's posterior value, which agents in the group exploit by changing asset demand in the opposite direction of the endowment shock.

¹³Specifically, expected volume is $E[Vol] = \frac{1}{2} \sum_m E[Trade_m]$ and $E[Trade_m] = \sqrt{\frac{2\text{Var}(Trade_m)}{\pi}}$ just as in the baseline model.

liquidity trading as well.¹⁴ These effects are small. In practice, overconfidence has very little impact on liquidity trading.

The relationship between private information (τ_s) and trade is more complicated. Liquidity trading always decreases with τ_s . When overconfidence is high ($\psi > \frac{9}{8}$ for informed trade and $\psi > 2$ for overall trade), informed and overall trading increase monotonically with τ_s . When overconfidence is moderate ($1 < \psi < \frac{9}{8}$ for informed trade and $1 < \psi < 2$ for overall trade), informed and overall trading increase with τ_s if τ_s is large relative to V .¹⁵ Without overconfidence, all types of trading decrease with τ_s . Essentially, private information always induces trade when there is high overconfidence. Under moderate overconfidence, private information decreases trade in liquidity trading (high V) environments but increases trade in disagreement trading (low V) environments. The Akerlof (1970) logic that private information destroys trade applies when agents are fully rational and in liquidity trading environments when agents are only moderately overconfident. When trade primarily stems from overconfident disagreement, the opposite effect prevails, and private information increases trade.

Liquidity

Liquidity has the same definition as in the baseline model, and illiquidity is the price impact of a trade resulting from an exogenous shock, formally measured as:

$$\lambda \equiv \frac{\frac{dP}{dy_m}}{\frac{dTrade_m}{dy_m}}$$

Illiquidity could just as easily be defined using endowment (x_m) shocks instead of information (y_m) shocks. The resulting λ is the same. In either case, the shocked agents interact with the rest of the market only through their trading demand. Thus, a given shock to trading must have the same price impact regardless of what motivated the trade. This can be verified algebraically by taking derivatives with respect to x_m instead of y_m .

¹⁴Specifically, $\frac{dVar[Liquidity_Trade_m]}{d\psi} < 0$ if $1 < \psi < \frac{1}{\eta} \sqrt{\frac{MV}{\tau_s}}$ and is positive for larger ψ .

¹⁵ Assuming $\psi > 1$, $\frac{dVar[Informed_Trade_m]}{d\tau_s} < 0$ if and only if $\psi < \frac{9}{8}$ and $\tau_s < \frac{MV(3-2\psi-\sqrt{9-8\psi})}{2\eta^2\psi^2(\psi-1)}$; and $\frac{dVar[Trade_m]}{d\tau_s} < 0$ if and only if $\psi < 2$ and $\tau_s < \frac{(2-\psi)MV}{\eta^2\psi^2(\psi-1)}$.

Taking derivatives of price and net asset demand with respect to underlying shocks yields:

$$\begin{aligned}
\lambda &= \left\{ \frac{M}{(M-1)\eta(\tau_p + \psi\tau_s + \tau_A)} \right\} \\
&\quad + \left\{ \frac{\eta^2\psi^2\tau_s}{\eta^2\psi^2\tau_s + MV} \right\} \left\{ \frac{\tau_s}{\tau_p + \psi\tau_s + \tau_A} \right\} \left\{ \frac{\eta}{M} \left(\psi\tau_s - \frac{\beta}{M}(\tau_p + \psi\tau_s + \tau_A) \right) \right\}^{-1} \\
&= \{S\} + \{B1\}\{B2\}\{B3\}^{-1}
\end{aligned} \tag{17}$$

Equation (17) expresses lambda as the sum of a supply channel and a belief channel. The belief channel is further decomposed into the probability that a trade is informed ($B1$) times the impact a known shock to y_m would have on the posteriors of other agents ($B2$) divided by how aggressively shocked agents trade on y_m shocks ($B3$). S , $B2$, and $B3$ were present in the baseline model. As before, S and $\frac{B2}{B3}$ tend to decrease as total information ($\tau_p + \psi\tau_s + \tau_A$) increases regardless of whether the information is public or private. S and $\frac{B2}{B3}$ also decrease with overconfidence.

$B1$ is new and deserves special consideration. First, note that $B1 = \frac{\eta^2\psi^2\tau_s}{\eta^2\psi^2\tau_s + MV}$ is the probability that a trade is informed. Specifically, $B1$ is the ratio of informed trading variance to total trading variance (see equation (16) to verify this). $B1$ is increasing in ψ and τ_s and represents the channel through which they can destroy liquidity. Private information and overconfidence increase informed trading as a share of overall trading. When V is close to $\frac{\eta^2\psi^2\tau_s}{M}$, this ratio is highly sensitive to ψ and τ_s . By contrast, for very small or large V , $B1$ is close to 1 or 0 and fairly stable. The traditional logic that private information destroys liquidity applies only when increases in private information have a large impact on the ratio of informed to total trade. In particular, the traditional logic does not apply to disagreement trading environments because in those environments trade is primarily informed regardless of the exact level of private information. Similarly, overconfidence and private information enhance liquidity in extreme liquidity trading environments. The Internet Appendix considers derivatives of λ with respect to ψ , τ_s , and τ_p and derives parameter regions in which the derivatives are positive and negative. The Internet Appendix also shows that private information can enhance liquidity even without overconfidence. The basic result hinges on how much trading is informed, not on the presence of overconfidence.

Numerical examples

To better understand how overconfidence and information structure affect trading and liquidity, it is useful to consider numerical examples. Consider the following baseline parameter values:

$$\mu = 1$$

$$\mu_X = 1$$

$$M = 10$$

$$\tau_p = 100$$

$$\tau_s = 10$$

$$\psi = 2$$

$$\eta = 0.1$$

$$V \in \{0.0000001, 0.1, 10\}$$

μ and μ_X are normalized to one, and there are 10 groups of agents. Prior precision of 100 yields a public prior standard deviation of 10%. Private precision of 10 makes the private signals in aggregate as valuable as the public prior. Overconfidence of 2 means that agents attribute twice as much value to their own signals as they do to the signals of agents in other groups. Aggregate risk tolerance of 0.1 produces a risk premium of 10% under public information without supply shocks.

The low variance environment ($V = 0.0000001$) roughly corresponds to the baseline constant endowment model. At baseline values, expected turnover is 12% and $\lambda = 58\%$ (meaning an exogenous trade of 1% of aggregate asset supply would change the asset's price by 58 basis points). The first row of Figure 1 plots these quantities as functions of overconfidence, varying ψ from 1 to 10 while holding all other parameters at baseline values. Panel A shows that trading is approximately zero when $\psi = 1$ and trading increases close to proportionally with ψ . When $\psi = 2$, expected turnover is 12%. When $\psi = 10$, expected turnover is 108%. The solid line in Panel B plots λ as a function of ψ . As predicted by the baseline model, the market is highly illiquid when $\psi = 1$ (approaching infinity as $V \rightarrow 0$) and becomes more liquid as ψ increases. When $\psi = 2$, $\lambda = 58\%$. When $\psi = 10$, $\lambda = 8\%$. Beyond $\psi = 10$, λ continues to decrease with ψ , approaching 0 as $\psi \rightarrow \infty$. The dotted and dashed lines in panel B decompose λ into its supply and belief channels. The belief channel (dashed line) is the dominant source of illiquidity. The second row of Figure 1 repeats

the same exercise, varying private signal precision from 0 to 100. As private signals become more precise, trading increases (panel C) and illiquidity decreases (panel D). The final row of Figure 1 considers public prior precision. As predicted, turnover is unaffected and illiquidity decreases as τ_p increases.

Figure 2 plots expected turnover and illiquidity as functions of overconfidence, private information precision, and public prior precision in a moderate supply variance environment. Baseline parameter values are the same as before except that endowment variance is now 0.1, which corresponds to an aggregate asset supply standard deviation of 0.32, compared to its mean of 1. These examples capture a market in which trading comes from both informational and liquidity motives. Panels A, C, and E plot expected total turnover (solid line), expected informed turnover (dashed line), and expected liquidity turnover (dotted line). Informed turnover is the turnover that would prevail if agents differed only in their information shocks (i.e., if $x_m = \bar{x} \forall m$). Analogously, liquidity turnover is the turnover that would prevail if agents differed only in their supply shocks. Note that total turnover is less than the sum of informed turnover and liquidity turnover because these two types of trade partially offset one another.

The $\psi = 1$ starting point of panel A plots turnover in the absence of overconfidence, which recreates the Diamond and Verrecchia (1981) model. As alluded to earlier in the paper, liquidity turnover represents 95% of total turnover. As overconfidence (panel A) and private information precision (panel C) increase, informed turnover increases, driving up total turnover. By contrast, private information decreases liquidity turnover as τ_s increases. Liquidity turnover is fairly insensitive to overconfidence. At first it slightly decreases with ψ , then it slightly increases with ψ . Both forms of turnover are unaffected by public prior precision.

As expected, liquidity trading enhances liquidity. The baseline λ decreases from 58% to 20% when V increases from 0.0000001 to 0.1. The reduction in λ is entirely driven by the belief channel. Belief and supply illiquidity now have similar magnitudes. Most interestingly, liquidity trading changes the relationships between ψ , τ_s , and illiquidity. Illiquidity is now a hump-shaped function of ψ (panel B) and τ_s (panel D). Overconfidence and private information precision initially decrease liquidity before eventually enhancing it. Panels A and C show why. Initial increases in ψ and τ_s dramatically increase the ratio of informed trading to total trading, which increases the probability that any given trade is informed. At higher levels of ψ and τ_s , this ratio and probability are fairly

stable.

Figure 3 replicates the moderate variance example under a no-overconfidence ($\psi = 1$) baseline. Panel A shows that without overconfidence private information decreases turnover. This is an illustration of the general result derived above. Overconfidence is a necessary ingredient for private information to increase trading. Nonetheless, panel B shows that private information enhances liquidity for τ_s above 32. Overconfidence amplifies this liquidity enhancement but is not necessary for the basic result.

Figure 4 plots expected turnover and illiquidity in a high supply variance ($V = 10$) environment in which trading is primarily liquidity-driven. In this environment, liquidity is enhanced (the baseline λ is 9.4%, and this is almost entirely from the supply channel). λ decreases in ψ , τ_s , and τ_p (see panels B, D, and F) because of the supply channel. Panels A, C, and E show that expected total turnover is high (its baseline value is 378%) and insensitive to ψ , τ_s , and τ_p even though ψ and τ_s increase informed trading.

5 Stylized facts

How does the model relate to the data? To start with, I propose three stylized facts about stock, bond, and option markets:

1. Trade and liquidity are positively correlated;
2. As private information increases relative to public information, trade increases and liquidity decreases, trade increases and liquidity decreases; and
3. Trade and liquidity are elevated following high past returns.

These facts are not new, especially with respect to the stock market. For example, Sadka and Scherbina (2007) show that analyst forecast dispersion (one proxy for the relative importance of private information) decreases liquidity. Frazzini and Lamont (2007) show that turnover is elevated around earnings announcements (another proxy for private information). Evidence of decreased liquidity around earnings announcements includes Lee, Mucklow, and Ready (1993), Krinsky and Lee (1996), So and Wang (2014), and Johnson and So (2018). Statman, Thorley, and Vorkink

(2006) and Griffin, Nardari, and Stulz (2007) show that trading increases following high returns. Hong and Stein (2007) also observe that returns and trading volume are correlated.

I extend these findings and show that trading increases with analyst dispersion and liquidity increases with past returns. I also show that the stylized facts hold across stock, corporate bond, and option asset classes. Finally, the Internet Appendix shows that results are robust across different types of stocks by sorting stocks on size, book-to-market ratios, past returns, and other characteristics.

The sources for stock data are CRSP for return and volume data, Compustat for industry and earnings announcement data, I/B/E/S for analyst earnings forecast data, and TAQ for intraday trade and quote data. The stock sample is New York Stock Exchange (NYSE) stocks with prices above \$5 at the end of the previous month. The sample starts in 1926, but most of the analysis is limited by analyst earnings forecast data, which start in 1976, and bid-ask spread data, which start in 1993 for the primary bid-ask measure based on intraday TAQ data. Stock turnover is calculated as monthly volume divided by shares outstanding. The primary liquidity measure is the effective bid-ask spread (*ebidask*) of Chordia, Roll, and Subrahmanyam (2000), which is calculated at the transaction level as twice the difference between a trade’s price and the midpoint of the prevailing quote before the trade. Because this measure is only available starting in 1993, Amihud’s (2002) $illiq_{it} = \frac{|Return_{it}|}{\$Volume_{it}}$ measure is also used for some analyses. The Internet Appendix considers two alternative bid-ask spread measures.

Bond data come from TRACE, supplemented by Mergent FISD bond characteristics. The TRACE data start in 2002. The sample is investment grade U.S. corporate bonds and medium term notes without asset backing or security enhancement features that are at least one year old and have at least one year of maturity left. The sample also requires that a bond be actively traded (defined as having at least two buy trades and two sell trades) on at least 15 days during the previous month. I consider only transactions between dealers and their customers. Turnover is the total par value of all trades in a bond scaled by the bond’s outstanding par value. Effective bid-ask spread (*ebidask*) is the difference between the weighted average prices of a day’s buy and sell trades.¹⁶

¹⁶The TRACE data I use is the enhanced dataset available on WRDS, which identifies whether a trade is with a customer or another dealer and which side of the transaction the reporting dealer was on. The enhanced dataset also includes all volume data instead of truncating large trades.

Option data comes from Ivy DB OptionMetrics, available starting in 1996. Analysis is at the stock (as opposed to option contract) level. Option turnover as total option dollar volume divided by stock market capitalization. Quoted end-of-day proportional bid-ask spreads are weighted by dollar volume when averaged to the stock level.

Figure 5 plots equally-weighted average monthly turnover and bid-ask spreads for stocks, bonds, and options. Stock turnover (panel A) averages 14% during the plotted 1993 to 2011 time period and reaches as high as 40% late in the sample. Stock effective bid-ask spreads (panel B) average 0.4% and decrease over the sample period. Corporate bonds also have significant trading activity and moderate bid-ask spreads. Bond turnover (panel C) is typically near the 5-10% range, and average bond effective bid-ask spreads (panel D) range from 0.5% to 2.5%. In contrast to stocks and bonds, dollar transaction volumes are relatively small for stock options. Average dollar option volume (panel E) is typically under 1% of stock market capitalization. Average option bid-ask spreads hover around 20% of contract value. However, these spreads are relative to option contract value as opposed to stock price exposure. In practice, options often deliver large stock price exposure with minimal up-front contract value.

The remainder of this section summarizes findings most relevant for each of the three stylized facts. Details and additional stock analysis are in the Internet Appendix.

Fact 1: Trade and liquidity are positively correlated

Traditional reasoning predicts a strong positive relationship between trading and liquidity. The two quantities are mutually reinforcing. More noise trading improves liquidity, and enhanced liquidity attracts additional trading activity. Admati and Pfleiderer (1988) show that when timing is discretionary, liquidity trading concentrates in periods of high trading and enhanced liquidity. Collin-Dufresne and Fos (2016) show that informed traders also time their trades to preference periods of elevated noise trading and enhanced liquidity.

Positive correlation between trade and liquidity is clear in the stock, bond, and option data. As turnover increases, bid-ask spreads tend to decrease. Table 1 reports results for panel regressions of log bid-ask spreads on log turnover. The regressions include stock, bond, and time fixed effects. Because turnover and bid-ask spreads are both expressed as logs, the results can be interpreted as elasticities. The elasticity of stock bid-ask spreads with respect to turnover is -15.8%. The equivalent

coefficients for bonds and options are -12.8% and -21.3%, respectively. All three estimates are highly significant. The Internet Appendix includes cross-sectional tests and alternative measures of illiquidity.

Fact 2: As private information increases relative to public information, trade increases and liquidity decreases

While trading and liquidity are positively correlated, they do not always move in the same direction. In particular, as private information increases relative to public information, trading tends to increase, but liquidity tends to decrease. The prevalence of private information relative to public information is frequently described as asymmetric information in the empirical literature, consistent with models of informed and uninformed investors such as Kyle (1985), which are prevalent in the market microstructure literature. However, private information need not be asymmetric. Building on Grossman (1976) and Diamond and Verrecchia (1981), in the models presented in this paper, private information creates information heterogeneity among informed investors as opposed to information asymmetry between informed and uninformed investors. Information asymmetry and heterogeneity both stem from the prevalence of private information relative to public information, which is what the empirical proxies discussed below capture.

Changes in private information relative to public information are identified in two ways. First, periods around earnings announcements are likely to have elevated private information. Prior to announcements, private information can be in the form of leaks and insider trading. After announcements, investors process different pieces of information at different paces using different models, keeping private information high until the announcement is fully digested and reflected in prices. Public uncertainty is also high around earnings announcements because asset values are highly sensitive to the announcements. Second, I follow Sadka and Scherbina (2007) and use dispersion of analyst forecasts as a proxy for private information relative to public information. Analyst dispersion may represent or cause public uncertainty and could also stem from private information. Dispersion is measured as the standard deviation across analysts of current year earnings forecasts scaled by the mean forecast. Firms are included if they are covered by at least two analysts, have a non-zero mean earnings forecast, and have a December fiscal year end. The December fiscal year requirement ensures that all stocks have the same amount of time remaining in the current fiscal

year.

For earnings announcement analysis, turnover and bid-ask spreads are scaled by average values over the three calendar months prior to an earnings announcement and analyzed over a 21-day trading window around earnings announcements. Figure 6 plots equally weighted average scaled turnover and bid-ask spreads in event time around earnings announcements. Day 0 is the announcement day or first trading day after the announcement. Other days represent trading days relative to the announcement. 95% confidence intervals are plotted in dashed lines.

Panels A and B plot stock data. Consistent with Frazzini and Lamont (2007), turnover starts to increase the day before an announcement, spikes to 80% above normal levels on the announcement day, stays at that level for another day, and then decays. Bid-ask spreads widen around earnings announcements with a peak of 13% above normal levels on the day of the announcement, and are also elevated the day before and after the announcement. The Internet Appendix (Figure IA.3) shows that the same pattern emerges for alternative bid-ask spread measures. Moreover, the results are robust across different types of stocks. In sorts on size, book-to-market ratios, past returns, and institutional ownership, the same basic patterns emerge for all groups (Figures IA.4 to IA.8).

Though not always as pronounced, turnover and bid-ask spreads also tend to increase around earnings announcements for bonds and options. Panel C of Figure 6 shows that bond turnover peaks at 40% above normal levels the day after an earnings announcement. Bond bid-ask spreads (panel D) are slightly elevated around earnings announcements, particularly the day before an announcement. Option volumes (panel E) surge to over 250% normal levels around earnings announcements, and option bid-ask spreads (panel F) widen to 6% above normal levels, with a peak the day before the earnings announcement.

Analyst earnings forecast dispersion is a second proxy for private information relative to public information. Sadka and Scherbina (2007) show that in sorts on analyst dispersion, high dispersion stocks tend to be less liquid. I employ panel regressions to control for stock, bond, and time fixed effects and add a test of the effect of analyst dispersion on trading volumes. I also extend the analysis to bonds and options. Table 2 presents the results. The analyzed variables are logs so coefficients can be interpreted as elasticities. For stocks, the turnover coefficient on lagged analyst dispersion (column 1) is 1.6% and the bid-ask spread coefficient on lagged analyst dispersion (column 2) is 11.0%. Bonds and options respond similarly to analyst dispersion. For bonds, the turnover coeffi-

cient (column 3) is 7.1%, and the bid-ask spread coefficient (column 4) is 5.6%. For options, the turnover coefficient (column 5) is 14.4%, and the bid-ask spread coefficient (column 6) is 3.6%. All coefficients are highly significant. The Internet Appendix repeats this analysis for additional measures of stock illiquidity, performs cross-sectional analysis on stocks, and confirms that baseline results are robust across most sorts on stock characteristics.

Fact 3: Trade and liquidity are elevated following high past returns

Statman, Thorley, and Vorkink (2006) use a market vector autoregression (VAR) and individual stock-level VARs to show that market turnover increases following high market returns and individual stock turnover increases following both high market returns and high individual stock returns. I add illiquidity to Statman, Thorley, and Vorkink’s market VAR methodology and apply it to stocks, bonds, and options. The market VAR model is:

$$Y_t = \alpha + \sum_{k=1}^2 A_k Y_{t-k} + e_t \quad (18)$$

where Y_t is a 3×1 vector of detrended log turnover, detrended log illiquidity, and excess market returns.¹⁷ Using two lags is optimal according to the Bayesian information criteria. In all cases, the market return variable is excess stock market returns. Thus, the VAR estimates the impact of stock market returns on future trading and liquidity in stocks, bonds, and options. The baseline stock analysis uses Amihud’s *illiq* instead of bid-ask spread because it is available for the full sample instead of just after 1993. Bonds and option analysis uses the same bid-ask spread measures as before.

Figure 7 plots impulse response functions for stock, bond, and option market VARs. The plots show how a one standard deviation stock market return shock affects forecasted turnover and illiquidity.¹⁸ Consistent with Statman, Thorley, and Vorkink (2006), panel A shows that stock

¹⁷Turnover and illiquidity measures are detrended using a Hodrick and Prescott (1997) filter. Following Statman, Thorley, and Vorkink (2006) and common practice, I use a penalty value of 14,400 for the filter. Also following Statman, Thorley, and Vorkink (2006), I employ a 2-sided filter. The 2-sided filter, which makes use of future data, would be problematic if I was using it for forecasting purposes, but I am not. To verify that my results are unaffected by the use of future data, in untabulated analysis, I replicate my market VAR with a 1-sided HP filter proposed by Stock and Watson (1999) with unchanged results.

¹⁸Other impulse-response combinations are omitted for brevity. Coefficient estimates and plots of all stock impulse response functions are included in the internet appendix (Table IA.11 and Figure IA.9).

turnover responds positively to market returns. A one standard deviation return shock predicts a turnover increase of 5% in the next month. Turnover increases slightly in the following month and then decays toward normal levels. Panel B shows that positive market returns also predict enhanced liquidity. A one standard deviation return shock predicts decreases in *illiq* of 8% in the next month. Positive stock return shocks also predict decreased bond and option bid-ask spreads (panels D and F). Bond and option volumes (panels C and E) do not significantly respond to past stock returns. The Internet Appendix reports coefficients for the stock VAR (Table IA.11) and results for stock VARs using bid-ask spread measures instead of *illiq* (Figure IA.10).

In addition to responding to market returns, individual stock turnover and liquidity respond positively to past stock and industry level returns. Using separately estimated stock-level VARs, Statman, Thorley, and Vorkink (2006) show that stock turnover is positively related to both past market returns and past individual stock returns. I employ a different econometric strategy and estimate a single panel VAR that includes stock-level turnover, illiquidity, and returns as well as industry returns. The panel VAR specification allows for stock and time fixed effects, eliminating the need for detrending the data. Specifically, the panel VAR model is:

$$Y_{i,t} = \alpha_t + f_i + \sum_{k=1}^2 A_k Y_{i,t-k} + e_t \quad (19)$$

where $Y_{i,t}$ is a 4×1 vector of log stock turnover, log stock *illiq*, stock returns, and industry returns for stock i in month t . α_t and f_i are 4×1 vectors of time and stock fixed effects for each variable. The model employs two lags for consistency with the market model. The time fixed effects control for the effect of market returns as well as any other market-level time variation. Prior to estimation, all variables are de-meaned to remove time fixed effects. Stock fixed effects are trickier because stock demeaned lag variables are not orthogonal to the regression residual. Similarly, directly estimating stock fixed effects would produce biased and non-consistent estimates for all coefficients. Following Holtz-Eakin, Newey, and Rosen (1988), I take the first differences of all variables, resulting in:

$$Y_{i,t} - Y_{i,t-1} = (\alpha_t - \alpha_{t-1}) + \sum_{k=1}^2 A_k (Y_{i,t-k} - Y_{i,t-1-k}) + e_t \quad (20)$$

which can be estimated using $Y_{i,t-2}$ and $Y_{i,t-3}$ as instruments. The sample include all observations

with at least three lagged observations. When there are breaks in the data, observations before and after the break are treated as separate stocks. Lagged variables control for autocorrelation in the data, but there is likely cross-sectional correlation within time periods. To account for this, standard errors are bootstrapped with randomly sampled (with replacement) time periods. When a time period is drawn, all observations in that time period are included. This preserves the data's cross-sectional correlation structure.

Figure 8 plots the most relevant impulse response functions of the panel VAR. As before, the impulse shocks are all one standard deviation. In panel A, turnover is unaffected by individual stock returns. In the other three panels, past returns forecast increased turnover and decreased illiquidity. A one standard deviation shock to an individual stock's return forecasts a future decline in *illiq* of 2.7% (panel B). A one standard deviation shock to an industry's return forecasts a 0.7% increase in turnover (panel C) and a 0.4% decrease in *illiq* (panel D). Coefficient estimates are reported in the Internet Appendix (Table IA.12), which also includes related cross-sectional momentum analysis (Table IA.13).

Past returns and overconfidence

The past returns evidence is consistent with high past returns increasing overconfidence, which in turn increases trading and liquidity. This connection between past returns and overconfidence is predicted by the learning models of Daniel, Hirshleifer, and Subrahmanyam (1998) and Gervais and Odean (2001). Self-attribution bias causes investors to be particularly overconfident following high returns. Because investors hold the market on average, aggregate overconfidence should track market returns. Similarly, individual stock returns could affect stock-level overconfidence. To the extent that investors specialize in certain industries or have industry-specific confidence levels, past industry returns also affect stock-level overconfidence.

To test the overconfidence interpretation, I analyze investor-level returns and trading activity. Self-attribution bias predicts that an investor's overconfidence will increase following positive returns to his own portfolio. This hypothesis can be tested using account-level trading data from a discount brokerage firm (this is the Barber and Odean, 2000, data). Specifically, I look at how trading intensity responds to market returns, individual stock returns, and an investor's own portfolio returns. Using six months of trailing trade data, an investor's portfolio is estimated to be the

net positions his trades over that period would result in. All short positions are set to zero because shorting is uncommon for retail investors. The outcome variable of interest is whether buying intensity, measured as the total dollar value of all buy trades, increases between month t and month $t+2$. The returns considered are excess returns over the risk free rate in the interim month.

Table 3 reports the results. Column (1) regresses increases in overall buying intensity at the investor-month level on market and portfolio returns with account fixed effects. Unconditionally, buying intensity increases 15% of the time. 1% shocks to portfolio and market returns increase this probability by 30.4 and 33.4 basis points, respectively. Column (2) regresses increases in stock specific buying intensity at the investor-month-stock level on market, portfolio, and individual stock returns. Individual stock buying intensity increases 4% of the time, and 1% shocks to all three return variables increase this probability by 5 to 10 basis points. All coefficient estimates are statistically significant. The results are consistent with past returns increasing investor confidence, causing investors to trade more aggressively.

6 Model assessment

In contrast to traditional models, disagreement trading generates significant trading and liquidity in the baseline model even without exogenous noise. This resolves the puzzle that private information is frequently associated elevated trading is a potential explanation for the enhanced liquidity Collin-Dufresne and Fos (2015) and Kacpercyk and Pagnotta (2018) observe around informed trading. Moreover, disagreement trading is consistent with the stylized facts developed in Section 2.

First, high turnover is generally associated with high liquidity. This is easiest to see by comparing the low, medium, and high variance numerical examples. As liquidity trading increases across the scenarios, trading and liquidity increase dramatically. Overconfidence also typically moves trading and liquidity in the same direction.

Second, the disagreement model is consistent with the prevalence of private information relative to public information increasing trade while decreasing liquidity. Proposition (2) of the baseline model predicts that private information increases trade and liquidity. Proposition (3) predicts that public uncertainty decreases liquidity while having no impact on trading. Jointly, propositions (2)

and (3) predict that information heterogeneity (which consists of private information and public uncertainty) can only increase trading and has an ambiguous effect on liquidity. Unlike traditional models, these predictions are consistent with observed empirical evidence.

Third, overconfidence increases trade and liquidity. If overconfidence increases following high past returns, as self-attribution bias theory predicts, this delivers the prediction that turnover and liquidity will increase following high past returns.

In addition to conforming with the stylized facts, the model generates new predictions about how the impact of information structure on trading and liquidity varies with the type of information shock, the level of liquidity trading, and the existing level of private information. Testing these predictions is challenging because they involve unobservable quantities and it is difficult to separately identify private information, public uncertainty, and liquidity trading. Fortunately, the model itself provides guidance for differentiating private information changes from public uncertainty changes. In the baseline model, private information increases trading, whereas public uncertainty has no impact on it. Thus, information changes with a larger trading impact are more likely to be driven by changes in private information. The baseline model predicts that private information enhances liquidity whereas public uncertainty reduces it. Thus, we should expect information heterogeneity to have the least negative impact on liquidity (and potentially even enhance it) when it has the most impact on trading.

This prediction can be tested in the data by sorting stocks based on their past turnover responses to changes in information heterogeneity. Past turnover response is based on rolling 5-year regressions of stock-level turnover on lagged analyst forecast dispersion, controlling for aggregate dispersion. Stocks are annually sorted into low, medium, and high responsiveness groups based on 30th and 70th percentile breakpoints. I then estimate analyst forecast dispersion panel regressions with interactions between analyst forecast dispersion and past turnover responsiveness groups.

Table 4 reports the results. Column (1) shows that turnover responsiveness to analyst forecast dispersion is persistent. Low past turnover responsiveness stocks have a slightly negative response to analyst forecast dispersion (the coefficient of log turnover on log lagged analyst dispersion is -0.9% for the low responsiveness group). As past turnover responsiveness increases, this coefficient increases by 2.4 ppt for the medium responsiveness group and by 5.8 ppt for the high responsiveness group. Both results are highly significant. The model predicts that as turnover responsiveness

increases, illiquidity responsiveness should decrease. This is consistent with what I find in the data. Column (2) presents results for *illiq*. High turnover response stocks have a highly significant 4.1 ppt lower illiquidity response to analyst dispersion compared to low turnover response stocks. Medium turnover response stocks have about the same dispersion coefficient as low turnover response stocks. Columns (3) to (5) present results for different bid-ask spread measures with inconclusive results. While signs are the same, coefficients are smaller, and results are insignificant.

7 Conclusion

Liquidity plays an increasingly important role in asset pricing and macro finance. Yet, we lack a clear understanding of some of the most basic drivers of liquidity in informationally sensitive markets. Existing models and intuition suggest that private information destroys trading and liquidity. Though less well understood, overconfidence is generally associated with enhanced liquidity. Most liquidity models involve exogenous noise trading and do not explain the positive correlation between private information and trading volume in the data.

The disagreement literature proposes that overconfidence-driven disagreement provides a rationale for trade that does not require exogenous noise traders or uncertain asset supply. This paper shows that overconfidence is also sufficient for generating and thinking about liquidity. In the baseline model, agents differentially weight their own signals even though prices perfectly reveal the average signal. This causes them to disagree about the asset's value and trade. The market is liquid even without liquidity trading. All of this is within an intentionally simple market setup. All agents are homogenous until receiving signals. No outside parties are needed for noise trading or market making. The same agents simultaneously serve as informed traders, noise traders, and market makers.

The baseline model rationalizes heavy trading and liquidity with private information and no exogenous liquidity trading. The model also produces three predictions: (1) as overconfidence increases, trading and liquidity both increase; (2) as private information becomes more precise, trading and liquidity both increase; and (3) as public uncertainty increases, trading is unaffected and liquidity decreases. Consistent with stylized facts about stocks, corporate bonds, and stock options, these predictions jointly imply that information heterogeneity increases trading and to the

extent that past returns increase overconfidence they also increase trading and enhance liquidity. The predictions also imply that information heterogeneity changes with the largest trading impact should have the smallest (and potentially even negative) illiquidity impact, which is consistent with the data.

The baseline model's predictions are at odds with some of the recent literature on financial crises. Kacperczyk and Schnabl (2010) and Gorton and Metrick (2011) document the collapse of trade in asset backed commercial paper and repurchase agreements during the 2007-2009 financial crisis. Both of these markets previously facilitated liquidity trading in instruments that were perceived to be safe and informationally insensitive. The authors reasonably argue that the market collapses were at least in part driven by increases in asymmetric information. Dang, Gorton, and Holmstrom (2015) propose a model of liquidity in which debt contracts optimally facilitate trade in part by minimizing asymmetric information. Consistent with traditional intuition, if the debt contracts become informationally sensitive trade and liquidity dry up. This narrative of the financial crisis contradicts the baseline model and is also difficult to reconcile with the liquid trade observed in equity markets despite significant private information. The general model provides a way to bridge this gap. In the face of moderate liquidity trading, adding private information to a market at first destroys liquidity by increasing the likelihood any given trade is informed. Once most trades are already informed, further increases in private information enhance liquidity.

One aspect of disagreement trading that this paper doesn't address is welfare. Overconfidence-driven disagreement trading clearly has some negative consequences. Unequal risk sharing causes optimistic agents to hold higher variance portfolios than they would without overconfidence, which diminishes welfare under the criterion of Brunnermeier, Simsek, and Xiong (2014). On the other hand, overconfidence facilitates liquidity, which is likely beneficial. Overconfidence also makes prices more informative by causing agents to trade on their information more aggressively. Additional work connecting the microfoundations of liquidity and trading to their welfare implications is necessary to fully understand these trade-offs.

References

- Admati, Anat and Paul Pfleiderer. 1988. "A Theory of Intraday Patterns: Volume and Price Variability." *Review of Financial Studies* 1 (1): 3–40.
- Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84 (3): 488–500.
- Amihud, Yakov. 2002. "Illiquidity and Stock Returns: Cross-section and Time-series Effects." *Journal of Financial Markets* 5 (1): 31–56.
- Bai, Jennie, Thomas Philippon, and Alexi Savov. 2016. "Have Financial Markets Become More Informative?" *Journal of Financial Economics* 122 (3): 625–654.
- Baker, Malcolm, and Jeremy C Stein. 2004. "Market Liquidity as a Sentiment Indicator." *Journal of Financial Markets* 7 (3): 271–299.
- Banerjee, Snehal. 2011. "Learning from Prices and the Dispersion in Beliefs." *Review of Financial Studies* 24 (9): 3025–3068.
- Banerjee, Snehal, and Ilan Kremer. 2010. "Disagreement and Learning: Dynamic Patterns of Trade." *Journal of Finance* 65 (4): 1269–1302.
- Barber, Brad M, and Terrance Odean. 2000. "Trading Is Hazardous to Your Wealth: The Common Stock Investment Performance of Individual Investors." *The Journal of Finance* 55 (2): 773–806.
- Biais, Bruno, Larry Glosten, and Chester Spatt. 2005. "Market Microstructure: A Survey of Microfoundations, Empirical Results, and Policy Implications." *Journal of Financial Markets* 8 (2): 217–264.
- Boot, Arnoud, Radhakrishnan Gopalan, and Anjan Thakor. 2008. "Market Liquidity, Investor Participation, and Managerial Autonomy: Why Do Firms Go Private?" *Journal of Finance* 63 (4): 2013–2059.
- Brunnermeier, Markus K., Alp Simsek, and Wei Xiong. 2014. "A Welfare Criterion for Models with Distorted Beliefs." *Quarterly Journal of Economics* 129 (4): 1753–1797.

- Chordia, Tarun, Richard Roll, and Avanidhar Subrahmanyam. 2000. "Commonality in Liquidity." *Journal of Financial Economics* 56 (1): 3–28.
- Collin-Dufresne, Pierre and Vyacheslav Fos. 2015. "Do Prices Reveal the Presence of Informed Trading?" *Journal of Finance* 70 (4): 1555–1582.
- Collin-Dufresne, Pierre and Vyacheslav Fos. 2016. "Insider Trading, Stochastic Liquidity, and Equilibrium Prices." *Econometrica* 84 (4): 1441–1475.
- Dang, Tri Vi, Gary Gorton, and Bengt Holmström. 2015. "Ignorance, Debt, and Financial Crises." Working Paper.
- Daniel, Kent, David Hirshleifer, and Avanidhar Subrahmanyam. 1998. "Investor Psychology and Security Market Under- and Overreactions." *The Journal of Finance* 53 (6): 1839–1885.
- DeBondt, Werner F.M., and Richard H. Thaler. 1995. "Financial Decision Making in Markets and Firms: A Behavioral Perspective." In *Finance*, edited by R. Jarrow, V. Maksimovic, and W.T. Ziemba. Elsevier - North Holland: Series of Handbooks in Operations Research and Management Sciences.
- Diamond, Douglas W, and Robert E Verrecchia. 1981. "Information Aggregation in a Noisy Rational Expectations Economy." *Journal of Financial Economics* 9 (3): 221–235.
- Eyster, Erik, Matthew Rabin, and Dimitri Vayanos. 2018. "Financial Markets where Traders Neglect the Informational Content of Prices." *Journal of Finance*, forthcoming.
- Frazzini, Andrea, and Owen Lamont. 2007. "The Earnings Announcement Premium and Trading Volume." Working Paper.
- Gervais, Simon, and Terrance Odean. 2001. "Learning to Be Overconfident." *Review of Financial Studies* 14 (1): 1–27.
- Glosten, Lawrence R., and Paul R. Milgrom. 1985. "Bid, Ask and Transaction Prices in a Specialist Market with Heterogeneously Informed Traders." *Journal of Financial Economics* 14 (1): 71–100.
- Gorton, Gary, and Andrew Metrick. 2012. "Securitized Banking and the Run on Repo." *Journal of Financial Economics* 104 (3): 425–451.

- Griffin, John M., Federico Nardari, and Rene M. Stulz. 2007. "Do Investors Trade More When Stocks Have Performed Well? Evidence from 46 Countries." *The Review of Financial Studies* 20 (3): 905-951.
- Grossman, Sanford. 1976. "On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information." *The Journal of Finance* 31 (2): 573-585.
- Grossman, Sanford J, and Joseph E Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *The American Economic Review* 70 (3): 393-408.
- Harris, Milton, and Artur Raviv. 1993. "Differences of Opinion Make a Horse Race." *Review of Financial Studies* 6 (3): 473-506.
- Hellwig, Martin F. 1980. "On the Aggregation of Information in Competitive Markets." *Journal of Economic Theory* 22 (3): 477-498.
- Hodrick, Robert J, and Edward C Prescott. 1997. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit and Banking* 29 (1): 1-16.
- Holmström, Bengt. 2008. "Discussion of 'The Panic of 2007,' by Gary Gorton." Working Paper.
- Holtz-Eakin, Douglas, Whitney Newey, and Harvey S Rosen. 1988. "Estimating Vector Autoregressions with Panel Data." *Econometrica* 56 (6): 1371-1395.
- Hong, Harrison, and Jeremy C Stein. 1999. "A Unified Theory of Underreaction, Momentum Trading, and Overreaction in Asset Markets." *The Journal of Finance* 54 (6): 2143-2184.
- . 2007. "Disagreement and the Stock Market." *Journal of Economic Perspectives* 21 (2): 109-128.
- Johnson, Travis and Eric So. 2018. "Asymmetric Trading Costs Prior to Earnings Announcements: Implications for Price Discovery and Returns." *Journal of Accounting Research* 56 (1): 217-263.
- Kacperczyk, Marcin, and Emiliano Pagnotta. 2018. "Chasing Private Information." Working Paper.
- Kacperczyk, Marcin, and Philipp Schnabl. 2010. "When Safe Proved Risky: Commercial Paper During the Financial Crisis of 2007-2009." *Journal of Economic Perspectives* 24 (1): 29-50.

- Kandel, Eugene, and Neil D Pearson. 1995. "Differential Interpretation of Public Signals and Trade in Speculative Markets." *Journal of Political Economy* 103 (4): 831–872.
- Kelley, Eric, and Paul Tetlock. 2013. "Why Do Investors Trade?" Working Paper.
- Krinsky, Itzhak and Jason Lee. 1996. "Earnings Announcements and the Components of the Bid-Ask Spread." *Journal of Finance* 51 (4): 1523–1535.
- Kyle, Albert S. 1985. "Continuous Auctions and Insider Trading." *Econometrica: Journal of the Econometric Society* 53 (6): 1315–1335.
- Kyle, Albert S, Anna A Obizhaeva, and Yajun Wang. 2017. "Smooth Trading with Overconfidence and Market Power." *Review of Economic Studies* 85 (1): 611–662.
- Kyle, Albert S, and F Albert Wang. 1997. "Speculation Duopoly with Agreement to Disagree: Can Overconfidence Survive the Market Test?" *The Journal of Finance* 52 (5): 2073–2090.
- Lee, Charles, Belinda Mucklow, and Mark Ready. 1993. "Spreads, Depths, and the Impact of Earnings Information: An Intraday Analysis." *Review of Financial Studies* 6 (2): 345–374.
- Milgrom, Paul, and Nancy Stokey. 1982. "Information, Trade and Common Knowledge." *Journal of Economic Theory* 26 (1): 17–27.
- Odean, Terrance. 1998. "Volume , Volatility , Price , and Profit When All Traders Are Above Average." *The Journal of Finance* 53 (6): 1887–1934.
- Peng, Lin, and Wei Xiong. 2006. "Investor Attention, Overconfidence and Category Learning." *Journal of Financial Economics* 80 (3): 563–602.
- Sadka, Ronnie, and Anna Scherbina. 2007. "Analyst Disagreement, Mispricing, and Liquidity." *The Journal of Finance* 62 (5): 2367–2403.
- Scheinkman, Jose A., and Wei Xiong. 2003. "Overconfidence and Speculative Bubbles." *Journal of Political Economy* 111 (6): 1183–1220.
- So, Eric and Sean Wang. 2014. "News-driven Return Reversals: Liquidity Provision Ahead of Earnings Announcements." *Journal of Financial Economics* 114: 20–35.
- Statman, Meir, Steven Thorley, and Keith Vorkink. 2006. "Investor Overconfidence and Trading Volume." *Review of Financial Studies* 19 (4): 1531–1565.

- Stock, James H, and Mark W Watson. 1999. “Forecasting Inflation.” *Journal of Monetary Economics* 44 (2): 293–335.
- Vayanos, Dimitri, and Jiang Wang. 2012. “Liquidity and Asset Prices under Asymmetric Information and Imperfect Competition.” *Review of Financial Studies* 25 (5): 1339–1365.

Figure 1: Trading and Illiquidity Under Low Variance. Expected turnover and illiquidity (λ) are calculated using values indicated on the horizontal axis for the parameter in parentheses and low variance baseline values ($\mu = 1$, $\mu_x = 1$, $M = 10$, $\tau_p = 100$, $\tau_s = 10$, $\psi = 2$, $\eta = 0.1$, $V = 0.0000001$) for all other parameters. Vertical lines represent the baseline. Expected total turnover is a solid line; expected informed turnover is a dashed line; and expected liquidity turnover is a dotted line. Overall illiquidity is a solid line; the illiquidity belief channel is a dashed line; and the illiquidity supply channel is a dotted line.

Figure 2: Trading and Illiquidity Under Moderate Variance. Expected turnover and illiquidity (λ) are calculated using values indicated on the horizontal axis for the parameter in parentheses and moderate variance baseline values ($\mu = 1$, $\mu_x = 1$, $M = 10$, $\tau_p = 100$, $\tau_s = 10$, $\psi = 2$, $\eta = 0.1$, $V = 0.1$) for all other parameters. Vertical lines represent the baseline. Expected total turnover is a solid line; expected informed turnover is a dashed line; and expected liquidity turnover is a dotted line. Overall illiquidity is a solid line; the illiquidity belief channel is a dashed line; and the illiquidity supply channel is a dotted line.

Figure 3: Trading and Illiquidity Without Overconfidence. Expected turnover and illiquidity (λ) are calculated using values indicated on the horizontal axis for the parameter in parentheses and moderate variance baseline values without overconfidence ($\mu = 1$, $\mu_x = 1$, $M = 10$, $\tau_p = 100$, $\tau_s = 10$, $\psi = 1$, $\eta = 0.1$, $V = 0.1$) for all other parameters. Vertical lines represent the baseline. Expected total turnover is a solid line; expected informed turnover is a dashed line; and expected liquidity turnover is a dotted line. Overall illiquidity is a solid line; the illiquidity belief channel is a dashed line; and the illiquidity supply channel is a dotted line.

Figure 4: Trading and Illiquidity Under High Variance. Expected turnover and illiquidity (λ) are calculated using values indicated on the horizontal axis for the parameter in parentheses and high variance baseline values ($\mu = 1$, $\mu_x = 1$, $M = 10$, $\tau_p = 100$, $\tau_s = 10$, $\psi = 2$, $\eta = 0.1$, $V = 10$) for all other parameters. Vertical lines represent the baseline. Expected total turnover is a solid line; expected informed turnover is a dashed line; and expected liquidity turnover is a dotted line. Overall illiquidity is a solid line; the illiquidity belief channel is a dashed line; and the illiquidity supply channel is a dotted line.

Figure 5: Monthly Time Series. Stock turnover is relative to stock market capitalization. Bond turnover is relative to outstanding bond value. Option turnover is option contract value relative to stock market capitalization. Bid-ask spreads are all proportional to the value of the security being traded. Stock and bond bid-ask spreads are intraday effective spreads. Option bid-ask spreads are end of day quoted spreads.

Figure 6: Turnover and Liquidity Around Earnings Announcements. Turnover and bid-ask spreads are scaled by average daily values over the three calendar months before the earnings announcement. Solid lines are equally weighted averages. Dashed lines are 95% confidence intervals. Day 0 is the day of the earnings announcement.

Figure 7: Market VAR Responses to Market Return Impulse. Each VAR includes detrended log market turnover ($turn$), CRSP value-weighted market returns in excess of the risk free rate ($rmrf$), and a detrended log measure of market illiquidity ($illiq$ for the stock VAR in the first row, effective bid-ask spread for the bond VAR in the second row, and quoted bid-ask spread for the option VAR in the third row). The solid lines are responses to one standard deviation shocks to $rmrf$ after the number of lags indicated on the horizontal axis. The dashed lines are 95% confidence intervals.

Figure 8: Stock Panel VAR Impulse Response Functions. The panel VAR includes stock log turnover, stock log *illiq* (an illiquidity measure), stock returns, industry returns, and stock and time fixed effects. The first variable in each panel title is the impulse variable. The second variable is the response variable. The solid lines are responses to one standard deviation shocks to the impulse variables after the number of lags indicated on the horizontal axis. The dashed lines are 95% confidence intervals. For brevity, only the most relevant impulse response functions are shown.

Table 1: Turnover Panel Regressions

Results are for stock- and bond-level regressions of log bid-ask spread measures on log turnover. Stock data is for NYSE stocks with lagged prices greater than \$5. Bond data is for actively traded U.S. corporate bonds without credit enhancements. Option data is for all traded stock options. Bid-ask spreads are proportional to security value. Stock and bond bid-ask spreads are intraday effective spreads. Option bid-ask spreads are end of day quoted spreads. Robust clustered (by firm) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	Stocks	Bonds	Options
	(1)	(2)	(3)
	Log Bid-Ask	Log Bid-Ask	Log Bid-Ask
Log Turnover	-0.158*** (0.006)	-0.128*** (0.010)	-0.213*** (0.002)
Stock/Bond FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Observations	331,238	12,831	363,702
R^2	0.889	0.860	0.707
Date Range	1993-2011	2002-2011	1996-2011

Table 2: Analyst Dispersion Panel Regressions

Results are for stock- and bond-level regressions of log turnover and log bid-ask spread measures on lagged (by one month) log dispersion of analyst earnings forecasts. Data is limited to NYSE stocks with lagged prices greater than \$5, at least 2 analyst forecasts, and fiscal years that end in December. Bid-ask spreads are proportional to security value. Stock and bond bid-ask spreads are intraday effective spreads. Option bid-ask spreads are end of day quoted spreads. Robust clustered (by firm) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	Stocks		Bonds		Options	
	(1) Log Turnover	(2) Log Bid-Ask	(3) Log Turnover	(4) Log Bid-Ask	(5) Log Turnover	(6) Log Bid-Ask
Lagged Log Analyst Dispersion	0.016*** (0.004)	0.110*** (0.004)	0.071** (0.028)	0.056*** (0.020)	0.144*** (0.014)	0.036*** (0.005)
Stock/Bond FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	341,431	201,948	4,836	4,824	107,397	107,363
R^2	0.729	0.901	0.506	0.882	0.578	0.636
Date Range	1976-2011	1993-2011	2002-2011	2002-2011	1996-2011	1996-2011

Table 3: Impact of Returns on Buying Intensity

The analyzed data is discount brokerage trades in 100,000 accounts between 1991 and 1996. Dependent variables are indicators for an investor increasing his dollar buying activity over a two-month period. A value of 1 indicates that the dollar sum of all an investor's buy trades in month $t+2$ is greater than the dollar sum of all his buy trades in month t . Column (1) analyzes total buys at the investor-month level. Column (2) analyzes stock-specific buys at the investor-month-stock level. To be included, the investor must have at least two buy trades in month t (overall for column (1) and of the specific stock for column (2)). The explanatory variables represent excess returns over the risk free rate in the interim month ($t+1$). The portfolio return is the return to a portfolio consisting of the investor's cumulative net trades over the six months leading up to month t . Implied short positions are set to zero. Robust clustered (by account) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	(1) Overall Buy Increase	(2) Stock-Specific Buy Increase
Portfolio Return	0.304*** (0.014)	0.079*** (0.017)
Market Return	0.334*** (0.047)	0.101* (0.053)
Stock Return		0.054*** (0.011)
Account FE	Yes	Yes
Observations	180,952	57,415
R^2	0.208	0.172
Date Range	1991-1996	1991-1996

Table 4: Analyst Dispersion Panel Regressions by Lagged Turnover Response

Results are for stock-level regressions of log turnover and log illiquidity measures on lagged (by one month) log dispersion of analyst earnings forecasts. Medium and high turnover responsiveness indicator variables are based on turnover responsiveness to dispersion changes over the past five years relative to 30th and 70th percentile breakpoints. Data is limited to NYSE stocks with lagged prices greater than \$5, at least 2 analyst forecasts, and fiscal years that end in December. Robust clustered (by firm) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	(1) Log Turnover	(2) Log <i>illiq</i>	(3) Log <i>bidask</i>	(4) Log <i>qbidask</i>	(5) Log <i>ebidask</i>
Lagged Log Analyst Dispersion	-0.009* (0.005)	0.215*** (0.010)	0.102*** (0.005)	0.111*** (0.005)	0.112*** (0.005)
Lagged Log Dispersion × Medium Turn Response	0.024*** (0.006)	-0.001 (0.010)	-0.006 (0.005)	-0.005 (0.005)	-0.005 (0.005)
Lagged Log Dispersion × High Turn Response	0.058*** (0.007)	-0.041*** (0.013)	-0.007 (0.007)	-0.010 (0.006)	-0.008 (0.006)
Medium Turn Response	0.032 (0.021)	0.036 (0.036)	0.000 (0.021)	0.010 (0.020)	0.014 (0.021)
High Turn Response	0.191*** (0.026)	-0.124** (0.048)	0.021 (0.026)	0.008 (0.025)	0.022 (0.025)
Stock FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Observations	294,401	294,374	191,101	193,412	193,412
R^2	0.720	0.886	0.921	0.928	0.903
Date Range	1981-2011	1981-2011	1993-2011	1993-2011	1993-2011

Internet Appendix for: “Disagreement and Liquidity”

A Model details

General model solution

I assume that the price function is linear:

$$P = \alpha + \beta \bar{y} - \gamma (X - \mu_X) \quad (\text{IA.1})$$

CLAIM 1. $\beta \neq 0$.

PROOF. Assume $\beta = 0$. Thus, price is uninformative about private signals, and all agents have posterior beliefs of $E_i[\theta|y_{m(i)}, x_{m(i)}, P] = \frac{\tau_p \mu + \psi \tau_s y_{m(i)}}{\tau_p + \psi \tau_s}$ and $\text{Var}_i[\theta|y_{m(i)}, x_{m(i)}, P] = (\tau_p + \psi \tau_s)^{-1}$, resulting in asset demand $D_i[y_{m(i)}, x_{m(i)}, P] = \frac{E_i[\theta|y_{m(i)}, x_{m(i)}, P] - P}{\frac{N}{\eta} \text{Var}_i[\theta|y_{m(i)}, x_{m(i)}, P]} = \frac{\eta}{N} [\tau_p \mu + \psi \tau_s y_{m(i)} - (\tau_p + \psi \tau_s) P]$. Solving $\sum_i D_i = X$ for the market clearing price implies $P = \frac{\tau_p \mu + \psi \tau_s \bar{y} - \frac{X}{\eta}}{\tau_p + \psi \tau_s}$. Thus, $\beta = \frac{\psi \tau_s}{\tau_p + \psi \tau_s} \neq 0$, a contradiction. ■

Given the price function described by equation (IA.1), agent i extracts a noisy signal for $\bar{y}_{-m(i)}$ from observing price, $y_{m(i)}$, and $x_{m(i)}$:

$$\begin{aligned} A_{m(i)} &= \frac{M}{\beta(M-1)} (P - \alpha) - \frac{1}{M-1} y_{m(i)} + \frac{\gamma M}{\beta(M-1)} \left(x_{m(i)} - \frac{\mu_X}{M} \right) \\ &= \bar{y}_{-m(i)} - \frac{\gamma M}{\beta} \left(\bar{x}_{-m(i)} - \frac{\mu_X}{M} \right) \end{aligned} \quad (\text{IA.2})$$

Note that $A_{m(i)} \sim_i \mathcal{N} \left(\theta, \frac{1}{(M-1)\tau_s} + \left(\frac{\gamma}{\beta} \right)^2 \frac{M}{M-1} V \right)$ and $A_{m(i)}$ is independent of $y_{m(i)}$ and $x_{m(i)}$. Using Bayesian updating with signals $y_{m(i)}$ and $A_{m(i)}$ and substituting P , $y_{m(i)}$, $x_{m(i)}$ for $A_{m(i)}$ using (IA.2), agent i 's posterior beliefs as a function of P , $y_{m(i)}$, $x_{m(i)}$ are:

$$E_i[\theta|y_{m(i)}, x_{m(i)}, P] = \frac{\left(\tau_p \mu - \frac{M\tau_A}{\beta(M-1)} \alpha + \left(\psi \tau_s - \frac{\tau_A}{M-1} \right) y_{m(i)} + \frac{\gamma M\tau_A}{\beta(M-1)} \left(x_{m(i)} - \frac{\mu_X}{M} \right) + \frac{M\tau_A}{\beta(M-1)} P \right)}{\tau_p + \psi \tau_s + \tau_A} \quad (\text{IA.3a})$$

$$\text{Var}_i[\theta|y_{m(i)}, x_{m(i)}, P] = (\tau_p + \psi \tau_s + \tau_A)^{-1} \quad (\text{IA.3b})$$

where $\tau_A = \left(\frac{1}{(M-1)\tau_s} + \left(\frac{\gamma}{\beta} \right)^2 \frac{M}{M-1} V \right)^{-1}$ is the precision agent i attributes to $A_{m(i)}$.

Agent i 's asset demand is:

$$\begin{aligned} D_i &= \frac{E_i[\theta|y_{m(i)}, x_{m(i)}, P] - P}{\frac{N}{\eta} \text{Var}_i[\theta|y_{m(i)}, x_{m(i)}, P]} \\ &= \frac{\eta}{N} \left(\begin{aligned} &\tau_p \mu - \frac{M\tau_A}{\beta(M-1)} \alpha + \left(\psi \tau_s - \frac{\tau_A}{M-1} \right) y_{m(i)} + \frac{\gamma M \tau_A}{\beta(M-1)} \left(x_{m(i)} - \frac{\mu_X}{M} \right) \\ &- \left(\tau_p + \psi \tau_s + \left(1 - \frac{M}{\beta(M-1)} \right) \tau_A \right) P \end{aligned} \right) \end{aligned} \quad (\text{IA.4})$$

The market clearing price must solve $X = \sum_i D_i$. Thus,

$$P = \frac{\tau_p \mu - \frac{1}{\eta} \mu_X - \frac{M\tau_A}{\beta(M-1)} \alpha + \left(\psi \tau_s - \frac{\tau_A}{M-1} \right) \bar{y} - \left(\frac{1}{\eta} - \frac{\gamma \tau_A}{\beta(M-1)} \right) (X - \mu_X)}{\tau_p + \psi \tau_s + \left(1 - \frac{M}{\beta(M-1)} \right) \tau_A} \quad (\text{IA.5})$$

Equations (IA.1) and (IA.5) yield the following system of equations:

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} \left(\tau_p + \psi \tau_s + \left(1 - \frac{M}{\beta(M-1)} \right) \tau_A \right) = \begin{bmatrix} \tau_p \mu - \frac{1}{\eta} \mu_X - \frac{M\tau_A}{\beta(M-1)} \alpha \\ \psi \tau_s - \frac{\tau_A}{M-1} \\ \frac{1}{\eta} - \frac{\gamma \tau_A}{\beta(M-1)} \end{bmatrix} \quad (\text{IA.6})$$

CLAIM 2. The unique solution to equations (IA.6) is:

$$\alpha = \frac{(\eta^2 \psi^2 \tau_p \tau_s + MV \tau_p) \mu - \left(\eta \psi^2 \tau_s + \frac{MV}{\eta} \right) \mu_X}{\eta^2 \psi^2 \tau_s (\tau_p + (M + \psi - 1) \tau_s) + (\tau_p + \psi \tau_s) MV} \quad (\text{IA.7a})$$

$$\beta = \frac{\psi \tau_s (\eta^2 \psi (M + \psi - 1) \tau_s + MV)}{\eta^2 \psi^2 \tau_s (\tau_p + (M + \psi - 1) \tau_s) + (\tau_p + \psi \tau_s) MV} \quad (\text{IA.7b})$$

$$\gamma = \frac{\eta^2 \psi (M + \psi - 1) \tau_s + MV}{\eta (\eta^2 \psi^2 \tau_s (\tau_p + (M + \psi - 1) \tau_s) + (\tau_p + \psi \tau_s) MV)} \quad (\text{IA.7c})$$

PROOF. Consider $\Gamma \equiv \frac{\gamma}{\beta}$. We already established that $\beta \neq 0$ so Γ is finite. Substituting Γ into (IA.6) and dividing the γ equation by the β equation yields:

$$\begin{aligned} \Gamma &= \frac{\frac{1}{\eta} - \Gamma \frac{\tau_A}{M-1}}{\psi \tau_s - \frac{\tau_A}{M-1}} \\ &= \frac{1}{\eta \psi \tau_s} \end{aligned} \quad (\text{IA.8})$$

Plugging (IA.8) into the β equation of (IA.6) yields:

$$\beta = \frac{\psi \tau_s (\eta^2 \psi (M + \psi - 1) \tau_s + MV)}{\eta^2 \psi^2 \tau_s (\tau_p + (M + \psi - 1) \tau_s) + (\tau_p + \psi \tau_s) MV} \quad (\text{IA.9})$$

Plugging (IA.8) and (IA.9) into the γ equation of (IA.6) yields:

$$\gamma = \frac{\eta^2 \psi (M + \psi - 1) \tau_s + MV}{\eta (\eta^2 \psi^2 \tau_s (\tau_p + (M + \psi - 1) \tau_s) + (\tau_p + \psi \tau_s) MV)} \quad (\text{IA.10})$$

Finally, plugging (IA.8), (IA.9), and (IA.10) into the α equation of (IA.6) yields:

$$\alpha = \frac{(\eta^2 \psi^2 \tau_p \tau_s + MV \tau_p) \mu - \left(\eta \psi^2 \tau_s + \frac{MV}{\eta} \right) \mu_X}{\eta^2 \psi^2 \tau_s (\tau_p + (M + \psi - 1) \tau_s) + (\tau_p + \psi \tau_s) MV} \quad (\text{IA.11})$$

■

Trading

Using equation (IA.4),

$$\begin{aligned} Trade_m &\equiv \sum_{i:m(i)=m} \left[D_i - \frac{M}{N} x_m \right] \\ &= \left\{ \frac{\eta}{M} \left(\psi \tau_s - \frac{\tau_A}{M-1} \right) (y_m - \bar{y}) \right\} - \left\{ \left(1 - \frac{\eta \gamma \tau_A}{(M-1) \beta} \right) (x_m - \bar{x}) \right\} \end{aligned} \quad (\text{IA.12})$$

All other trading derivations are in the main text of the paper.

Liquidity

Recall that illiquidity is defined as:

$$\lambda \equiv \frac{\frac{dP}{dy_m}}{\frac{dTrade_m}{dy_m}}$$

Taking derivatives of P (eq. IA.1) and $Trade_m$ (eq. IA.12) with respect to y_m and plugging in γ from (IA.7c) and Γ from (IA.8) yields:

$$\begin{aligned}
\lambda &= \frac{-\gamma}{\left(\frac{N}{M}\right) \left(\frac{\eta}{N}\right) \left(\frac{\gamma M \tau_A}{\beta(M-1)} + \gamma \left(\tau_p + \psi \tau_s + \left(1 - \frac{M}{\beta(M-1)}\right) \tau_A\right)\right) - 1} \\
&= \frac{M \left(\eta^2 \psi^2 \tau_s + MV\right) \left(\eta^2 \psi (M + \psi - 1) \tau_s + MV\right)}{\eta (M - 1) \left(\eta^2 (\psi^2 - \psi) \tau_s + MV\right) \left(\eta^2 \psi^2 \tau_s (\tau_p + (M + \psi - 1) \tau_s) + (\tau_p + \psi \tau_s) MV\right)} \\
&= \left\{ \frac{M}{(M - 1) \eta (\tau_p + \psi \tau_s + \tau_A)} \right\} \\
&\quad + \left\{ \frac{\eta^2 \psi^2 \tau_s}{\eta^2 \psi^2 \tau_s + MV} \right\} \left\{ \frac{\tau_s}{\tau_p + \psi \tau_s + \tau_A} \right\} \left\{ \frac{\eta}{M} \left(\psi \tau_s - \frac{\beta}{M} (\tau_p + \psi \tau_s + \tau_A) \right) \right\}^{-1} \\
&= \{S\} + \{B1\} \{B2\} \{B3\}^{-1} \tag{IA.13}
\end{aligned}$$

Consistent with the baseline model, $\lim_{V \rightarrow 0} \frac{d\lambda}{d\psi} < 0$, $\lim_{V \rightarrow 0} \frac{d\lambda}{d\tau_s} < 0$, and $\lim_{V \rightarrow 0} \frac{d\lambda}{d\tau_p} < 0$. From (IA.13) one can see that $\frac{d\lambda}{d\tau_p} < 0$ for all V . However, $\frac{d\lambda}{d\psi}$ and $\frac{d\lambda}{d\tau_s}$ are not always negative. Their signs are determined by complicated functions of the parameters. Considering limiting cases is instructive. We have already seen that $\frac{d\lambda}{d\psi}$ and $\frac{d\lambda}{d\tau_s}$ are negative in the limit as $V \rightarrow 0$. Both are also negative in the limit as $V \rightarrow \infty$. For interim values of V (i.e., positive, finite V), $\frac{d\lambda}{d\psi}$ and $\frac{d\lambda}{d\tau_s}$ can be positive or negative. Both follow a similar pattern. As $\psi \rightarrow 0$ or $\tau_s \rightarrow 0$, $\lambda \rightarrow \frac{M}{(M-1)\eta\tau_p}$, which is solely a supply impact – it includes no belief price response.^{1,2} $\frac{d\lambda}{d\psi}$ and $\frac{d\lambda}{d\tau_s}$ initially have the same sign as $\eta^2 \tau_p - V$ (i.e., $\text{sign} \left[\lim_{\psi \rightarrow 0} \frac{d\lambda}{d\psi} \right] = \text{sign} \left[\lim_{\tau_s \rightarrow 0} \frac{d\lambda}{d\tau_s} \right] = \text{sign} [\eta^2 \tau_p - V]$). As ψ and τ_s increase, they eventually decrease λ , driving it to approach zero as $\psi \rightarrow \infty$ or $\tau_s \rightarrow \infty$.

Liquidity without overconfidence

Private information precision can enhance liquidity even without overconfidence. When $\psi = 1$ (which reproduces the model of Diamond and Verrecchia (1981)), illiquidity is:

$$\lambda_{\psi=1} = \frac{M \left(\eta^2 \tau_s + MV\right) \left(\eta^2 \tau_s + V\right)}{\eta (M - 1) \left(\eta^2 \tau_s (\tau_p + M \tau_s) + (\tau_p + \tau_s) MV\right) V} \tag{IA.14}$$

¹Though I restrict my attention to overconfidence ($\psi > 1$) in other parts of the paper, it is useful to generalize and consider underconfidence ($\psi < 1$) here to get a full picture of the relationship between λ and ψ .

²The total risk tolerance of agents not receiving the shock is $\frac{(M-1)\eta}{M}$ and their posterior variance is τ_p^{-1} .

and its derivative with respect to private information precision is:

$$\frac{d\lambda}{d\tau_s}_{\psi=1} = \frac{M (\eta^6 \tau_p \tau_s^2 + \eta^4 \tau_s (2\tau_p - M\tau_s) MV + \eta^2 (\tau_p - 2\tau_s) M^2 V^2 - M^2 V^3)}{\eta (M-1) (\eta^2 \tau_s (\tau_p + M\tau_s) + (\tau_p + \tau_s) MV)^2 V} \quad (\text{IA.15})$$

The $V \rightarrow 0$ limit is uninteresting because without endowment shocks or overconfidence, illiquidity is infinite.

As $V \rightarrow \infty$, $\lambda_{\psi=1} \rightarrow \frac{M}{(M-1)\eta(\tau_p+\tau_s)}$ and $\frac{d\lambda}{d\tau_s}_{\psi=1} \rightarrow \frac{-M}{(M-1)\eta(\tau_p+\tau_s)^2} < 0$. Under infinite supply variance, liquidity trading swamps informed trading so trades carry no information. Thus, only the supply channel is operative, and the supply illiquidity channel always decreases as information (public or private) increases.

For interim values of V , λ starts off as solely a supply effect: $\lim_{\tau_s \rightarrow 0} \lambda_{\psi=1} = \frac{M}{(M-1)\eta\tau_p}$. As τ_s increases, the supply illiquidity channel decreases, but the belief illiquidity channel increases at least initially. For large V , the decreasing supply channel is more powerful. For small V , the increasing belief channel is more powerful. Specifically, $\lim_{\tau_s \rightarrow 0} \frac{d\lambda}{d\tau_s}_{\psi=1} = \frac{M(\eta^2 \tau_p - V)}{\eta \tau_p^2 (M-1)V}$. For large τ_s , only the belief channel is operative, and $\lim_{\tau_s \rightarrow \infty} \lambda_{\psi=1} = \frac{\eta}{(M-1)V}$. Note that this is a positive constant whereas $\lim_{\tau_s \rightarrow \infty} \lambda = 0$ when $\psi > 1$. The belief channel consistently increases with τ_s when V is small, but when V is large, τ_s eventually decreases the belief channel, thereby decreasing overall illiquidity as well. Specifically, $\text{sign} \left[\lim_{\tau_s \rightarrow \infty} \frac{d\lambda}{d\tau_s}_{\psi=1} \right] = \text{sign} [\eta^2 \tau_p - M^2 V]$. Another point of interest is to compare illiquidity at the two limits of τ_s : $\frac{\lim_{\tau_s \rightarrow \infty} \lambda_{\psi=1}}{\lim_{\tau_s \rightarrow 0} \lambda_{\psi=1}} = \frac{\eta^2 \tau_p}{MV}$.

The overall relationship between private information and illiquidity without overconfidence is as follows: For high supply variance ($V > \eta^2 \tau_p$), private information decreases illiquidity; for low supply variance ($V < \frac{\eta^2 \tau_p}{M^2}$), private information increases illiquidity; and for moderate supply variance ($\frac{\eta^2 \tau_p}{M^2} < V < \eta^2 \tau_p$), illiquidity is a hump-shaped function of private information. Within the moderate case, τ_s decreases illiquidity overall when $V > \frac{\eta^2 \tau_p}{M}$ and increases illiquidity overall when $V < \frac{\eta^2 \tau_p}{M}$.

B Empirical details and supplemental analysis

The paper proposes three stylized facts about trading and liquidity in stock, option, and bond markets. This appendix describes the paper’s empirical methodologies in more detail and reports stock results that are omitted in the main paper. It also reports additional stock market tests, including cross-sectional evidence and robustness checks across different types of stocks.

Data and methodology

I consider trading and liquidity of New York Stock Exchange (NYSE) stocks between 1926 and 2011. Focusing on the NYSE avoids differences in market structure and reporting across exchanges. Limiting attention to the NYSE also mitigates the impact of small stocks. The sample is also limited to stocks with share prices above \$5 at the end of the previous month.

Table IA.1 and Figure IA.1 summarize the data. Means represent equally weighted averages across stocks. Standard deviations are cross-sectional. Trading volume is the most straight-forward data item because it is directly observable in CRSP monthly data from 1926 to 2011. Average monthly turnover is 6.2% over the full sample with a standard deviation of 8.1%. On average, there are 1,100 observations per month. Turnover was relatively high (10%+) during the 1926 to 1935 period, dropped to the 2-5% range during the middle of the twentieth century, and then rose to 20% in 2011 with a peak of 40% in 2008. The time trend is clearest in Figure IA.1, panel A. Table IA.1 shows that cross sectional standard deviations followed a similar trend.

Within the model, the relevant concept of illiquidity is the price impact caused by a buy or sell trade. While illiquidity is not directly observable, a large literature has identified different ways of estimating it. The first measure of illiquidity I consider is $illiq_{it} = \frac{|Return_{it}|}{\$Vol_{it}}$, proposed by Amihud (2002). *illiq* has the advantage of directly relating price changes to volume. It is also readily computable using daily return and volume data for the full 1926 to 2011 time series. Table IA.1 shows that equally-weighted average *illiq* has decreased dramatically over time from 134% per \$100K of volume in the 1926 to 1945 period to 0.56% in the most recent period. Figure IA.1, panel B plots $\log(illiq)$ over time.

Bid-ask spread is another measure of illiquidity, and corresponds to price response to trading in Glosten and Milgrom (1985). I consider three variations of bid-ask spread. *bidask* is a daily

measure of a stock’s final quoted bid-ask spread, scaled by its closing price. Quoted bid-ask spread (*qbidask*) is the transaction-level counterpart to *bidask*, using intraday trades and quotes instead of end-of-day data. Using TAQ data, I match trades to quotes prevailing two seconds earlier. *qbidask* is the equally weighted average across transactions of quoted bid-ask spread scaled by transaction price. Finally, I consider the effective bid-ask spread (*ebidask*) measure of Chordia, Roll, and Subrahmanyam (2000). I calculate *ebidask* by matching trades with prevailing quotes from two seconds earlier. *ebidask* is twice the deviation between price and the bid-ask midpoint, scaled by price. Effective bid-ask spread takes into account that many trades are executed within quoted spreads and that large trades can take place outside of the spread if they exceed quoted depth. The main assumption behind *ebidask* is that deviations in price from the bid-ask midpoint represent buying or selling pressure. Like *qbidask*, *ebidask* is an equally weighted average across all of a stock’s transactions in a given day. Table IA.1 and panels C and D of Figure IA.1 summarize and plot the bid-ask data. TAQ transaction data is available for *qbidask* and *ebidask* starting in 1993. Daily data on *bidask* is available from CRSP for this period and prior to 1942, but is missing in the interim period. In the overlapping post-1993 period, daily bid-ask spreads average 1.1%. Transaction data on quoted spreads record intraday spreads half as large. Effective spreads are smaller still at 0.4%. All three time series follow a similar pattern with significant decreases in the early 2000’s. The drop-off in daily bid-ask spreads was particularly large. Where the series overlap, *illiq* and the bid-ask spread time series follow a similar pattern.

For overconfidence proxies, I rely on the learning models of Daniel, Hirshleifer, and Subrahmanyam (1998) and Gervais and Odean (2001). Self-attribution bias causes investors to be particularly overconfident following high returns. Because investors hold the market on average, aggregate overconfidence should track market returns. Individual stock returns may also affect stock-level overconfidence. To the extent that investors specialize in certain industries or have industry-specific confidence levels, past industry returns will also affect stock-level overconfidence. Statman, Thorley, and Vorkink (2006) employ market and individual stock VAR analyses to show that aggregate turnover increases following high market returns and stock-level turnover increases following past market and stock-level returns. I extend Statman, Thorley, and Vorkink’s market VAR to include measures of illiquidity as well as turnover. For stock-level analysis, I employ a panel VAR with stock-level turnover, stock-level illiquidity, stock-level returns, industry returns, and time and firm

fixed effects. I also test the effects of stock-level returns by sorting stocks into cross-sectional momentum deciles. Stock-level return data comes from CRSP. Excess market return (CRSP value weight market return less the risk free rate) and industry return data is from Ken French’s website. Return data is available starting in 1926. Industries are defined using Ken French’s 10 industry groups and assigned based on Compustat (and where missing CRSP) SIC codes (available after 1950).³

Changes in private information relative to public information are identified in two ways. First, periods around earnings announcements are likely to have elevated private information and public uncertainty. Prior to announcements, private information can be in the form of leaks and insider trading. After announcements, investors process different pieces of information at different paces using different models, keeping private information high until the announcement is fully digested and reflected in prices. Public uncertainty is also high around earnings announcements because asset values are highly sensitive to the announcements. Second, I follow Sadka and Scherbina (2007) and use dispersion of analyst forecasts as a proxy for private information relative to public information. Analyst dispersion may represent or cause public uncertainty and could also stem from more private information. Dispersion is measured as the standard deviation across analysts of current year earnings forecasts scaled by the mean forecast. Stocks are included if they are covered by at least two analysts, have a non-zero mean earnings forecast, and have a December fiscal year end. The December fiscal year requirement ensures that all stocks have the same amount of time remaining in the current fiscal year. Monthly analyst forecasts, summarized in Table IA.1, are available from I/B/E/S starting in 1976. Valid dispersion data is available for about 60% of the stocks in the other samples. Analyst dispersion has decreased over time from 20.6% in the 1976 to 1992 period to 13.8% in the 1993 to 2011 period.

All of the variables in Table IA.1 and Figure IA.1 show significant time trends and changing standard deviations. Regressions analyze log variables to make changes more comparable over time. Panel analysis controls for time and firm fixed effects. For market vector autoregression (VAR) analysis, log time series are detrended using a Hodrick and Prescott (1997) filter. The dashed lines in Figure IA.1 plot the trends. Figure IA.2 shows the detrended time series.

³I follow Ken French’s methodology of updating industry definitions at the end of each June based on SIC codes from the end of the previous year.

Fact 1: Trade and liquidity are positively correlated

I first examine the reduced form relationship between turnover and liquidity in the cross section. At the end of each month, stocks are sorted into decile portfolios based on turnover in the past month. Table IA.2 shows the results. Turnover is persistent over the next month, and high turnover stocks are more liquid. Across all four measures, illiquidity monotonically decreases as turnover increases and differences between deciles 1 and 10 are highly significant.

A potential concern with the decile sort methodology is that lagged turnover is likely correlated with omitted variables. To control for omitted variables, the panel regressions in Table 1 include stock, bond, and month fixed effects. The resulting identification is based on changes to turnover over time for a given firm. Table IA.3 reports supplemental stock results using *illiq* and alternative bid-ask spread measures. Consistent, with Table 1, turnover is negatively related to all illiquidity measures.

Fact 2: As private information increases relative to public information, trade increases and liquidity decreases

Figure IA.3 plots daily and quoted transaction-level bid-ask spreads around earnings announcements. Results are similar to the effective bid-ask spread results plotted in Figure 6.

Table IA.4 reports turnover and illiquidity for decile portfolios formed monthly by sorts lagged analyst dispersion. Consistent with Sadka and Scherbina (2007), illiquidity increases with dispersion. The relationship is monotonic for all illiquidity measures with the exception of reductions between deciles 1 and 2. Decile 10 is significantly more illiquid than decile 1 across all illiquidity measures. Daily bid-ask spreads are 0.5% higher in decile 10. Intraday spreads are 0.2% to 0.3% higher. *illiq* is 0.5% higher. Turnover also increases monotonically as dispersion increases, and decile 10 turnover is a significant 5.3% higher than decile 1.

Table IA.5 reports analyst dispersion panel regression results for alternative illiquidity measures. Consistent with results in Table 2, analyst earnings forecast dispersion is positively related to all illiquidity measures.

The results show that as information heterogeneity (proxied by analyst forecast dispersion and earnings announcements) increases, trade increases and liquidity decreases. These relationships

could vary across stocks. To test for differences across stocks, I add stock characteristic interactions to the analyst forecast dispersion regressions and replicate the earnings announcement event studies for different subsets of stocks. Results are robust across most stock characteristics. Across almost all sorts, as information heterogeneity increases, trading tends to increase and liquidity tends to decrease.

Size

I sort stocks at the end of June in each year based on their current market capitalization. Large stocks have market capitalizations above the NYSE median. Small stocks have market capitalizations below the median. Results are in Table IA.6 and Figure IA.4.

- Turnover increases with analyst forecast dispersion only for large stocks.
- Analyst forecast dispersion decreases liquidity for all stocks.
- Size does not have a major impact on earnings announcement results.

Book-to-market ratios

I sort stocks at the end of June based on book to market ratios at the end of the previous calendar year. Breakpoints for growth (low B/M ratio), neutral (medium B/M ratios), and value (high B/M ratios) are the NYSE 30th and 70th percentiles. Results are in Table IA.7 and Figure IA.5.

- Analyst dispersion decreases liquidity for all groups, but the effect decreases with B/M ratios. Thus, growth stock liquidity is more sensitive to analyst dispersion than value stock liquidity is.
- Earnings announcement illiquidity results are generally unaffected by B/M ratios.

Momentum

I sort stocks monthly based on returns from twelve months ago to one month ago. Breakpoints for low, medium, and high momentum are the NYSE 30th and 70th percentiles. Results are in Table IA.8 and Figure IA.6.

- Turnover only increases with analyst forecast dispersion for medium and (most significantly) high momentum stocks.
- Analyst forecast dispersion has less of an impact on liquidity for medium and (most significantly) high momentum stocks.
- Earnings announcement results are insensitive to momentum.

Lagged returns

I sort stocks based on returns in the past month. Breakpoints for low, medium, and high prior month returns are the NYSE 30th and 70th percentiles. Results are in Table IA.9 and Figure IA.7.

- Results are consistent with the momentum sorts. Stocks with high past returns have higher turnover sensitivity and lower liquidity sensitivity to analyst forecast dispersion, and earnings announcement results are insensitive to past returns.

Institutional ownership

I calculate institutional ownership (IO) for each stock at the end of each year using Thomson 13F data. I sort stocks at the end of the following June into low, medium, and high IO groups based NYSE 30th and 70th percentile breakpoints. Results are in Table IA.10 and Figure IA.8.

- High IO stocks tend to be more liquid and have higher turnover, but analyst forecast dispersion and earnings announcement results are insensitive to IO levels.

Fact 3: Trade and liquidity are elevated following high past returns

Table IA.11 reports parameter estimates and bootstrapped standard errors for the stock market VAR.

Figure IA.9 plots impulse response functions for the stock market VAR, including responses that are omitted from Figure 7 for brevity. All shocks are one standard deviation in magnitude. The plot summarizes how the shocks affect forecasts 1-5 months in the future. Turnover and *illiq* are logs so their responses can be interpreted as percent changes. Solid lines represent the impulse response functions. Dashed lines are bootstrapped 95% confidence intervals. The first row of Figure IA.9

shows responses to a one standard deviation shock to turnover. Turnover itself remains elevated for 2-3 months. *illiq* decreases by about 2% and then returns to normal. Returns are largely unaffected. Row 2 shows that *illiq* impulses are persistent and have a slight negative effect on turnover and positive effect on market returns. By contrast, the return impulse (row 3) is barely persistent at all, but it significantly decreases illiquidity and increases turnover for about five months. The initial response is +5% for turnover and -8% for *illiq*. Assuming return shocks increase overconfidence, this is what proposition 1 predicts.

As a robustness check, I repeat the market VAR analysis of Table IA.11 and Figure IA.9 on bid-ask spreads, which are unfortunately only available in uninterrupted time series starting in 1993. Figure IA.10 shows impulse responses to one standard deviation return shocks for daily, quoted, and effective bid-ask spreads. Standard errors are larger, but the market return impulse significantly reduces all bid-ask spread forecasts, and the responses last a full five months, supporting the initial market VAR results. In the shortened 1993 to 2011 time period, returns no longer positively predict future trading. The result goes slightly the opposite way during this sample.

Table IA.12 reports panel VAR coefficient estimates and standard errors. The primary coefficients of interest are turnover and illiquidity on lagged returns and lagged industry returns. All four coefficients have the expected signs. Lagged stock returns predict decreased illiquidity and slightly (though insignificantly) predict increased turnover. Similarly, lagged industry returns predict increased turnover and decreased illiquidity (with an insignificant t-statistic for illiquidity).

As a final test, stocks are sorted cross-sectionally based on momentum (stock-level returns from 12 months ago to 1 month ago). Table IA.13 reports turnover and illiquidity for portfolios formed from monthly momentum sorts. Turnover follows a U-shaped pattern, suggesting portfolio rebalancing or other trading motivations in extreme portfolios. Nonetheless, turnover is highest in the high momentum portfolios. Monthly turnover is 2.1% higher in decile 10 than in decile 1 with a 0.2% standard error. All illiquidity measures monotonically decrease with momentum except in the highest momentum deciles, where there is a small increase in illiquidity. Differences between deciles 1 and 10 are all negative and highly significant.

Figure IA.1: Monthly Time Series. Solid lines are the actual data. Dashed lines are trends calculated using a Hodrick Prescott filter with a penalty value of 14,400 on the log data.

Figure IA.2: Detrended Log Monthly Time Series. The log time series were detrended using a Hodrick Prescott filter with a penalty value of 14,400.

Figure IA.3: Turnover and Liquidity Around Earnings Announcements. Turnover and bid-ask spreads are scaled by average daily values over the three calendar months before the earnings announcement. Solid lines are equally weighted averages across all stocks. Dashed lines are 95% confidence intervals. Day 0 is the day of the earnings announcement.

Figure IA.4: Turnover and Liquidity Around Earnings Announcements by Size. Turnover and bid-ask spreads are scaled by average daily values over the three calendar months before the earnings announcement. Solid lines are equally weighted averages across all stocks. Dashed lines are 95% confidence intervals. Day 0 is the day of the earnings announcement.

Figure IA.5: Turnover and Liquidity Around Earnings Announcements by Book-to-Market Ratio. Turnover and bid-ask spreads are scaled by average daily values over the three calendar months before the earnings announcement. Solid lines are equally weighted averages across all stocks. Dashed lines are 95% confidence intervals. Day 0 is the day of the earnings announcement.

Figure IA.6: Turnover and Liquidity Around Earnings Announcements by Momentum. Turnover and bid-ask spreads are scaled by average daily values over the three calendar months before the earnings announcement. Solid lines are equally weighted averages across all stocks. Dashed lines are 95% confidence intervals. Day 0 is the day of the earnings announcement.

Figure IA.7: Turnover and Liquidity Around Earnings Announcements by Prior Month Returns. Turnover and bid-ask spreads are scaled by average daily values over the three calendar months before the earnings announcement. Solid lines are equally weighted averages across all stocks. Dashed lines are 95% confidence intervals. Day 0 is the day of the earnings announcement.

Figure IA.8: Turnover and Liquidity Around Earnings Announcements by Institutional Ownership. Turnover and bid-ask spreads are scaled by average daily values over the three calendar months before the earnings announcement. Solid lines are equally weighted averages across all stocks. Dashed lines are 95% confidence intervals. Day 0 is the day of the earnings announcement.

Figure IA.9: Impulse Response Functions from *illiq* VAR. The first variable in each panel title is the impulse variable. The second variable is the response variable. The solid lines are responses to one standard deviation shocks to the impulse variables after the number of lags indicated on the horizontal axis. The dashed lines are 95% confidence intervals. *turn* and *illiq* are detrended log market turnover and market illiq (a measure of illiquidity), respectively. *rmrf* is the excess return of the CRSP value weighted market index over the risk free rate.

Figure IA.10: Impulse Response Functions from *bidask*, *qbidask*, and *ebidask* VARs. Each VAR includes detrended log market turnover (*turn*), CRSP value-weighted market returns in excess of the risk free rate (*rmrf*), and a detrended log measure of market illiquidity (daily bid-ask spreads in the first row, and intra-day quoted and effective bid-ask spreads in the second and third rows, respectively). The first variable after the colon in each panel title is the impulse variable. The second variable is the response variable. The solid lines are responses to one standard deviation shocks to the impulse variables after the number of lags indicated on the horizontal axis. The dashed lines are 95% confidence intervals. For brevity, only the most relevant impulse response functions are shown.

Table IA.1: Data Summary

This table reports the sources, availability, and summary statistics for the main data items analyzed in this paper. The data sample is NYSE stocks with lagged prices greater than \$5. Means and standard deviations are equally weighted across stocks. The reported values are averages of means and cross-sectional standard deviations calculated monthly in the relevant time period.

Data Item	Source	Date Range	Full Sample					1926-1945		
			Obs	Obs/Month	Mean	Std Dev	Obs/Month	Mean	Std Dev	
turn	CRSP monthly	1926-2011	1,130,018	1,095	6.22%	8.07%	585	6.03%	13.45%	
illiq	CRSP daily	1926-2011	1,130,018	1,095	37.18%	182.76%	585	134.25%	688.01%	
bidask	CRSP daily	1926-1941; 1993-2011	433,171	1,031	2.54%	2.89%	554	4.27%	5.31%	
qbidask	TAQ	1993-2011	331,239	1,453	0.51%	0.46%				
ebidask	TAQ	1993-2011	331,239	1,453	0.36%	0.35%				
disp	IBES	1976-2011	347,864	805	17.14%	100.81%				

Data Item	Obs/Month	1946-1975			1976-1992			1993-2011		
		Obs/Month	Mean	Std Dev	Obs/Month	Mean	Std Dev	Obs/Month	Mean	Std Dev
turn	1,091	2.32%	3.39%		1,302	4.90%	4.93%	1,453	13.78%	12.62%
illiq	1,091	15.54%	59.98%		1,302	2.10%	5.36%	1,453	0.56%	3.50%
bidask								1,433	1.09%	0.85%
qbidask								1,453	0.51%	0.46%
ebidask								1,453	0.36%	0.35%
disp					696	20.64%	123.86%	903	14.02%	80.18%

Table IA.2: Turnover Deciles

Decile portfolios are formed at the end of each month by sorting stocks by turnover in the previous month. The table reports market capitalization, turnover, and illiquidity measures for the next month. The reported values are equally weighted averages of all stocks in the decile portfolio. Standard errors for the 10-1 portfolio difference are reported in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance. The data is for NYSE stocks with lagged prices greater than \$5.

Decile	Lagged Turnover	Market Cap (\$B)	Turnover	Illiq	Bid-Ask (Daily)	Bid-Ask (Quoted)	Bid-Ask (Effective)
1	0.84%	1.14	1.22%	166.62%	6.08%	0.99%	0.73%
2	1.76%	2.64	2.25%	46.77%	3.51%	0.62%	0.44%
3	2.41%	2.60	2.93%	31.12%	2.69%	0.52%	0.36%
4	3.04%	2.28	3.60%	24.15%	2.33%	0.46%	0.32%
5	3.75%	2.01	4.30%	21.54%	2.15%	0.43%	0.30%
6	4.62%	1.77	5.15%	20.24%	2.01%	0.42%	0.29%
7	5.76%	1.55	6.23%	18.19%	1.87%	0.41%	0.29%
8	7.41%	1.34	7.68%	15.59%	1.73%	0.41%	0.29%
9	10.27%	1.12	10.04%	14.06%	1.62%	0.42%	0.29%
10	22.71%	0.89	18.87%	12.06%	1.42%	0.43%	0.31%
10-1	21.87%*** (0.63%)	-0.24*** (4.56%)	17.64%*** (0.54%)	-154.56%*** (16.52%)	-4.65%*** (0.31%)	-0.56%*** (0.01%)	-0.42%*** (0.01%)
Date Range	1926-2011	1926-2011	1926-2011	1926-2011	1926-1941; 1993-2011	1993-2011	1993-2011

Table IA.3: Turnover Panel Regressions

Results are for stock-level panel regressions of log illiquidity measures on log turnover. Robust clustered (by stock) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance. Data includes all NYSE stocks with lagged prices greater than \$5.

	(1) $\log(illiq)$	(2) $\log(bidask)$	(3) $\log(qbidask)$	(4) $\log(ebidask)$
Log Turnover	-0.667*** (0.010)	-0.112*** (0.007)	-0.164*** (0.006)	-0.158*** (0.006)
Stock FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Observations	1,128,984	326,781	331,239	331,238
R^2	0.925	0.908	0.914	0.889
Date Range	1926-2011	1993-2011	1993-2011	1993-2011

Table IA.4: Analyst Dispersion Deciles

Decile portfolios are formed at the end of each month by sorting stocks by analyst dispersion in the past month. Analyst dispersion is the standard deviation of current year analyst earnings forecasts scaled by the mean forecast. The table reports market capitalization, turnover, and illiquidity measures for the next month. The reported values are equally weighted averages of all stocks in the decile portfolio. Standard errors for the 10-1 portfolio difference are reported in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance. The data is for NYSE stocks with lagged prices greater than \$5, coverage by at least two analysts, and December fiscal years.

Decile	Lagged Dispersion	Market Cap (\$B)	Turnover	Illiq	Bid-Ask (Daily)	Bid-Ask (Quoted)	Bid-Ask (Effective)
1	0.01	9.20	7.69%	0.54%	0.85%	0.37%	0.25%
2	0.02	7.25	8.09%	0.35%	0.79%	0.33%	0.23%
3	0.02	5.83	8.42%	0.36%	0.82%	0.35%	0.24%
4	0.03	4.96	8.86%	0.37%	0.85%	0.36%	0.25%
5	0.04	4.32	9.17%	0.40%	0.88%	0.38%	0.26%
6	0.05	4.11	9.58%	0.42%	0.91%	0.39%	0.28%
7	0.07	4.44	10.09%	0.47%	0.98%	0.42%	0.30%
8	0.10	4.09	10.90%	0.59%	1.05%	0.46%	0.33%
9	0.16	2.83	12.18%	0.67%	1.13%	0.51%	0.36%
10	1.22	1.96	13.04%	1.01%	1.32%	0.62%	0.45%
10-1	1.21*** (0.04)	-7.23*** (27.18%)	5.34%*** (0.32%)	0.46%*** (0.03%)	0.47%*** (0.03%)	0.26%*** (0.01%)	0.20%*** (0.01%)
Date Range	1976-2011	1976-2011	1976-2011	1976-2011	1993-2011	1993-2011	1993-2011

Table IA.5: Analyst Dispersion Panel Regressions

Results are for stock-level regressions of log turnover and log illiquidity measures on lagged (by one month) log analyst forecast dispersion. Robust clustered (by stock) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance. Data includes all NYSE stocks with lagged prices greater than \$5, at least 2 analyst forecasts, and fiscal years that end in December.

	(1) log(<i>turn</i>)	(2) log(<i>illiq</i>)	(3) log(<i>bidask</i>)	(4) log(<i>qbidask</i>)	(5) log(<i>ebidask</i>)
Lagged Log Analyst Dispersion	0.016*** (0.004)	0.207*** (0.007)	0.100*** (0.004)	0.107*** (0.003)	0.110*** (0.004)
Stock FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Observations	341,431	341,398	199,582	201,948	201,948
R^2	0.729	0.888	0.919	0.926	0.901
Date Range	1976-2011	1976-2011	1993-2011	1993-2011	1993-2011

Table IA.6: Analyst Dispersion Panel Regressions – Size

Results are for stock-level regressions of log turnover and log illiquidity measures on lagged (by one month) log analyst forecast dispersion. Robust clustered (by stock) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	(1) log(<i>turn</i>)	(2) log(<i>illiq</i>)	(3) log(<i>bidask</i>)	(4) log(<i>qbidask</i>)	(5) log(<i>ebidask</i>)
Lagged Log Analyst Dispersion	-0.012** (0.005)	0.180*** (0.008)	0.086*** (0.004)	0.098*** (0.004)	0.099*** (0.004)
Lagged log(<i>disp</i>) × Big	0.065*** (0.007)	-0.027** (0.011)	0.009 (0.006)	-0.005 (0.006)	-0.001 (0.006)
Big	0.305*** (0.026)	-1.121*** (0.046)	-0.220*** (0.025)	-0.321*** (0.023)	-0.311*** (0.024)
Stock FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Observations	331,395	331,363	193,007	195,294	195,294
R^2	0.737	0.902	0.922	0.931	0.908
Date Range	1976-2011	1976-2011	1993-2011	1993-2011	1993-2011

Table IA.7: Analyst Dispersion Panel Regressions – Book-to-Market Ratios

Results are for stock-level regressions of log turnover and log illiquidity measures on lagged (by one month) log analyst forecast dispersion. Robust clustered (by stock) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	(1) log(<i>turn</i>)	(2) log(<i>illiq</i>)	(3) log(<i>bidask</i>)	(4) log(<i>qbidask</i>)	(5) log(<i>ebidask</i>)
Lagged Log Analyst Dispersion	0.020*** (0.005)	0.145*** (0.010)	0.083*** (0.006)	0.090*** (0.005)	0.092*** (0.005)
Lagged log(<i>disp</i>) × Neutral BM	-0.001 (0.006)	0.034*** (0.011)	0.009 (0.006)	0.004 (0.005)	0.005 (0.006)
Lagged log(<i>disp</i>) × Growth	0.009 (0.008)	0.075*** (0.015)	0.017** (0.008)	0.017** (0.007)	0.020*** (0.007)
Neutral BM	0.013 (0.019)	-0.212*** (0.036)	-0.093*** (0.020)	-0.133*** (0.019)	-0.133*** (0.020)
Growth	0.109*** (0.026)	-0.330*** (0.056)	-0.119*** (0.030)	-0.163*** (0.029)	-0.160*** (0.030)
Stock FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Observations	331,309	331,277	192,974	195,261	195,261
R^2	0.735	0.893	0.921	0.929	0.905
Date Range	1976-2011	1976-2011	1993-2011	1993-2011	1993-2011

Table IA.8: Analyst Dispersion Panel Regressions – Momentum

Results are for stock-level regressions of log turnover and log illiquidity measures on lagged (by one month) log analyst forecast dispersion. Robust clustered (by stock) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	(1) log(<i>turn</i>)	(2) log(<i>illiq</i>)	(3) log(<i>bidask</i>)	(4) log(<i>qbidask</i>)	(5) log(<i>ebidask</i>)
Lagged Log Analyst Dispersion	0.004 (0.004)	0.210*** (0.007)	0.097*** (0.004)	0.103*** (0.004)	0.106*** (0.004)
Lagged log(<i>disp</i>) × Medium Momentum	0.019*** (0.003)	-0.033*** (0.005)	-0.014*** (0.003)	-0.013*** (0.003)	-0.015*** (0.003)
Lagged log(<i>disp</i>) × High Momentum	0.039*** (0.004)	-0.049*** (0.007)	-0.013*** (0.004)	-0.008** (0.003)	-0.011*** (0.004)
Medium Momentum	0.001 (0.011)	-0.222*** (0.018)	-0.150*** (0.012)	-0.145*** (0.010)	-0.158*** (0.010)
High Momentum	0.195*** (0.015)	-0.495*** (0.023)	-0.226*** (0.014)	-0.196*** (0.012)	-0.218*** (0.013)
Stock FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Observations	332,972	332,939	192,876	195,170	195,170
R^2	0.738	0.893	0.922	0.930	0.907
Date Range	1976-2011	1976-2011	1993-2011	1993-2011	1993-2011

Table IA.9: Analyst Dispersion Panel Regressions – Prior Month Returns

Results are for stock-level regressions of log turnover and log illiquidity measures on lagged (by one month) log analyst forecast dispersion. Robust clustered (by stock) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	(1) log(<i>turn</i>)	(2) log(<i>illiq</i>)	(3) log(<i>bidask</i>)	(4) log(<i>qbidask</i>)	(5) log(<i>ebidask</i>)
Lagged Log Analyst Dispersion	-0.002 (0.004)	0.229*** (0.007)	0.103*** (0.004)	0.112*** (0.004)	0.115*** (0.004)
Lagged log(<i>disp</i>) × Medium Lag Ret	0.018*** (0.002)	-0.033*** (0.003)	-0.008*** (0.002)	-0.011*** (0.002)	-0.012*** (0.002)
Lagged log(<i>disp</i>) × High Lag Ret	0.030*** (0.002)	-0.031*** (0.003)	-0.005** (0.002)	-0.004** (0.002)	-0.005*** (0.002)
Medium Lag Ret	-0.040*** (0.008)	-0.153*** (0.011)	-0.095*** (0.008)	-0.106*** (0.007)	-0.107*** (0.007)
High Lag Ret	0.089*** (0.008)	-0.242*** (0.010)	-0.101*** (0.007)	-0.095*** (0.006)	-0.098*** (0.006)
Stock FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Observations	341,232	341,200	199,427	201,789	201,789
R^2	0.731	0.888	0.920	0.927	0.902
Date Range	1976-2011	1976-2011	1993-2011	1993-2011	1993-2011

Table IA.10: Analyst Dispersion Panel Regressions – Institutional Ownership

Results are for stock-level regressions of log turnover and log illiquidity measures on lagged (by one month) log analyst forecast dispersion. Robust clustered (by stock) standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance.

	(1) log(<i>turn</i>)	(2) log(<i>illiq</i>)	(3) log(<i>bidask</i>)	(4) log(<i>qbidask</i>)	(5) log(<i>ebidask</i>)
Lagged Log Analyst Dispersion	0.024*** (0.008)	0.182*** (0.012)	0.096*** (0.007)	0.108*** (0.006)	0.111*** (0.006)
Lagged log(<i>disp</i>) × Medium IO	0.001 (0.008)	0.016 (0.013)	0.004 (0.007)	-0.004 (0.007)	-0.005 (0.007)
Lagged log(<i>disp</i>) × High IO	-0.015 (0.009)	0.024 (0.015)	0.001 (0.008)	-0.002 (0.008)	-0.002 (0.008)
Medium IO	0.119*** (0.028)	-0.260*** (0.050)	-0.089*** (0.027)	-0.120*** (0.026)	-0.132*** (0.026)
High IO	0.193*** (0.034)	-0.426*** (0.060)	-0.154*** (0.032)	-0.173*** (0.031)	-0.199*** (0.031)
Stock FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Observations	293,155	293,124	193,030	195,317	195,317
R^2	0.721	0.887	0.921	0.928	0.904
Date Range	1976-2011	1976-2011	1993-2011	1993-2011	1993-2011

Table IA.11: VAR Results

turn and *illiq* are detrended log turnover and *illiq* (a measure of illiquidity), respectively. *rmrf* is the excess return of the CRSP value weighted market return over the risk free rate. *turn* and *illiq* were detrended using a Hodrick and Prescott (1997) filter with a penalty value of 14,400. Reported results are for a 2-lag VAR of *turn*, *illiq*, and *rmrf*. Bootstrapped standard errors are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance. Turnover and illiq are equally weighted averages. Sample includes all NYSE stocks with lagged prices greater than \$5 from 1926 to 2011.

	(1) <i>turn</i>	(2) <i>illiq</i>	(3) <i>rmrf</i>
Lag 1			
<i>turn</i>	0.4835*** (0.0392)	-0.1083** (0.0437)	0.0137 (0.0120)
<i>illiq</i>	-0.0634 (0.0392)	0.4329*** (0.0507)	0.0225** (0.0110)
<i>rmrf</i>	0.8625*** (0.1905)	-1.5128*** (0.1687)	0.1133* (0.0612)
Lag 2			
<i>turn</i>	-0.0059 (0.0432)	0.0810** (0.0383)	-0.0049 (0.0094)
<i>illiq</i>	-0.0043 (0.0366)	0.2444*** (0.0450)	0.0167 (0.0103)
<i>rmrf</i>	0.3903** (0.1849)	0.0793 (0.1640)	0.0161 (0.0540)
Observations	1024	1024	1024
R^2	0.380	0.520	0.045
Date Range	1926-2011	1926-2011	1926-2011

Table IA.12: Panel VAR Results

turn and *illiq* are monthly stock-level log turnover and *illiq* (a measure of illiquidity), respectively. *ret* is the monthly individual stock returns. *ret_ind* is the monthly return on the stock's industry. Industries are defined using the 10 industry groups on Ken French's website. Reported results are for a 2-lag VAR of *turn*, *illiq*, *ret*, and *ret_ind*. Bootstrapped standard errors controlling for cross-sectional correlation are in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance. Sample includes all NYSE stocks with lagged prices greater than \$5 from 1951 to 2011.

	(1) <i>turn</i>	(2) <i>illiq</i>	(3) <i>ret</i>	(4) <i>ret_ind</i>
Lag 1				
<i>turn</i>	0.2295*** (0.0312)	-0.1645*** (0.0333)	0.0214* (0.0125)	0.0007 (0.0035)
<i>illiq</i>	-0.1209 (0.0803)	0.2187*** (0.0696)	0.0250 (0.0226)	0.0006 (0.0055)
<i>ret</i>	0.0558 (0.0935)	-0.2166** (0.0873)	-0.1021*** (0.0278)	-0.0001 (0.0066)
<i>ret_ind</i>	0.1776*** (0.0638)	-0.1159 (0.0740)	0.1119** (0.0486)	0.0515 (0.0713)
Lag 2				
<i>turn</i>	-0.0196 (0.0204)	0.0192 (0.0199)	0.0080 (0.0070)	0.0004 (0.0019)
<i>illiq</i>	-0.0410 (0.0343)	0.0626** (0.0297)	0.0139 (0.0097)	0.0005 (0.0024)
<i>ret</i>	-0.0116 (0.0250)	-0.0036 (0.0249)	-0.0325*** (0.0090)	-0.0018 (0.0021)
<i>ret_ind</i>	0.0840** (0.0403)	-0.0388 (0.0463)	0.0054 (0.0293)	-0.0212 (0.0413)
Stock FE	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes
Observations	751,221	751,221	751,221	751,221
Date Range	1951-2011	1951-2011	1951-2011	1951-2011

Table IA.13: Momentum Deciles

Decile portfolios are formed at the end of each month by sorting stocks by the 11-month return from 12 months ago to 1 month ago. The table reports market capitalization, turnover, and illiquidity measures for the next month. The reported values are equally weighted averages of all stocks in the decile portfolio. Standard errors for the 10-1 portfolio difference are reported in parentheses. * represents 10% significance, ** represents 5% significance, *** represents 1% significance. The data is for NYSE stocks with lagged prices greater than \$5.

Decile	Momentum	Market Cap (\$B)	Turnover	Illiq	Bid-Ask (Daily)	Bid-Ask (Quoted)	Bid-Ask (Effective)
1	-30.36%	0.85	7.63%	62.37%	3.56%	0.76%	0.55%
2	-14.21%	1.53	5.97%	42.26%	2.94%	0.60%	0.43%
3	-5.75%	1.81	5.51%	32.60%	2.59%	0.52%	0.37%
4	1.14%	1.96	5.22%	30.29%	2.41%	0.48%	0.34%
5	7.53%	2.03	5.12%	26.29%	2.23%	0.46%	0.33%
6	14.10%	2.11	5.23%	26.01%	2.16%	0.44%	0.31%
7	21.46%	2.16	5.39%	22.80%	2.08%	0.43%	0.30%
8	30.73%	2.15	5.82%	20.96%	2.01%	0.43%	0.30%
9	44.71%	1.98	6.68%	23.09%	2.05%	0.43%	0.31%
10	91.76%	1.32	9.75%	33.66%	2.20%	0.48%	0.34%
10-1	122.12%*** (1.87%)	0.47*** (7.30%)	2.12%*** (0.19%)	-28.71%*** (4.29%)	-1.36%*** (0.08%)	-0.28%*** (0.02%)	-0.21%*** (0.01%)
Date Range	1926-2011	1926-2011	1926-2011	1926-2011	1926-1941; 1993-2011	1993-2011	1993-2011