Uncertainty about What's in the Price

Joël Peress and Daniel Schmidt*

December 31, 2021

Abstract

Speculators face uncertainty about which signals are already reflected in the price. We present a model in which speculators update the probability that their information is truly novel rather than stale based on recent price movements and market makers are aware that speculators may be trading on stale news. The model predicts an asymmetric price response to past price movements: after a recent price increase, buy volume—because it may result from speculators trading on stale news—has a lower price impact than sell volume (and vice versa after a recent price decrease). Using a comprehensive sample of order flow imbalances and price impact costs, we find strong support for this prediction.

JEL classification: G11, G14

Keywords: Strategic trading, learning from prices

*Joël Peress is at INSEAD, Boulevard de Constance, 77300 Fontainebleau, France. Email: joel.peress@insead.edu. Daniel Schmidt is at HEC Paris, 1 Rue de la Libération, 78350 Jouy-en-Josas, France. Email: schmidt@hec.fr. We thank Dion Bongaerts, Thierry Foucault, Johan Hombert, Augustin Landier, Jeongmin Lee (WFA discussant), Shiyang Huang (EWFC discussant), Francesco Sangiorgi, Esad Smajlbegovic, and seminar/conference participants at Aarhus University, Erasmus University, Frankfurst School of Finance & Management brownbag, HEC Paris brownbag, the 2020 European Winter Finance Conference, 2020 the Western Finance Association Meeting.

1 Introduction

Asset prices reflect information. Yet, in a complex world it is seldom clear *which* information is already reflected in the price and which one is not. While there is a large literature on information asymmetry, informed trading, and learning from prices (e.g., Grossman, 1976; Grossman and Stiglitz, 1980; Hellwig, 1980; Kyle, 1985), this type of uncertainty is rarely captured in existing models. Indeed, almost all of the theoretical literature on this subject relies on an arguably implausible degree of common knowledge about the information structure faced by market participants. For example, it is typically assumed that all market participants know what type of signals, if any, are observed by *all* other market participants.[1] In practice, however, uncertainty about a stock is multidimensional and may depend on a variety of factors such as consumer demand, competition, takeover opportunities, technological changes, regulation etc. Given this complexity, it appears unrealistic to think that all investors know precisely how many other investors have information about each and every one of these dimensions of uncertainty. In other words, the assumption of complete knowledge of a stock's information environment—although common in the literature—is surely too restrictive.

This paper belongs to a nascent literature attempting to relax this restrictive common knowledge assumption. Prior work in this field has looked at the asset pricing implications of the uncertainty that results when *uninformed* investors are not sure about the presence of informed investors. In contrast, this paper focuses on the uncertainty faced by *informed* investors about how informed they really are: do they possess genuinely novel information—on which it would be profitable to trade—or do they possess *stale* information that is already reflected in the price? Such type of uncertainty

[1] For instance, in Grossman and Stiglitz (1980), investors are assumed to know both the *exact fraction* of informed and uninformed investors as well as *what* the informed investors are informed about (i.e., the fundamental is $u = \theta + \epsilon$ and informed investors are assumed to know θ).

should be very common. After all, prices can move for a myriad of reasons and it is difficult, nay impossible for investors to know the precise extent to which a recent price move is driven by this or that piece of information. Hence, when investors contemplate trading on a signal, they will not know how many other investors have traded on this information before them and thus how novel their signal truly is. In this paper, we put forth a parsimonious trading model in which investors face this type of uncertainty and describe the resulting equilibrium implications.

Two key insights emerge from the model. The first insight concerns investors' updating and was first pointed out by Treynor and Ferguson (1985): investors rely on past price movements to assess the novelty of their trading signals.[2] To see the intuition for this, consider an investor that has just unearthed positive information about a stock. Not knowing whether this information is already reflected in the price or not, the investor looks at recent price changes for guidance. If the stock has just gone up, it is possible that other investors have learned the same information before him, implying that his information is stale. In contrast, if the stock has gone down, the price movement must be explained by some different information, and so the investor concludes that his information is novel. These considerations lead the investor with positive news to trade more (less) aggressively after recent price downturns (upturns).

The second insight involves market makers' assessment of adverse selection risk and is, to the best of our knowledge, new to the literature: after recent price increases (decreases), market makers consider positive (negative) order flow to be less informative, because it could come from investors trading on stale news. Hence our model predicts that equilibrium price functions and trading strategies are asymmetric and depend on prior price movements. We argue that this type of asymmetric dependance only arises when there is uncertainty about what's in the price, making it a unique "footprint" to look for in empirical data. Moreover, we show that, by making investors reluctant to trade, uncertainty about what's in the price reduces stock price informativeness.

[2] Treynor and Ferguson (1985) make this point in defense of technical analysis. In their analysis, they take prices as given and do not spell out the equilibrium implications of such behavior.

The literature to date has focused on two different sources for an asymmetric price impact. The first is dynamic speculation (e.g., Llorente et al., 2002): when informed investors gradually establish their positions, a string of orders in the same direction signals their presence, prompting market makers to increase price impact. The second is inventory risk (e.g., Ho and Stoll, 1981; Madhavan and Smidt, 1993; Hendershott and Menkveld, 2014): when market makers are loath to deviate from a given target inventory level, they will require bigger price concessions for accommodating orders that push their inventory further away from target as compared to orders that allow them to move toward the target. In both cases buy (sell) orders that come after buy orders are expected to come with a larger (lower) price impact. Our model makes the exact opposite prediction: buy orders that come after prior buy orders are less informative (since they could come from speculators trading on stale news), implying a lower price impact. In practice, all of these channels are expected to co-exist. It is thus an empirical question whether our prediction regarding uncertainty about what's in the price prevails in the data.

We shed light on this question by examining price impact costs in an exhaustive sample of NYSE-traded stocks for the 1993 to 2014 period. Using the Lee and Ready (1991) algorithm to infer trade direction, we compute daily measures of trade imbalances from intraday TAQ data. This allows us to compare price impact costs on days with net-buying and net-selling activity as a function of past returns. Using four distinct proxies of price impact that reflect adverse selection costs, we find strong evidence that these costs behave asymmetrically in a manner consistent with our model.[3] Specifically, we document that on days with net-buying activity price impact costs are negatively related with past returns, while on days with net-selling activity price impact costs are positively related with past returns. Put differently, buys elicit a lower price impact

[3]Our four price impact measures are: i) the signed Amihud (2002) illiquidity ratio, named *price impact costs*, defined as the ratio of a stock's daily return (adjusted for autocorrelation) to its signed trade imbalance; ii) *lambda*, the slope coefficient from a regression of stock returns on signed order flow over five-minute intervals; iii) *quote-based price impact*, the percentage change in the mid-quote from before to five minutes after the transaction; iv) *ln(Amihud)*, the standard Amihud (2002) illiquidity ratio, defined as the logarithm of the stock's absolute return divided by its dollar volume.

when prior returns were positive, consistent with market makers understanding that investors are potentially buying based on stale news; for the same reason, sells elicit a lower price impact when prior returns were negative. We test auxiliary predictions of the model, namely that the asymmetric price pattern—if it is indeed caused by uncertainty about what's in the price—should be less pronounced when such uncertainty abates. We report two pieces of evidence consistent with this prediction. First, the asymmetric price impact pattern weakens immediately after earnings announcements when investors know better what information is already reflected in stock prices. Second, it also weaker for stocks for which more information is public; that is, for larger stocks and stocks with more analyst coverage. Overall, our results indicate that uncertainty about what's in the price is a real concern for investors. Indeed, we are not aware of any other theory that could explain why price impact costs for buys and sells are asymmetrically related to past price movements in the direction we report, and why this asymmetry should be consistently reduced after public earnings announcements or for stocks that are in the public limelight.

Encouraged by these results, we construct a cross-sectional measure of the extent of uncertainty about what's in the price based on the difference in sensitivities of price impact for buys and sells with respect to past returns. We then relate this measure to the information content of prices around earnings announcements. Following Weller (2018), we construct a measure of stock price informativeness—called the price jump ratio—defined as the fraction of the total earnings-related return change that occurs in the immediate aftermath of the announcement date. The higher this measure, the less information has entered stock prices before the announcement, indicating lower price informativeness. In panel regressions, we find that heightened uncertainty about what's in the price is consistently associated with a higher jump price ratio and thus with less informative stock prices. This confirms our model prediction that investors concerned about the novelty of their signals trade more cautiously and thereby slow down the information capitalization into stock prices.

We contribute to the literature on informed trading in financial markets. More specifically, our model is related to (but different from) existing models of uncertainty in investors' information environment. Easley and O'Hara (1992) present a sequential trade model in which a signal about the fundamental may or may not have been observed by speculators. In the model, market makers update the probability that such an "information event" occurred based on the direction and frequency of incoming orders. More recently, Banerjee and Green (2015) solve a dynamic model in which some investors learn about whether other investors are trading on signals or noise. Importantly, the learning investors are themselves uninformed—hence, they cannot use their own signal realization in combination with what is revealed through the price to update on that probability. It is this interplay between an investor's own signal realization and recent price changes that lies at the heart of our model.

The feature that investors update based on past prices is shared with the theoretical literature on "technical analysis" (e.g., Brown and Jennings, 1989; Grundy and McNichols, 1989; Brunnermeier, 2005). In these models, past prices are shown to have an *independent* signal value that is not subsumed by the current price, but the information structure remains common knowledge. Investors are therefore not worried that their signals may be stale; they simply use past prices to try to obtain a better estimate of the signal realizations observed by other investors. In contrast, the nature of learning from past prices is very different in our model: investors use them to update on the probability that others have seen their signal before them (thereby rendering it stale). Because this behavior is anticipated by market makers, prices respond asymmetrically to positive and negative order flow as a function of past price movements. This latter prediction—for which we find strong support in the data—distinguishes our model from other papers with learning from past prices.

Our paper further relates to Abreu and Brunnermeier (2002) and Abreu and Brunnermeier (2003). In these models, investors need to coordinate their arbitrage activity in order to correct a mispricing about which they become aware sequentially. Importantly, they assume that investors don't know their exact position in this sequence and show that the resulting synchronization risk gives rise to endogenous trading delays. We similarly assume that investors do not know whether they observe a given signal early or late, but instead of studying the coordination problem between investors, we focus on how investors update their beliefs about the extent of their information advantage based on (endogenous) past price movements.

Finally, our paper is related to Tetlock (2011), who shows that especially retail investors appear to be trading on stale news reprinted in the media. While his findings are consistent with an irrational overreaction to news, we argue that even sophisticated investors may find it difficult to judge the true value of a privately-acquired signal. We explore the ramifications of this idea in a model that is a entirely rational (apart from the usual assumption about uninformed noise trader demand) and shed a first light on its empirical relevance.

The paper proceeds as follows. Section 2 describes a simple trading game and solves it under different assumptions about the information structure. Section 3 presents empirical evidence in favor of our key model prediction regarding asymmetric price impact costs as a function of past returns. Section 4 tests whether uncertainty about what's in the price hurts stock price informativeness. Section 5 concludes.

2 Model

We develop a parsimonious model in which investors face uncertainty about what's in the price. We deliberately keep our model as simple as possible for ease of exposition.

2.1 Setup

There are three dates, denoted 0, 1, and 2, three categories of agents, namely market makers, speculators (or insiders), and noise traders, and a single stock. The stock pays a dividend of $\theta = \theta_1 + \theta_2$ at date t=2, where θ_1 and θ_2 are independent and both pay off $+\sigma$ or $-\sigma$ with equal probability. Hence, θ is -2σ with probability $1/4$, 0 with probability $1/2$, or $+2\sigma$ with probability $1/4$.

At dates t=0 and t=1, prices are set by competitive market makers (henceforth M) as in Kyle (1985). At date t=0, there are no speculators and no noise traders. Market makers observe a part of the fundamental θ_m where $m \in \{1, 2\}$ with equal probability, and equate the price of the asset, p_0, to their expectation of the dividend: $p_0 = E(\theta|\theta_m) = \theta_m$. At date t=1, a unit-mass of informed speculators (henceforth S) with mean-variance utility and risk aversion parameter γ all observe the same part of the fundamental θ_s where $s \in \{1, 2\}$ with equal probability. They then submit market orders conditional on the realization of their signal and the price at t=0, p_0 (that is, θ_m).[4] The variables m and s are drawn independently, implying that M and S observe with equal probability the same part of the fundamental ($m = s$) or different parts ($m \neq s$). At date t=1, there are also noise traders who submit a random market order, n, uniformly distributed over the interval $[-1, +1]$. Therefore, the total order flow at date t=1, ω_1, consists of the orders of the speculators and of the noise traders.

The following assumption ensures that the aggregate order flow at t=1 does not become fully revealing:

Assumption 1. *Let* $\gamma\sigma > 3$.

Our way of modeling the stock dividend as being determined by two parts, θ_1 and θ_2, captures in stylized fashion the idea that a stock's fundamental value depends on multiple sources uncertainty. Which bits and pieces are known to M and S, respectively, and whether they know what the others know are central elements to the model. Figure 1 summarizes the model setup.

[4] The assumption that all speculators observe the same part of the fundamental is not crucial for our argument. Indeed, the model's main prediction about investors' updating on the novelty of their signal based on past price movements remains valid if one, for example, assumes that each speculator i observes θ_{s_i} with $s_i \in \{1, 2\}$ being independent from m and s_j for all speculators $j \neq i$.

2.2 Benchmark: No uncertainty about what's in the price

We describe a benchmark in which speculators face no uncertainty about what's in the price. We first assume that both speculators and market makers know which component of the fundamental the other group observes. Then we assume that only market makers face uncertainty about what's in the price; that is, speculators know which component market makers observes, but not vice versa.

2.2.1 Both m and s are common knowledge

We start by assuming that both M and S know m and s. That is, both M and S know which part of the fundamental is observed by M at t=0 and whether S have observed the same part or not. Consider first the case $m = s$, meaning that both M and S know the same part of the fundamental, $\theta_m = \theta_s$. S then have no information advantage over M and thus refrain from trading ($\omega_1 = n$). Therefore, $p_0 = p_1 = \theta_m$.

Next, suppose M and S observe different parts of the fundamental; that is, $m \neq s$. In this case, M do not know the exact realization of θ_s, but they know that S have an information advantage (since they observe $p_0 = \theta_m$) they will trade on. M then try to back out θ_s from the order flow. We conjecture that S trade in a symmetric fashion and buy $x_{(1)}$ (sell $-x_{(1)}$) with $|x_{(1)}| < 1$ when they know $\theta_s = \sigma$ ($\theta_s = -\sigma$). Hence, the order flow is $\omega_1 = x_{(1)} + n$ if $\theta_s = \sigma$, and $\omega_1 = -x_{(1)} + n$ if $\theta_s = -\sigma$. If M observe $\omega_1 > -x_{(1)} + 1$ ($\omega_1 < x_{(1)} - 1$), then they infer that $\theta_s = \sigma$ ($\theta_s = -\sigma$). If instead $x_{(1)} - 1 \leq \omega_1 \leq -x_{(1)} + 1$, then it is both possible that S bought or sold and M learn nothing about θ_s. Accordingly, the equilibrium price is given by:

$$p_1 = \begin{cases} \theta_m + \sigma & \text{for} \quad -x_{(1)} + 1 < \omega_1 \leq x_{(1)} + 1 \\ \theta_m & \text{for} \quad x_{(1)} - 1 \leq \omega_1 \leq -x_{(1)} + 1 \\ \theta_m - \sigma & \text{for} \quad -x_{(1)} - 1 \leq \omega_1 < x_{(1)} - 1 \end{cases}$$

Each speculator i from the set S chooses her order x_i to maximize expected utility, taking the price function as given. Imposing rational expectations (i.e., $x_i = x_{(1)}$ for all

i in S) on the first-order condition yields the following equilibrium condition:

$$x_{(1)} = \frac{E\left[\theta - p_1 \,|\, p_0, \theta_s\right]}{\gamma\, Var\left[\theta - p_1 \,|\, p_0, \theta_s\right]}$$

Consider the case $m \neq s$ and $\theta_s = \sigma$ (the case $\theta_s = -\sigma$ is symmetric). We have $E\left[\theta - p_1 \,|\, p_0, \theta_s = \sigma, m \neq s\right] = \sigma(1 - x_{(1)})$ and $Var\left[\theta - p_1 \,|\, p_0, \theta_s = \sigma, m \neq s\right] = \sigma^2 x_{(1)}(1 - x_{(1)})$. Plugging these expressions into the previous equation yields $x_{(1)} = \sqrt{1/(\gamma\sigma)}$. The following proposition summarizes the equilibrium.

Proposition 2. *Assume that market makers M and speculators S know which part of the fundamental is observed by M and S; that is, m and s are common knowledge. At t=0, M set $p_0 = \theta_m$. At t=1:*

- *If $m = s$, then S refrain from trading, and M set $p_1 - p_0 = 0$.*
- *If $m \neq s$ and $\theta_s = \sigma$ ($\theta_s = -\sigma$), then S buy (sell) an amount $x_{(1)}$ ($-x_{(1)}$) where $x_{(1)} = \sqrt{1/(\gamma\sigma)}$. M set the price change, $p_1 - p_0$, independently of the realization of p_0, according to:*

$$p_1 - p_0 = \begin{cases} +\sigma & \textit{for} \quad -x_{(1)} + 1 < \omega_1 \leq x_{(1)} + 1 \\ 0 & \textit{for} \quad x_{(1)} - 1 \leq \omega_1 \leq -x_{(1)} + 1 \\ -\sigma & \textit{for} \quad -x_{(1)} - 1 \leq \omega_1 < x_{(1)} - 1 \end{cases} \tag{1}$$

Intuitively, speculators S trade less aggressively ($x_{(1)}$ lower) when they are more averse to risk (γ larger) and when the final payoff is more uncertain (σ larger). Note that $x_{(1)} < 1$ by Assumption 1, and hence the order flow $\omega_1 = x_{(1)} + n$ does not fully reveal θ_s, implying that S derive positive expected utility from trading. The solution resembles Vives (1995), who also models trading by a continuum of risk-averse speculators in the presence of competitive market makers (but with normally-distributed payoff and noise). The key feature of the equilibrium is that the price change, $p_1 - p_0$, does not depend on the lagged price p_0; that is, a buy or sell order triggers a price change of the same magnitude regardless of the lagged price.

2.2.2 Only m is common knowledge

We now assume that S know m and s, but that M only know m. In other words, both M and S know which part of the fundamental is observed by M (and thus reflected in p_0), but only S know whether they observed the same part or not. In essence, this setting features uncertainty for market makers about whether (better) informed speculators are present or not.

If M and S observe the same part of the fundamental (i.e., $m = s$), then S have no information advantage over M and thus refrain from trading ($\omega_1 = n$). If $m \neq s$, then S have additional information about θ and are conjectured to buy (sell) an amount $x_{(2)}$ ($-x_{(2)}$) when $\theta_s = \sigma$ ($\theta_s = -\sigma$). Market makers M do not know, however, which of these cases has occurred and try to learn from the order flow. If M observe $\omega_1 > 1$ ($\omega_1 < -1$), then they infer that $\theta_s = \sigma$ ($\theta_s = -\sigma$). If $-x_{(2)} + 1 \leq \omega_1 \leq 1$ ($-1 \leq \omega_1 < x_{(2)} - 1$), then M know that S did not sell $-x_{(2)}$ (did not buy $x_{(2)}$). In other words, M know that either $m = s$ (when S do not trade) or $\theta_s = \sigma$ ($\theta_s = -\sigma$). The conditional expectation of θ_s is then $\frac{1}{3}\sigma$ ($-\frac{1}{3}\sigma$) (see Appendix A.2). Finally, if $x_{(2)} - 1 \leq \omega_1 \leq -x_{(2)} + 1$, then M learns nothing about θ_s. Thefore, the equilibrium price function is given by Equation (2) below.

Given this price function, S choose the order size x that maximize their expected utility. Consider the case when $m \neq s$ and they know $\theta_s = \sigma$ (the case $\theta_s = -\sigma$ is symmetric). In this case, S are expected to buy, implying that the order flow is drawn at random from the interval $[x_{(2)} - 1, x_{(2)} + 1]$. It follows that (See Appendix A.2):

$$\begin{aligned} E(\theta - p_1 | p_0, \theta_s = \sigma, m \neq s) &= \sigma \left(1 - \frac{2}{3} x_{(2)}\right) \\ Var(\theta - p_1 | p_0, \theta_s = \sigma, m \neq s) &= \sigma^2 x_{(2)} \left(\frac{5}{9} - \frac{4}{9} x_{(2)}\right) \end{aligned}$$

Plugging these expressions into the first-order condition for S' profit maximization problem and imposing rational expectations (i.e., $x_i = x_{(2)}$ for all i in S) yields the optimal order size x. The following proposition summarizes the resulting equilibrium.

Proposition 3. *Assume that only m is common knowledge; that is, speculators S know which part of the fundamental is observed by market makers M but not vice versa. At t=0, M set $p_0 = \theta_m$. At t=1:*

- *If $m = s$, then S refrain from trading.*
- *If $m \neq s$ and $\theta_s = \sigma$ ($\theta_s = -\sigma$), then S buy (sell) an amount $x_{(2)}$ ($-x_{(2)}$) , where $x_{(2)}$ is defined in Appendix A.2. Market makers M don't know whether $m = s$ or $m \neq s$ and set the price change, $p_1 - p_0$, independently of the realization of p_0, according to:*

$$p_1 - p_0 = \begin{cases} +\sigma & \text{for} \quad 1 < \omega_1 \leq x_{(2)} + 1 \\ +\frac{1}{3}\sigma & \text{for} \quad -x_{(2)} + 1 < \omega_1 \leq 1 \\ 0 & \text{for} \quad x_{(2)} - 1 \leq \omega_1 \leq -x_{(2)} + 1 \\ -\frac{1}{3}\sigma & \text{for} \quad -1 \leq \omega_1 < x_{(2)} - 1 \\ -\sigma & \text{for} \quad -x_{(2)} - 1 \leq \omega_1 < -1 \end{cases} \tag{2}$$

Proof. See Appendix A.2. □

As before, the equilibrium trading aggressiveness, $x_{(2)}$, is decreasing in γ and σ. Hence, S trade less aggressively when they are more risk averse or when the stock's payoff is more uncertain. Importantly, the price change, $p_1 - p_0$, again does not depend on the lagged price p_0. As we shall see, this is no longer the case when speculators face uncertainty about what's in the price.

We briefly compare this version of the model to Banerjee and Green (2015), who also solve a rational expectation equilibrium model in which there is uncertainty about whether informed traders are present or not. One key difference is that here prices are assumed to be set by competitive market makers, whereas Banerjee and Green

(2015) rely on market-clearing by risk-averse investors. Since in Banerjee and Green (2015) the asset is in positive supply, prices reflect a risk premium and there emerges an asymmetry: both high and low price signals increase investors' beliefs about there being informed traders and thus command a larger price discount, which attenuates (resp. amplifies) the market response to positive (resp. negative) news. Since market makers are risk-neutral in our model, this risk discount effect is shut down here and the price function remains symmetric despite of the uncertainty about whether there are informed traders.

2.3 Uncertainty about what's in the price: Neither m nor s are common knowledge

We now tackle the case of interest—that is, the case in which speculators face uncertainty about whether their trading signals are already reflected in the price. Specifically, we assume that M know m and S know s, but neither group knows which part of the fundamental was observed by the other. That is, as in the model solution discussed in Section 2.2.2, M do not know whether S observed the same part of the fundamental or not, but now this uncertainty also extends to S. As a result, when investors Ss' signal coincides with Ms' as revealed by the t=0 price ($\theta_m = \theta_s$), then investors S are unsure whether they observed the same part of the fundamental or whether they actually observed the other part and it just happens that this news goes in the same direction. When $\theta_m \neq \theta_s$, however, then S infer that $m \neq s$ and they understand that they have truly novel information.

To solve for the trading equilibrium, we conjecture that S buy (sell) an amount $x_{(3)}$ ($-x_{(3)}$) when $\theta_m \neq \theta_s$ and that they buy (sell) an amount $y_{(3)}$ ($-y_{(3)}$) when $\theta_m = \theta_s$ with $x \geq y$. Consider the case $\theta_m = \sigma$. Given the conjecture, M expect S to either buy $y_{(3)}$ (when $\theta_s = \sigma$) or sell $-x_{(3)}$ (when $\theta_s = -\sigma$). If $-x_{(3)} + 1 < \omega_1 \leq y_{(3)} + 1$, then M infer that S bought y. Ms' expection of the component of θ that they do not observe is then $\frac{1}{3}\sigma$ (see footnote 4), resulting in a price of $\sigma + \frac{1}{3}\sigma = \frac{4}{3}\sigma$. If $-x_{(3)} - 1 \leq \omega_1 < y_{(3)} - 1$, then M know that S sold $x_{(3)}$. Their expection of the component of θ that they do not observe is therefore $-\sigma$ and so the price equals $\sigma - \sigma = 0$. If $y_{(3)} - 1 \leq \omega_1 \leq -x_{(3)} + 1$, then it is both possible that S bought or sold; thus M learn nothing from the order flow and set the price according to their private signal, to σ. As a result, the price function is given by Equation (3) below. For the case $\theta_m = -\sigma$, the logic is reversed. S either buy $x_{(3)}$ or sell $-y_{(3)}$, and M draw analogous inferences from the order flow, leading to the price function (4).

The crucial feature of these price functions is their asymmetry: market makers anticipate that selling volume after a price increase is a more informative signal about the stock's fundamental compared to buying volume. The reason is that selling volume after a price increase indicates that speculators trade on genuine new information, whereas buying volume can also come from the speculators trading on stale information (i.e., information already reflected in p_0). Hence, price impact is larger for selling (buying) volume after recent price increases (decreases).

We now solve for the equilibrium trading quantities, $x_{(3)}$ and $y_{(3)}$. Consider the case $\theta_m = -\sigma$ (as usual, the case $\theta_m = \sigma$ is symmetric), which is revealed to investors S by the t=0 price. When $\theta_s = \sigma$, an investor i in set S expects the other investors in S to buy an amount $x_{(3)}$, resulting in an order flow drawn at random from the interval $[x_{(3)} - 1, x_{(3)} + 1]$. The investor then calculates (see Appendix A.3):

$$
\begin{aligned}
E(\theta - p_1 | \theta_m = -\sigma, \theta_s = \sigma) &= = \sigma \left(1 - \frac{x_{(3)} + y_{(3)}}{2}\right) \\
Var(\theta - p_1 | \theta_m = -\sigma, \theta_s = \sigma) &= \sigma^2 \left(1 - \frac{x_{(3)} + y_{(3)}}{2}\right) \frac{x_{(3)} + y_{(3)}}{2}
\end{aligned}
$$

When $\theta_s = -\sigma$, investors expect other investors in S to sell an amount $-y_{(3)}$, resulting in an order flow drawn at random from the interval $[-y_{(3)} - 1, -y_{(3)} + 1]$. The investor then calculates (see Appendix A.3):

$$\begin{aligned}
E(\theta - p_1|\theta_m = -\sigma, \theta_s = -\sigma) &= -\frac{1}{3}\sigma\left(1 - \frac{x_{(3)} + y_{(3)}}{2}\right) \\
Var(\theta - p_1|\theta_m = -\sigma, \theta_s = -\sigma) &= \frac{1}{9}\sigma^2\left(8 + \frac{x_{(3)} + y_{(3)}}{2}\left(1 - \frac{x_{(3)} + y_{(3)}}{2}\right)\right)
\end{aligned}$$

Investors first-order condition together with requiring rational expectations for the two cases $\theta_s = \sigma$ and $\theta_s = -\sigma$ (i.e., $x_i = x_{(3)}$ in the former case and $x_i = y_{(3)}$ in the latter) yields a system of two equations in $x_{(3)}$ and $y_{(3)}$. The following proposition summarizes the resulting equilibrium.

Proposition 4. *Assume that neither m nor s are common knowledge. At t=0, market makers M set $p_0 = \theta_m$. At t=1:*

- *If $\theta_m \neq \theta_s$ and $\theta_s = \sigma$ ($\theta_s = -\sigma$), then speculators S buy (sell) an amount $x_{(3)}$ ($-x_{(3)}$), where $x_{(3)}$ is defined in Appendix A.3.*
- *If $\theta_m = \theta_s$ and $\theta_s = \sigma$ ($\theta_s = -\sigma$), then S buy (sell) an amount $y_{(3)}$ ($-y_{(3)}$), where $y_{(3)} = \frac{2-\gamma\sigma x_{(3)}^2}{\gamma\sigma x_{(3)}} \leq x_{(3)}$.*

The price change, $p_1 - p_0$, depends on the realization of $p_0 = \theta_m$ as follows:

- *If $\theta_m = +\sigma$, then*

$$p_1 - p_0 = \begin{cases} \frac{1}{3}\sigma & \textit{for} \quad -x_{(3)} + 1 < \omega_1 \leq y_{(3)} + 1 \\ 0 & \textit{for} \quad y_{(3)} - 1 \leq \omega_1 \leq -x_{(3)} + 1 \\ -\sigma & \textit{for} \quad -x_{(3)} - 1 \leq \omega_1 < y_{(3)} - 1 \end{cases} \tag{3}$$

- *If $\theta_m = -\sigma$, then*

$$p_1 - p_0 = \begin{cases} +\sigma & \textit{for} \quad -y_{(3)} + 1 < \omega_1 \leq x_{(3)} + 1 \\ 0 & \textit{for} \quad x_{(3)} - 1 \leq \omega_1 \leq -y_{(3)} + 1 \\ -\frac{1}{3}\sigma & \textit{for} \quad -y_{(3)} - 1 \leq \omega_1 < x_{(3)} - 1 \end{cases} \tag{4}$$

Proof. See Appendix A.3. □

Both $x_{(3)}$ and $y_{(3)}$ are decreasing in γ and σ; that is, as in the previous cases, speculators trade more cautiously when they are more risk averse or when the stock's payoff is more uncertain. Moreover, since $x_{(3)} > y_{(3)}$ in equilibrium, investors trade more aggressively on their information when they are sure that it is novel as compared to the case when they are worried that market makers could have seen it first.

2.4 Comparison of equilibria and predictions

We now compare the equilibrium outcomes for the different degrees of common knowledge that we analyzed previously. For ease of reference, we refer to these equilibria as follows: the case with m and s being common knowledge (Section 2.2) is denoted by the subscript (1), the case with only m being common knowledge (Section 2.3) with subscript (2), and the case of interest with neither m nor s being common knowledge (Section 2.4) with subscript (3).

We begin by comparing speculators' trading aggressiveness across equilibria.

Corollary 5. *The trading aggressiveness of speculators S in equilibria (1), (2), and (3) compares as follows:* $0 < y_{(3)} < x_{(1)} < x_{(3)} < x_{(2)} < 1$.[5]

Speculators are most aggressive in case (2) when $m \neq s$; that is, when they have an information advantage but when market makers don't know this. Intuitively, for market makers the order flow appears less informative since unconditionally there is a 50% chance that it is pure noise (when $m = s$ so that speculators have no information advantage). The speculators respond to this by trading more aggressively when they do have an information advantage (i.e., when $m \neq s$).

With uncertainty about what's in the price (case (3)) and when $\theta_s \neq \theta_m$, speculators understand that they have an information advantage (i.e., that it must be $m \neq s$) and thus trade almost as aggressively as in case (2).[6] When $\theta_s = \theta_m$, speculators are

[5] The last inequality is due to Assumption 1.

[6] They trade slightly less aggressively because the equilibrium price function in case (3) entails a larger price impact compared to the one in case (2).

unsure whether their information is novel (i.e., $m \neq s$) or stale (i.e., $m = s$) and therefore trade less aggressively. Lastly, when both m and s are common knowledge (case (1)), speculators' trading aggressiveness when $m \neq s$ lies between the ones for $\theta_s \neq \theta_m$ and $\theta_s = \theta_m$ with uncertainty about what's in the price.

Next, we compare the price functions that obtain in the three equilibria.

Corollary 6. *In equilibria (1) and (2), price impact costs for buys and sells are symmetric and do not depend on the previous price update. In equilibrium (3)—i.e., under uncertainty about what's in the price—price impact costs are asymmetric and depend on the previous price update: after a price increase (decrease), price impact costs for buys (sells) are reduced, while those for sells (buys) are increased.*

This corollary highlights the key distinguishing feature of the model with uncertainty about what's in the price: price impact costs differ for buys and sells as a function of past price movements, as illustrated in Figure 2. Intuitively, when buying (selling) volume follows after a recent price uptick (downtick), market makers assign a positive probability to the possibility that speculators are trading on stale news and therefore charge a lower price impact. This property forms the basis for our empirical tests below. Indeed, by checking how price impact costs for buys and sells relate to past returns, we are able to assess the empirical relevance of uncertainty about what's in the price.

The corollary also paves the way for cross-sectional predictions. Indeed, in our model, whether there is uncertainty about what's in the price and hence whether price impact costs are asymmetric for buys and sells depends on whether speculators at $t = 1$ understand what information is already capitalized in the price and what is not (i.e., it depends on whether m is common knowledge). We hypothesize that, at certain times and for certain stocks, it should be easier for investors to understand what information is already reflected in the price; in those instances, the asymmetry for price impact costs should then be weaker. For instance, immediately after earnings announcements, investors should understand that recent price movements are driven by the (public) earnings news and they should find it easier to assess whether a given piece of information they possess is already priced in. We therefore posit that uncertainty about what's in the price, and the associated asymmetric pattern in price impact, should be weaker after earnings announcements. In similar vein, it seems natural that large stocks and stocks with high analyst coverage have more transparent prices and thus should exhibit a lower asymmetry in price impact costs.

Finally, we examine the price informativeness for the three different equilibria. We define price informativeness as $PI \equiv Var\left(E\left(\theta|p_1, p_0\right)\right)$.[7] The higher this measure, the more information prices contain, which lowers the residual uncertainty faced by investors and—to the extent that prices convey information to real decision makers (see e.g. Luo, 2005; Chen et al., 2007; Foucault and Fresard, 2012; Dessaint et al., 2018)—promotes real efficiency.

Corollary 7. *The price informativeness in equilibria (1), (2), and (3) is as follows:*

$$\begin{aligned} PI_{(1)} &= \sigma^2\left(1+\frac{1}{2}x_{(1)}\right) \\ PI_{(2)} &= \sigma^2\left(1+\frac{1}{3}x_{(2)}\right) \\ PI_{(3)} &= \sigma^2\left(1+\frac{1}{6}\left(x_{(3)}+y_{(3)}\right)\right) \end{aligned}$$

Moreover, we have $PI_{(3)} < PI_{(2)}$ and $PI_{(3)} < PI_{(1)}$ (whereas the comparison between $PI_{(1)}$ and $PI_{(2)}$ depends on the parameters).

The corollary shows that uncertainty about what's in the price unambiguously reduces price informativeness. There are two opposing effects that bear on price informativeness. On the one hand, when speculators are worried about whether their signal is stale, they trade less aggressively and thus impound less information into the price. On the other hand, compared to the case in which both speculators and market makers know s, speculators trade slightly more aggressively when they are sure that their signal is novel (i.e., when the signal goes against the most recent price change). This sec-

[7] Alternatively, price informativeness can be defined as $E\left(Var\left(\theta|p_1, p_0\right)\right)$. The definitions are equivalent and related through the Law of Total Variance: $E\left(Var\left(\theta|p_1, p_0\right)\right) = Var\left(\theta\right) - Var\left(E\left(\theta|p_1, p_0\right)\right) = 2\sigma^2 - Var\left(E\left(\theta|p_1, p_0\right)\right)$.

ond effect is indirect and comes from lower price impact costs since—with uncertainty about what's in the price—market makers expect a less informative order flow on average. Overall, the direct effect outweighs the indirect one and so price informativeness decreases.

2.5 Discussion of model assumptions

Our model is deliberately kept as simple as possible. Nonetheless, we conjecture that the main intution is robust to a number of assumptions.

First, we have assumed that market makers observe a part of the fundamental and set p_0 equal to their signal. This is just a convenient short cut. A more elaborate model would have different groups of speculators observing different or the same signals, and trading at different points in time. Such a model yields similar insights. The price at $t = 0$ then reflects the signals of speculators trading in that period. As in our model, speculators arriving at $t = 1$ would then compare their signal realizations with p_0 in order to assess whether other speculators have already traded on the same signal before them.

Regarding the model's distributional assumptions, the two pieces of the fundamental value, θ_1 and θ_2, are assumed to follow binary distributions. This renders speculators' inference particulary simple: when their signal is "high" and the price is "low," speculators infer that the signal must be novel; when speculators' signal is "high" and the price is "high" as well, then speculators are unsure about whether their signal is novel or stale. This intuition carries over naturally to continuous random variables provided we add some noise to the price. To see this, suppose that θ_1 and θ_2 are drawn from continuous distributions and that the price at $t = 0$ reflects market makers' signal with noise.[8] This noise might stem from market makers observing a noisy signal of θ_s or trading for reasons unrelated to the stock's fundamentals such as inventory concerns.[9] As before,

[8]Without noise, when $m \neq s$ and with continuous distributions, $\theta_m = \theta_s$ is a zero-probability event, implying that speculators at $t = 1$ know almost surely whether their signal is novel or stale (i.e., uncertainty about what's in the price disappears).

[9]Alternatively, and as noted above, noise in the $t = 0$ price could come from another group of

speculators arriving at $t = 1$ compare their signal with p_0. If the distributions from which θ_1, θ_2, and noise n are drawn satisfy the *monotone likelihood ratio property* (as is the case for example with normal distributions), then speculators' inference depends monotonically on the *distance* between θ_s and p_0: the larger this distance, the more likely it is that their signal is novel. Our key model prediction about asymmetric price impact costs is expected to go through in this setup.

Finally, note that our assumption about speculators having mean-variance preferences can be replaced by speculators being risk neutral but facing position limits. In that case, uncertainty about what's in the price again causes price impact to be asymmetric across buys and sells depending on θ_m (as in Proposition 3). The only difference with respect to our current setup is that speculators' trading aggressiveness is no longer asymmetric but dictated by the position limit.

3 How important is uncertainty about what's in the price?

In this section, we offer a first test to assess the empirical relevance of uncertainty about what's in the price (UWIP).

3.1 Data and methodology

Our sample comprises the union of the CRSP and TAQ databases for the 1993-2014 period. Throughout our analyses, we focus on common stocks (share codes 10 or 11) and exclude penny stocks (closing price < \$1). With regard to the TAQ data, we apply the filters and adjustments described by Holden and Jacobsen (2014) for dealing with withdrawn or canceled quotes, and we use their interpolated time technique to improve the accuracy of mid-quote prices.

We sign all TAQ trades using the Lee and Ready (1991) algorithm. To obtain signed

speculators trading with noise traders at $t = 0$.

dollar volumes, we multiply the number of shares traded with the prevailing mid-quote at the end of the 5-minute interval containing the trade. We then sum over all signed dollar volumes to obtain the daily trade imbalance, which captures the net buying or selling activity by liquidity consumers (i.e., market order users) on a given date.

3.1.1 Price impact measures

We employ four different price impact measures that are designed to capture adverse selection risk (as faced by the market makers in our model). The first three make use of TAQ data, and the last only requires CRSP. Our first measure is a signed version of the Amihud (2002) illiquidity ratio. Specifically, we define the *price impact costs* for stock i on date t as

$$\text{price impact}_{it} = \frac{\text{return}_{it}}{\text{trade imbalance}_{it}},$$

where the return in the numerator is adjusted for the autocorrelation in daily returns.[10] The difference with the Amihud ratio is that we use the signed trade imbalance rather than trading volume in the denominator, and accordingly also use signed returns in the numerator. This choice is motivated by our model, in which market makers set prices after observing the net order flow (i.e., the trade imbalance). Intuitively, our measure captures by how much the stock price increases (decreases) for one dollar of buying (selling) volume, with higher values indicating higher price impact costs.

Our second adverse selection measure, *lambda*, is the slope coefficient from a regression of stock returns on signed order flow over five-minute intervals; it can be interpreted as the cost of demanding a certain amount of liquidity over five minutes. The third measure is *quote-based price impact*, defined as the percentage change in the mid-quote from before to five minutes after the transaction. Our last measure, *Ln(Amihud)*, is the standard Amihud (2002) illiquidity ratio, defined as the logarithm of the stock's

[10] We explain below why adjust for autocorrelation in our definition. If we use raw returns instead, we find even stronger results (available upon request). The autocorrelation adjustment is done as follows: for a daily return of stock i in some month $\tau(t)$, we run a regression of stock i's return on its lagged return over the previous twelve month ($\tau - 12$ to $\tau - 1$) and record the autocorrelation coefficient $\hat{\beta}_{i\tau(t)}$. To return adjusted for autocorrelation is defined as $\text{return}_{it} - \hat{\beta}_{i\tau(t)} \times \text{return}_{it-1}$.

absolute return divided by its dollar volume.[11] Goyenko et al. (2009) show that it does a good job of capturing adverse selection. We winsorize all price impact measures at the 1% level on both sides.

3.1.2 Methodology

Our model's key prediction is that price impact costs for buys and sells depend on prior returns in opposite ways: whereas the price impact of buys should decrease when prior returns were positive (since these buys are more likely to have been triggered by stale news), the price impact of sells should increase (since these sells are then more likely to have been triggered by novel information). The model is agnostic about the horizon over which returns should be measured; they may be measured intra-day, over one day, or over multiple days. Accordingly, we consider in our empirical tests time windows ranging from one to ten trading days (to which we refer as the "lookback window").[12]

Specifically, we run the following regression separately for days with positive and negative net-buying activity:

$$Y_{it} = \alpha_{i\tau} + \alpha_t + \beta^b \, \text{past return}_{it} + \gamma \, X_{it-1} + \epsilon_{it} \quad \text{if} \quad \text{trade imbalance}_{it} > 0$$
$$Y_{it} = \alpha_{i\tau} + \alpha_t + \beta^s \, \text{past return}_{it} + \gamma \, X_{it-1} + \epsilon_{it} \quad \text{if} \quad \text{trade imbalance}_{it} < 0$$

where Y_{it} is one of our four price impact proxies measured for stock i on trading day t, $\alpha_{i\tau}$ and α_t are stock-month and day fixed effects, past return$_{it}$ is stock i's return on the prior trading day $(t-1)$ or cumulated over the previous five trading days $(t-5$ to $t-1)$ or ten trading days $(t-10$ to $t-1)$, and X_{it-1} is a vector of controls, which includes past turnover, past bid-ask spread, past squared return (as a proxy for volatility), and the natural logarithm of market capitalization. Our theory predicts $\beta^b < 0$ and $\beta^s > 0$. Note that, because our model is saturated with high-dimensional fixed effects

[11]Because this ratio can be zero, we add a small constant (0.00000001) before taking logs. The constant is chosen so as to make the Amihud ratio's distribution closer to a normal. Our results are robust to alternative choices for this constant, including dropping it altogether.

[12]Intra-day regressions over thousands of stocks and multiple years constitute a significant computational burden, so we leave this avenue for future research.

(in particular, stock-month fixed effects), we are controlling for stock-specific variations in illiquidity costs. Our identification thus comes from the difference between buy- and sell-days in the incremental effect of past returns on price impact costs.

When the past return is simply the stock's prior-day return, our approach is potentially confounded by the negative autocorrelation of returns (reversals) observed in individual stock return data. Indeed, a negative return yesterday predicts a positive return today, which enters the numerator of the first of our price impact measures. Since the denominator of this measure is by definition positive (negative) in the sample of days with positive (negative) trade imbalance, this alone could explain why one may find $\beta^b < 0$ and $\beta^s > 0$. This is the reason why we used autocorrelation-adjusted returns in the construction of this price impact measure.

In our baseline, we carry out our tests with raw returns. For robustness, we also show results for CAPM alphas, Fama-French 3-factor alphas, Carhart 4-factor alphas, and Fama-French 5-factor alphas.[13]

3.1.3 Descriptive statistics

Table 1 reports summary statistics for our price impact measures and lagged stock returns—averaged over all days and separately for days with net-buying or net-selling activity. For better visibility, *price impact costs* are scaled by 10^6. This implies that the price impact measures can be interpreted as representing the return change that is triggered by a one million USD net order flow. For instance, based on the statistics for the overall sample, a one million USD net buy would be expected to push up the price by 0.64%. For better visibility, the *lambda* measure is scaled by 10^2. The table further shows that, for three out of four price impact measures, the price impact is on average slightly higher on days with net selling activity, as compared to days with net buying

[13]For each factor model (CAPM, Fama-French 3-factor, Carhart 4-factor, Fama-French 5-factor), we estimate alphas as follows. First, for stock i at date t in month $\tau(t)$, we estimate factor loadings using stock i's returns and factor returns (taken from Kenneth French's website) over months $\tau - 12$ to $\tau - 1$. Second, we use the factor loadings together with the factor returns on date t to compute the alpha, $\alpha_{it} = \text{ret}_{it} - \sum_j \hat{\beta}_j r_{jt}$, where the sum is always taken over the factors pertaining to a certain factor model.

activity.

3.2 Baseline results

Table 2 Panel (a) shows the results based on lagged one-day returns as the key independent variable. For three out of the four price impact measures (*price impact costs*, *lambda*, and *ln(Amihud)*) we obtain results in line with the model's prediction: on days with a positive trade imbalance (net buys), price impact costs are significantly negatively related to the prior-day return (columns 1, 3, 5, and 7); while on days with a negative trade imbalance (net sells), they are significantly positively related to the prior-day return (columns 2, 4, and 8). The only exception is the *quote-based price impact*, for which the regression coefficient on the lagged return for net sells is negative (column 6), but of a smaller absolute magnitude and statistical significance than the one for net buys (column 5). The economic magnitude of our findings is meaningful. For instance, a one-standard deviation increase in the lagged return decreases (increases) the *price impact costs* on days with a positive (negative) net trade imbalance by about 5% of its standard deviation, thus driving a wedge between the *price impact costs* on buy- and sell-days of about 10% of its standard deviation.

Panels (b) and (c) show results employing longer lookback windows; that is, based on regressing price impact measures on lagged returns cumulated over either the preceeding five (Panel b) or ten trading days (Panel c). In those regressions, price impact remains measured on day t, even though returns are now measured over a window of several days, and the past turnover and past bid-ask spread control variables are then also averaged over five or ten days, respectively. In both panels, results are strongly aligned with our prediction: for all four price impact measures, price impact costs are negatively related to past returns on net-buy days, while being positively related to past returns on net-sell days. The economic magnitude is, if anything, increased. For instance, over the 5-day window, a one-standard deviation increase in the cumulated lagged return drives a wedge between the *price impact costs* on buy- and sell-days of about 12% of its standard deviation. The fact that the results are stronger than those for the lagged one-day return (Panel a) suggests that uncertainty about what's in the price is not only a short-term concern for market participants.

Table 3 presents robustness checks for our price impact tests. First, our results are robust to excluding stock-level controls from the baseline regression. Regardless of the lookback window, price impact costs remain asymmetrically related to past returns for buys and sells. In this specification, even the results for the *quote-based price impact measure* over the one-day window (which displayed the wrong sign in the baseline) work according to our theory. Notwithstanding, we use as baseline the specification with controls because it is more conservative. Second, our findings are robust to using various measures of abnormal returns instead of raw returns as key independent variables. Regardless of the factor model employed (CAPM, Fama-French 3-factor, Carhart 4-factor, Fama-French 5-factor) and of the length of the lookback window, price impact costs are lower (higher) for net buys (net sells) when past alphas are larger (the only exception occurs again for the *quote-based price impact* measure for the lagged one-day window). All in all, our baseline empirical findings align well with the model's key prediction, and suggest that uncertainty about what's in the price is an economically relevant phenomenon.

3.3 Cross-sectional results

In this subsection, we conduct powerful auxiliary tests of our theory. If, as we argue, the asymmetric price impact pattern for buy- and sell-days is caused by uncertainty about what's in the price, then it should be less pronounced i) at times when there is less such uncertainty, and ii) for stocks that display less such uncertainty.

Our first test tracks stocks over time and investigates whether the asymmetric price impact pattern weakens immediately after earnings announcements when investors know better what information is already reflected in stock prices. To implement this test, we amend our regressions for buy and sell days as follows:

$$\begin{aligned} Y_{it} &= \alpha_{i\tau} + \alpha_t + \beta_1^b \,\text{past return}_{it} + \beta_2^b \,\text{EA}_{it} + \beta_3^b \,\text{past return}_{it} \times \text{EA}_{it} + \gamma\, X_{it-1} + \epsilon_{it} \\ &\qquad\qquad \text{if trade imbalance}_{it} > 0 \\ Y_{it} &= \alpha_{i\tau} + \alpha_t + \beta_1^s \,\text{past return}_{it} + \beta_2^s \,\text{EA}_{it} + \beta_3^s \,\text{past return}_{it} \times \text{EA}_{it} + \gamma\, X_{it-1} + \epsilon_{it} \\ &\qquad\qquad \text{if trade imbalance}_{it} < 0 \end{aligned}$$

where EA_{it} is a dummy variable that takes the value of one when stock i on date t announced its earnings in the past one, five, or ten trading days depending on the lookback window.[14] The variable $\text{past return}_{it} \times \text{EA}_{it}$ denotes the interaction of the earnings announcement dummy with the past return. All other variables and fixed effects are as in our baseline regression. Based on our theory, we expect $\beta_1^b < 0$ and $\beta_1^s > 0$, but $\beta_3^b > 0$ and $\beta_3^s < 0$. In words, on buy-days (sell-days), price impact should be negatively (positively) related to past returns, but this relation should be weakened just after earnings announcements.

Table 4 provides strong support for this prediction. With only one exception (quote-based price impact for the one-day lookback period), price impact reacts asymmetrically to past returns on buy- and sell-days, but this asymmetric reaction is strongly muted after earnings announcements. Indeed, β_3 has consistently the opposite sign of β_1 and has a magnitude that, while lower than β_1, remains important. For instance, for *price impact costs* and a one-day lookback period (Panel a, columns 1 and 2), the results indicate that the initial effect of past returns on price impact is about 70% lower when there was an earnings announcement on the previous day.

Our second set of tests exploits variations across stocks (and over time) in the degree of uncertainty about what's in the price. Here, we argue that investors face less such uncertainty about stocks with more public scruteny; i.e., larger stocks and stocks with more analysts. The idea is that such stocks have a stronger public information

[14] We take earnings announcement dates from I/B/E/S data. Accordingly, we run this test only for stocks with I/B/E/S data.

environment, implying that the scope for information asymmetry and thus uncertainty about what's in the price is reduced. The tests are similar to the preceeding ones, except that we now interact past returns with market capitalization and analyst coverage, instead of an earnings announcement dummy.[15] The findings are reported, respectively, in Tables 5 and 6 for market capitalization and analyst coverage. They lend again support to our mechanism: the asymmetry in price impact for buy- and sell-days is reduced for larger stocks and for stocks with more analyst coverage. This is consistent with the idea that uncertainty about what's in the price is lower for stocks that receive more attention (from investors and analysts).

Overall, the results in this section strongly suggest that uncertainty about what's in the price is a real concern for investors. Indeed, we are not aware of any other theory that could explain why price impact costs for buys and sells are asymmetrically related to past price movements, and why this asymmetry should be consistently reduced after public earnings announcements or for stocks that investors follow more closely.

4 UWIP and price informativeness

In this section, we investigate how uncertainty about what's in the price (UWIP) relates to stock price informativeness.

4.1 Data and methodology

Our model predicts that uncertainty about what's in the price is associated with less informative stock prices. We test this prediction in the context of earnings announcements. Specifically, we measure the price jump around earnings announcement dates, and regress it on a self-constructed proxy for the extent of uncertainty about what's in the price.

We follow Weller (2018) and construct the price jump ratio as the fraction of

[15]In both tables, the level effects of market capitalization and analyst coverage are difficult to interpret as our regression specification continues to have stock-month fixed effects.

earnings-related information that is incorporated into the stock price *prior* to an earnings announcement. The intuition behind this measure follows straight from models of informed trading (e.g., Kyle, 1985; Back, 1992): the price drifts toward the post-announcement asset value ahead of the announcement as investors trade on their private information. Competition among informed traders accelerates this process, resulting in even more information impounded into prices before the announcement (Holden and Subrahmanyam, 1992). As such, the price jump ratio is a direct measure of the information content of stock prices (Weller, 2018). In contrast, widely used measures such as pricing error variance (Hasbrouck, 1993) or variance ratio tests (Lo and MacKinlay, 1988) only measure price efficiency (i.e., whether stock prices follow a random walk and thus accurately reflect available public information) and are therefore not suitable for our purpose.

We construct the price jump ratio as described in Weller (2018). Here, we provide a brief summary of his approach and refer the reader to his paper for more detail. We start from the sample of quarterly earnings announcements over the years 1995 to 2014. We estimate abnormal returns relative to the Fama and French (1992) three-factor model using daily returns over a 365-calendar day window ending 90 days before the earnings announcement. We retain the estimated factor loadings if at least 63 non-missing return observations are available in the estimation window. Abnormal returns around earnings announcements are then cumulated in event-time. Finally, the price jump ratio for stock i and event date t is defined as:

$$\text{jump}_{it} = \frac{CAR_{it}^{(T-1,T+2)}}{CAR_{it}^{(T-21,T+2)}}$$

A high price jump ratio corresponds to a large announcement-date jump relative to the pre-announcement drift and thus indicates a *low* level of price informativeness. As explained by Weller (2018), the price jump ratio is only meaningful for announcements with a sufficiently large information content. We therefore only retain announcement events that satisfy

$$\left|CAR_{it}^{(T-21,T+2)}\right| > \sqrt{24}\hat{\sigma}_{it}$$

where $\hat{\sigma}_{it}$ is the stock's daily return volatility calculated over trading days $T-42$ to $T-22$. In our final sample, the price jump ratio has a mean of 35%, suggesting that, for the average announcement event, a significant fraction of the information enters prices before the announcement date. This figure is in line with what is reported in Weller (2018).

Our measure of uncertainty about what's in the price builds on the key intuition of our model: the more unsure speculators are about what's in the price, the more asymmetrically price impact for buys and sells responds to past returns. To capture this effect, we employ the price impact measure that yielded the strongest results in Table 3—*price impact costs*—and estimate for each announcement event the following interaction specification using daily data over the calendar quarter prior to the earnings announcement date:

$$\text{price impact costs}_{i\tau} = \beta_0 + \beta_1 \, \text{return}_{i\tau-1} + \beta_2 \, \text{netsell}_{i\tau} + \beta_3 \, \text{return}_{i\tau-1} \times \text{netsell}_{i\tau} + \epsilon_{i\tau}$$

where $\text{netsell}_{i\tau}$ is a dummy variable that takes a value of one if stock i on date τ exhibited a negative trade imbalance (if $\text{trade imbalance}_{i\tau} < 0$) and zero otherwise. The coefficient of interest, β_3, captures the difference in the sensitivity of price impact for sells and buys with respect to the prior-day return. Our model predicts that $\beta_3 > 0$ and hence a higher coefficient estimate indicates a larger uncertainty about what's in the price. We thus define $\text{uwip}_{it}^{raw} = \hat{\beta}_3$. To mitigate the effect of outliers, we winsorize uwip_{it}^{raw} (as well as all other variables used in our regression below) at the 1% level on both sides. For robustness, we define uwip_{it}^{dec} as the decile rank of uwip_{it}^{raw}; i.e., uwip_{it}^{raw} takes on values from one to ten depending on the corresponding decile of $\hat{\beta}_3$.

For our price informativeness tests, we regress the price jump ratio on our measure of the uncertainty about what's in the price, as well as on a host of controls:

$$\text{jump}_{it} = \alpha_i + \alpha_t + \beta\,\text{uwip}_{it} + \gamma\,X_{it-1} + \epsilon_{it}$$

where α_i and α_t are stock and date fixed effects, respectively, and X_{it-1} is a vector of pre-determined control variables comprising market capitalization, analyst coverage, past returns, volatility, turnover, and bid-ask spreads (all control variables are defined in the header of Table 7 below).

4.2 Test results

Table 7 Panel (a) shows the results for the raw measure of uncertainty about what's in the price, uwip^{raw}. Looking at column (1), we find a significantly positive effect of uwip^{raw} on the price jump ratio, implying that uncertainty about what's in the price is associated with a lower stock price informativeness as predicted by our model. In terms of economic magnitude, a two-standard deviation increase (≈ 0.0001) in uwip^{raw} increases the price jump ratio by about 1 percentage point, or about 3% relative to its unconditional mean (35%) in our sample. Subsequently adding stock market controls or finer fixed effects does not alter this picture: the coefficient estimate for uwip^{raw} barely changes and always remains statistically significant at the 1% level.

Panel (b) shows that these findings are robust to using decile ranks of uwip^{raw}, denoted uwip^{dec} ; i.e., they are not driven by outliers. Here, the magnitude implies that moving from the 1st to the 10th decile of uncertainty about what's in the price increases the price jump ratio by about 1.4% percentage points, or about 4% relative to its unconditional sample mean. Overall, these results confirm our prediction that uncertainty about what's in the price slows down the incorporation of fundamental information into prices and hurts stock price informativeness.

5 Conclusion

This paper proposes a simple model in which speculators are unsure whether a given signal they observe is stale (i.e., already reflected in the price) or novel—and thus valuable to trade on. In equilibrium, speculators assess the novelty of their signal by comparing it to the most recent price movement and adjust the trading aggressiveness accordingly. Market makers, in turn, anticipate that speculators may be trading on stale news. The resulting price function is inherently asymmetric: after price increases (decreases), market makers consider incoming buy volume to be less (more) informative and thus charge a lower (higher) price impact compared to sell volume. Moreover, by making speculators reluctant to trade, uncertainty about what's in the price decreases stock price informativeness.

Using daily order flow data for a comprehensive panel of NYSE-traded stocks, we find strong support for the prediction that price impact costs depend on past returns. Specifically, we document that on days with a positive (negative) trade imbalance, price impact costs are negatively (positively) related to past stock returns (over the prior day or cumulated over the previous five or ten trading days). Moreover, we find that this asymmetry in price impact costs between buy- and sell-days is reduced after earnings announcements and for stocks with a large market capitalization and a high analyst coverage; i.e.,when speculators should arguably have a better sense about whether their private signals are novel or stale. Overall, our results strongly suggest that uncertainty about what's in the price is a common and widespread concern for stock market participants.

References

Abreu, D., Brunnermeier, M.K., 2002. Synchronization risk and delayed arbitrage. Journal of Financial Economics 66, 341–360.

Abreu, D., Brunnermeier, M.K., 2003. Bubbles and crashes. Econometrica 71, 173–204.

Amihud, Y., 2002. Illiquidity and stock returns. Journal of Financial Markets 5, 31–56.

Back, K., 1992. Insider trading in continuous time. Review of Financial Studies 5, 387–409.

Banerjee, S., Green, B., 2015. Signal or noise? Uncertainty and learning about whether other traders are informed. Journal of Financial Economics 117, 398–423.

Brown, D.P., Jennings, R.H., 1989. On technical analysis. Review of Financial Studies 2, 527–551.

Brunnermeier, M.K., 2005. Information leakage and information efficiency. Review of Financial Studies 18, 417–457.

Chen, Q., Goldstein, I., Jiang, W., 2007. Price informativeness and investment sensitivity to stock price. Review of Financial Studies 20, 619–650.

Dessaint, O., Foucault, T., Fresard, L., Matray, A., 2018. Noisy stock prices and corporate investment. Working Paper.

Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment. Journal of Finance 47, 577–605.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. Journal of Finance 47, 427–465.

Foucault, T., Fresard, L., 2012. Cross-listing, investment sensitivity to stock price, and the learning hypothesis. Review of Financial Studies 25, 3305–3350.

Goyenko, R.Y., Holden, C.W., Trzcinka, C.A., 2009. Do liquidity measures measure liquidity? Journal of Financial Economics 153-181.

Grossman, S., 1976. On the efficiency of competitive stock markets where trades have diverse information. Journal of Finance 31, 573–585.

Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. American Economic Review 70, 393–408.

Grundy, B.D., McNichols, M., 1989. Trade and the revelation of information through prices and direct disclosure. Review of Financial Studies 2, 495–526.

Hasbrouck, J., 1993. Assessing the quality of a security market: A new approach to transaction-cost measurement. Review of Financial Studies 6, 191–212.

Hellwig, M.F., 1980. On the aggregation of information in competitive markets. Journal of Economic Theory 22, 477–498.

Hendershott, T., Menkveld, A.J., 2014. Price pressures. Journal of Financial Economics 114, 405–423.

Ho, T., Stoll, H., 1981. Optimal dealer pricing under transaction and return uncertainty. Journal of Financial Economics 9, 47–73.

Holden, C.W., Jacobsen, S., 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. Journal of Finance 69, 1747–1785.

Holden, C.W., Subrahmanyam, A., 1992. Long-lived private information and imperfect competition. Journal of Finance 47, 247–270.

Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1461–1484.

Lee, C.M.C., Ready, M.J., 1991. Inferring trade dirdirect from intraday data. Journal of Finance 46, 733–746.

Llorente, G., Michaely, R., Saar, G., Wang, J., 2002. Dynamic volume-return relation of individual stocks. Review of Financial Studies 15, 1005–1047.

Lo, A.W., MacKinlay, A.C., 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Luo, Y., 2005. Do insiders learn from outsiders? Evidence from mergers and acquisitions. Journal of Finance 60, 1951–1982.

Madhavan, A., Smidt, S., 1993. An analysis of cchange in specialist invinvestor and quotations. Journal of Finance 48, 1595–1628.

Tetlock, P.C., 2011. All the news that's fit to reprint: Do investors react to stale information? Review of Financial Studies 24, 1481–1512.

Treynor, J.L., Ferguson, R., 1985. In defense of technical analysis. Journal of Finance 40, 757–775.

Vives, X., 1995. Short-term investment and the informational efficiency of the market. Review of Financial Studies 8, 125–170.

Weller, B.M., 2018. Does algorithmic trading reduce information acquisition? Review of Financial Studies 31, 2184–2226.

Figure 1: Model setup

This figure summarizes the model setup. At t=0, market makers M observe θ_m, where $m \in \{1, 2\}$ with equal probability, and set $p_0 = \theta_m$. At t=1, speculators S observe θ_s, where $s \in \{1, 2\}$ with equal probability, and submit market order to maximize expected utility. Market makers observe the net order flow, consisting of the sum of speculators' market orders and noise trades, and set $p_1 = E(\theta|\theta_m, \omega)$. At t=2, the stock's payoff $\theta = \theta_1 + \theta_2$ is realized and consumption takes place.

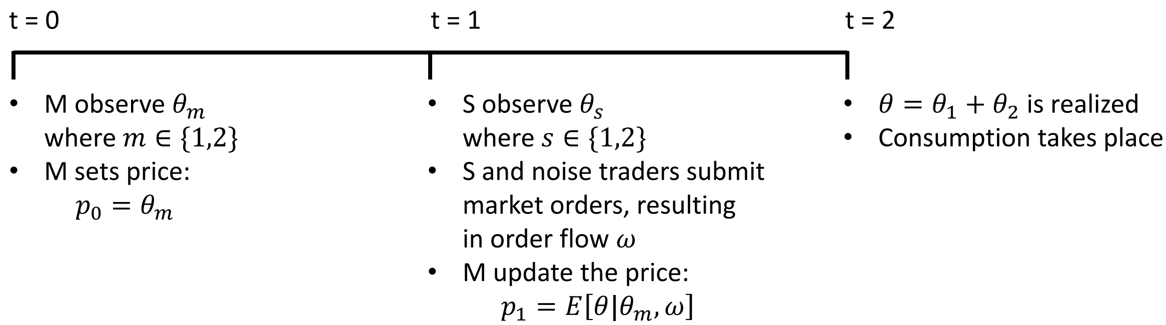
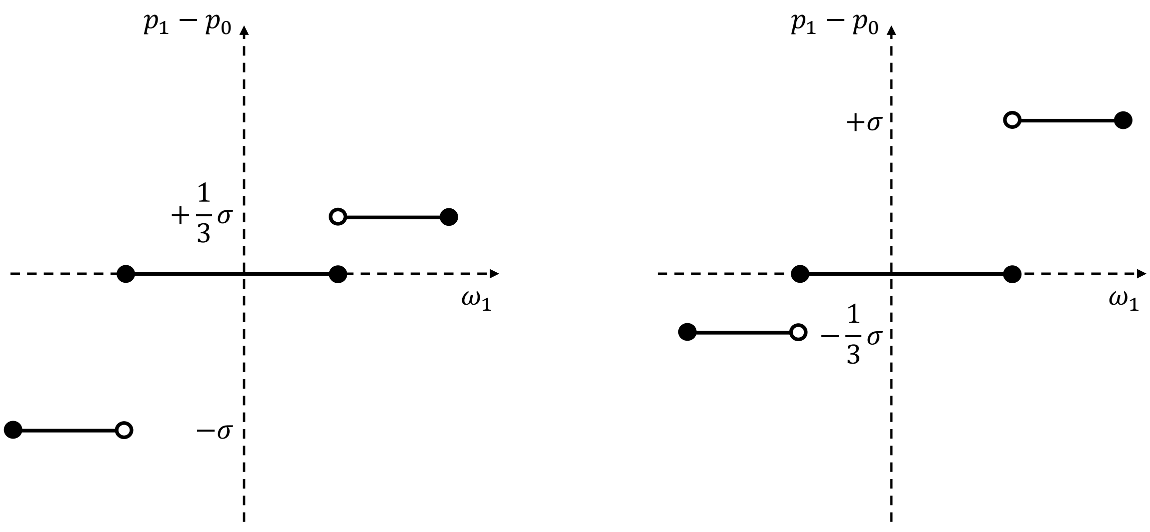


Figure 2: Equilibrium price function

This figure shows the equilibrium price function when speculators face uncertainty about whether their trading signal is already in the price. In Panel A, we show the price function for the case of prior positive news ($\theta_m = +\sigma$). In Panel B, we show the price function for prior negative news ($\theta_m = -\sigma$).



Panel A: For $\theta_m = +\sigma$ **Panel B**: For $\theta_m = -\sigma$

Table 1: **Price Impact Statistics**

This table reports descriptive statistics for our four price impact measures in the overall (stock-day) sample, as well as separately for days with positive and negative net trade imbalance. *Price impact costs* is defined as the ratio of the (autocorrelation-adjusted) return over the net trade imbalance. The measure is multiplied by 10^6 for better visibility. *Lambda* is defined as the slope coefficient of regressing stock returns on signed order flow over five-minute intervals. *Quote-based price impact* is defined as the dollar-weighted average of the percentage change in the mid-quote from right before the transaction to five minutes after the transaction. *Ln(Amihud)* is the Amihud (2002) illiquidity ratio, defined as the logarithm of (a small constant plus) the ratio of absolute return over dollar volume. In our sample, lagged one-day returns have a standard deviation of 3.5%, lagged cumulated five-day returns have a standard deviation of 7.5%, and lagged cumulated ten-day returns have a standard deviation of 10.5%. All price impact measures and lagged returns are winsorized at the 1% level on both sides.

	Mean	Median	Standard deviation
Overall sample (N =24,460,914)			
Price impact costs	0.7115	0.0094	6.2189
Lambda	0.0014	0.0002	0.0047
Quote-based price impact	0.0038	0.0013	0.0077
Ln(Amihud)	-16.9221	-17.7881	1.8502
Net Buys (N =11,928,063)			
Price impact costs	0.6411	0.0048	6.2257
Lambda	0.0012	0.0002	0.0045
Quote-based price impact	0.0037	0.0013	0.0074
Ln(Amihud)	-17.0658	-17.9460	1.7813
Net Sells (N =12,532,851)			
Price impact costs	0.7790	0.0172	6.2116
Lambda	0.0015	0.0002	0.0049
Quote-based price impact	0.0040	0.0014	0.0081
Ln(Amihud)	-16.7840	-17.5977	1.9038

Table 2: **Price Impact Tests for Raw Returns**

This table reports the results from regressing our four price impact measures on lagged raw returns as explained in Section 3.1. Panel (a) shows the results for using lagged one-day returns as the key independent variable. Panel (b) shows the results for using lagged cumulated five-day returns as the key independent variable. Panel (c) shows the results for using lagged cumulated ten-day returns as the key independent variable. In each panel, columns (1) and (2) show results for the *price impact costs* measure (scaled by 10^6 for better visibility), columns (3) and (4) show results for the *lambda* measure (scaled by 10^2 for better visibility), columns (5) and (6) show results for the *quote-based price impact* measure, and columns (7) and (8) show results for the *ln(Amihud)* measure. Odd columns show results for the net-buy subsample (i.e., all stock-day observations on which the trade imbalance is positive). Even columns show results for the net-sell subsample (i.e., all stock-day obsvervations on which the trade imbalance is negative). All regressions contain stock-level controls (past turnover, past bid-ask spread, and the natural logarithm of market capitalization), stock-month and date fixed effects. t-statistics are based on standard errors adjusted for double-clustering by stock and day. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Panel (a): Lagged one day

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-9.2697***	8.6408***	-0.0016***	0.0004***	-0.0021***	-0.0002**	-2.1571***	1.9709***
	(-64.58)	(62.21)	(-24.99)	(6.36)	(-21.54)	(-2.09)	(-86.18)	(65.48)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11,177,991	11,704,688	11,535,219	12,131,601	11,523,481	12,096,866	11,571,604	12,173,011
adj. R^2	0.14	0.13	0.33	0.29	0.35	0.29	0.70	0.64

Panel (b): Lagged cumulated five days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-5.5202***	4.7680***	-0.0017***	0.0006***	-0.0030***	0.0010***	-1.2662***	0.6420***
	(-68.22)	(65.18)	(-46.09)	(18.31)	(-52.44)	(19.03)	(-96.96)	(48.50)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11,280,290	11,810,488	11,638,880	12,239,140	11,627,138	12,204,400	11,675,258	12,280,549
adj. R^2	0.14	0.13	0.34	0.30	0.35	0.29	0.69	0.64

Panel (c): Lagged cumulated ten days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-6.3980***	5.5888***	-0.0019***	0.0007***	-0.0036***	0.0015***	-1.2278***	0.4677***
	(-70.74)	(68.79)	(-54.82)	(23.35)	(-67.55)	(29.35)	(-102.00)	(37.20)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	11,310,577	11,839,567	11,669,298	12,268,432	11,657,563	12,233,709	11,705,672	12,309,836
adj. R^2	0.14	0.13	0.34	0.30	0.35	0.29	0.69	0.63

Table 3: **Price Impact Tests – Robustness**

This table reports robustness checks for our price impact tests (see Section 3.1): 1) regressions without stock-level controls and 2) using different factor model alphas (CAPM alphas, Fama-French 3-factor alphas, Carhart 4-factor alphas, Fama-French 5-factor alphas) instead of raw returns. Each line thus corresponds to a different regression. Panel (a) shows the results for using lagged one-day returns or alphas as the key independent variable. Panel (b) shows the results for using lagged cumulated five-day returns or alphas as the key independent variable. Panel (c) shows the results for using lagged cumulated ten-day returns or alphas as the key independent variable. In each panel, columns (1) and (2) show results for the *price impact costs* measure (scaled by 10^6 for better visibility), columns (3) and (4) show results for the *lambda* measure (scaled by 10^2 for better visibility), columns (5) and (6) show results for the *quote-based price impact* measure, and columns (7) and (8) show results for the *ln(Amihud)* measure. Odd columns show results for the net-buy subsample (i.e., all stock-day observations on which the trade imbalance is positive). Even columns show results for the net-sell subsample (i.e., all stock-day observations on which the trade imbalance is negative). All regressions contain stock-month and date fixed effects. t-statistics are based on standard errors adjusted for double-clustering by stock and day. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Panel (a): Lagged one day

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1) Net Buys	(2) Net Sells	(3) Net Buys	(4) Net Sells	(5) Net Buys	(6) Net Sells	(7) Net Buys	(8) Net Sells
1) Without controls								
Raw return	-8.3838***	7.6351***	-0.0016***	0.0004***	-0.0016***	0.0002**	-2.3513***	1.7172***
	(-62.99)	(57.18)	(-23.57)	(7.63)	(-15.64)	(2.20)	(-91.78)	(57.74)
2) Risk-adjusted returns								
CAPM alpha	-8.8037***	8.1512***	-0.0016***	0.0004***	-0.0021***	-0.0002**	-2.0884***	1.8332***
	(-61.62)	(57.38)	(-24.56)	(5.59)	(-20.70)	(-2.30)	(-85.71)	(62.91)
FF 3-factor alpha	-8.8354***	8.1279***	-0.0016***	0.0004***	-0.0021***	-0.0002**	-2.0902***	1.8429***
	(-61.80)	(57.28)	(-24.25)	(5.63)	(-20.85)	(-2.01)	(-85.87)	(63.31)
Car 4-factor alpha	-8.8183***	8.0987***	-0.0016***	0.0004***	-0.0021***	-0.0002*	-2.0830***	1.8407***
	(-61.68)	(57.35)	(-24.19)	(5.62)	(-20.74)	(-1.83)	(-85.79)	(63.38)
FF 5-factor alpha	-8.8278***	8.1013***	-0.0016***	0.0003***	-0.0021***	-0.0002*	-2.0838***	1.8354***
	(-61.95)	(57.45)	(-24.17)	(5.48)	(-20.64)	(-1.90)	(-85.97)	(63.45)

Panel (b): Lagged cumulated five days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1) Net Buys	(2) Net Sells	(3) Net Buys	(4) Net Sells	(5) Net Buys	(6) Net Sells	(7) Net Buys	(8) Net Sells
1) Without controls								
Raw return	-4.8068***	3.4390***	-0.0017***	0.0006***	-0.0026***	0.0009***	-1.3417***	0.3601***
	(-66.48)	(54.66)	(-44.55)	(16.89)	(-48.31)	(18.90)	(-100.36)	(30.24)
2) Risk-adjusted returns								
CAPM alpha	-28.0233***	24.2387***	-0.0089***	0.0030***	-0.0150***	0.0051***	-6.2524***	3.0992***
	(-69.54)	(66.44)	(-46.15)	(18.34)	(-51.77)	(18.85)	(-97.63)	(46.80)
FF 3-factor alpha	-28.0952***	24.1333***	-0.0089***	0.0030***	-0.0149***	0.0051***	-6.2393***	3.1284***
	(-69.66)	(66.44)	(-45.94)	(18.21)	(-52.33)	(19.18)	(-97.81)	(47.39)
Car 4-factor alpha	-27.8903***	24.0123***	-0.0088***	0.0031***	-0.0149***	0.0052***	-6.1916***	3.1257***
	(-69.51)	(66.46)	(-45.53)	(18.33)	(-52.04)	(19.35)	(-97.51)	(47.62)
FF 5-factor alpha	-28.0187***	24.0255***	-0.0089***	0.0030***	-0.0149***	0.0052***	-6.2125***	3.1165***
	(-69.81)	(66.40)	(-45.84)	(18.09)	(-52.02)	(19.21)	(-98.17)	(47.50)

Panel (c): Lagged cumulated ten days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1) Net Buys	(2) Net Sells	(3) Net Buys	(4) Net Sells	(5) Net Buys	(6) Net Sells	(7) Net Buys	(8) Net Sells
1) Without controls								
Raw return	-4.9288***	3.6109***	-0.0019***	0.0007***	-0.0028***	0.0009***	-1.3417***	0.3601***
	(-68.73)	(60.85)	(-51.90)	(17.59)	(-60.94)	(22.53)	(-100.36)	(30.24)
2) Risk-adjusted returns								
CAPM alpha	-64.4518***	56.8810***	-0.0201***	0.0069***	-0.0363***	0.0147***	-12.0725***	4.6218***
	(-71.67)	(70.29)	(-55.14)	(21.68)	(-66.57)	(27.91)	(-102.33)	(36.56)
FF 3-factor alpha	-64.3103***	56.5617***	-0.0200***	0.0070***	-0.0361***	0.0148***	-11.9861***	4.6670***
	(-71.96)	(70.31)	(-54.65)	(21.74)	(-66.55)	(28.25)	(-102.52)	(37.23)
Car 4-factor alpha	-63.8055***	56.1589***	-0.0197***	0.0070***	-0.0358***	0.0147***	-11.8855***	4.6741***
	(-71.92)	(70.51)	(-54.38)	(21.98)	(-66.41)	(28.16)	(-102.48)	(37.51)
FF 5-factor alpha	-63.9482***	56.1799***	-0.0199***	0.0069***	-0.0358***	0.0147***	-11.9049***	4.6639***
	(-72.18)	(70.44)	(-54.52)	(21.81)	(-66.19)	(28.18)	(-102.68)	(37.48)

Table 4: **Price Impact Tests – Interaction with Earnings Announcements**

This table reports the results from regressing our four price impact measures on lagged raw returns, the EA dummy, the interaction of lagged returns and the EA dummy, and controls (unreported). Panel (a) shows the results for using lagged one-day returns as the key independent variable. Panel (b) shows the results for using lagged cumulated five-day returns as the key independent variable. Panel (c) shows the results for using lagged cumulated ten-day returns as the key independent variable. In each panel, columns (1) and (2) show results for the *price impact costs* measure (scaled by 10^6 for better visibility), columns (3) and (4) show results for the *lambda* measure (scaled by 10^2 for better visibility), columns (5) and (6) show results for the *quote-based price impact* measure, and columns (7) and (8) show results for the *ln(Amihud)* measure. Odd columns show results for the net-buy subsample (i.e., all stock-day obsvervations on which the trade imbalance is positive). Even columns show results for the net-sell subsample (i.e., all stock-day obsvervations on which the trade imbalance is negative). All regressions contain stock-level controls (past turnover, past bid-ask spread, the natural logarithm of market capitalization, and analyst coverage), stock-month and date fixed effects. t-statistics are based on standard errors adjusted for double-clustering by stock and day. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Panel (a): Lagged one day

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-6.8566***	6.9046***	-1.6184***	1.3935***	-0.0015***	-0.0005***	-1.6184***	1.3935***
	(-55.11)	(55.16)	(-71.24)	(50.71)	(-16.31)	(-5.11)	(-71.24)	(50.71)
EA	-0.0888***	0.0653***	-0.0110***	0.0038	0.0003***	0.0004***	-0.0110***	0.0038
	(-7.99)	(5.63)	(-5.03)	(1.50)	(20.27)	(23.01)	(-5.03)	(1.50)
Raw return*EA	4.3685***	-5.2129***	0.7205***	-1.1170***	-0.0003	-0.0006	0.7205***	-1.1170***
	(14.36)	(-16.04)	(14.90)	(-19.19)	(-0.97)	(-1.50)	(14.90)	(-19.19)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,398,341	9,419,490	9,671,770	9,730,123	9,638,338	9,678,670	9,671,770	9,730,123
adj. R^2	0.11	0.07	0.70	0.64	0.33	0.27	0.70	0.64

Panel (b): Lagged cumulated five days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-3.9414***	3.6935***	-0.0011***	0.0003***	-0.0022***	0.0007***	-1.0235***	0.4645***
	(-58.23)	(56.99)	(-36.80)	(10.63)	(-43.28)	(13.09)	(-84.17)	(37.78)
EA	-0.0626***	0.0444***	-0.0000***	-0.0000	0.0000***	0.0001***	-0.0239***	-0.0148***
	(-9.72)	(7.28)	(-4.33)	(-0.13)	(5.85)	(9.63)	(-18.65)	(-10.07)
Raw return*EA	1.7745***	-1.7609***	0.0004***	-0.0002***	0.0004***	-0.0003***	0.3545***	-0.2963***
	(20.82)	(-21.31)	(9.01)	(-3.32)	(4.83)	(-3.13)	(25.04)	(-18.57)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,476,649	9,497,732	9,723,237	9,778,537	9,717,304	9,757,464	9,750,818	9,809,069
adj. R^2	0.11	0.07	0.32	0.28	0.34	0.28	0.69	0.64

Panel (c): Lagged cumulated ten days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-4.3698***	4.1250***	-0.0012***	0.0004***	-0.0028***	0.0011***	-1.0129***	0.3401***
	(-60.82)	(60.83)	(-44.13)	(14.37)	(-58.20)	(21.85)	(-88.74)	(28.88)
EA	-0.0364***	0.0316***	-0.0000***	-0.0000***	-0.0000	0.0000	-0.0229***	-0.0184***
	(-6.65)	(5.78)	(-6.74)	(-3.38)	(-1.33)	(0.94)	(-20.87)	(-13.74)
Raw return*EA	1.0422***	-0.9656***	0.0003***	-0.0001***	0.0004***	-0.0003***	0.1973***	-0.1300***
	(18.20)	(-17.81)	(8.79)	(-3.21)	(6.09)	(-3.99)	(20.26)	(-11.65)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,508,375	9,528,948	9,755,062	9,809,779	9,749,103	9,788,661	9,782,680	9,840,365
adj. R^2	0.11	0.07	0.32	0.28	0.34	0.28	0.69	0.64

Table 5: **Price Impact Tests – Interaction with Market Capitalization**

This table reports the results from regressing our four price impact measures on lagged raw returns, the (natural logarithm of) market capitalization, the interaction of lagged returns and market capitalization, and controls (unreported). Panel (a) shows the results for using lagged one-day returns as the key independent variable. Panel (b) shows the results for using lagged cumulated five-day returns as the key independent variable. Panel (c) shows the results for using lagged cumulated ten-day returns as the key independent variable. In each panel, columns (1) and (2) show results for the *price impact costs* measure (scaled by 10^6 for better visibility), columns (3) and (4) show results for the *lambda* measure (scaled by 10^2 for better visibility), columns (5) and (6) show results for the *quote-based price impact* measure, and columns (7) and (8) show results for the *ln(Amihud)* measure. Odd columns show results for the net-buy subsample (i.e., all stock-day observations on which the trade imbalance is positive). Even columns show results for the net-sell subsample (i.e., all stock-day obsvervations on which the trade imbalance is negative). All regressions contain stock-level controls (past turnover, past bid-ask spread, the natural logarithm of market capitalization, and analyst coverage), stock-month and date fixed effects. (Because of the stock-month fixed effect, the coefficient on market capitalization alone should not be interpreted.) t-statistics are based on standard errors adjusted for double-clustering by stock and day. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Panel (a): Lagged one day

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-40.5160***	36.7527***	-0.0056***	0.0008	-0.0039***	-0.0030***	-9.0787***	7.9991***
	(-40.86)	(35.15)	(-11.41)	(1.57)	(-4.96)	(-3.88)	(-64.33)	(42.80)
Ln(mcap)	-0.7569***	0.7788***	0.0004***	-0.0001	-0.0006***	0.0003***	-0.0916***	0.0571***
	(-21.74)	(23.26)	(10.45)	(-1.45)	(-17.39)	(7.92)	(-15.54)	(9.01)
Raw return*Ln(mcap)	2.7258***	-2.4946***	-0.0002***	0.0001***	0.0002***	0.0002***	0.6032***	-0.5511***
	(36.64)	(-31.04)	(-9.98)	(4.07)	(3.33)	(3.53)	(57.35)	(-39.12)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,398,341	9,419,490	9,644,239	9,699,636	9,638,338	9,678,670	9,671,770	9,730,123
adj. R^2	0.11	0.07	0.31	0.27	0.33	0.27	0.70	0.64

Panel (b): Lagged cumulated five days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-24.0550***	18.5142***	-0.0073***	0.0011***	-0.0095***	0.0014***	-5.3141***	2.2557***
	(-44.66)	(36.04)	(-28.65)	(4.09)	(-23.53)	(3.28)	(-72.02)	(26.48)
Ln(mcap)	-1.3278***	1.2954***	-0.0004***	0.0001***	-0.0010***	0.0005***	-0.2591***	0.0997***
	(-33.92)	(34.66)	(-18.60)	(7.58)	(-27.55)	(11.85)	(-39.45)	(14.34)
Raw return*Ln(mcap)	-0.0636***	0.0470***	0.0005***	-0.0001***	0.0006***	-0.0001*	0.3461***	-0.1500***
	(-10.05)	(7.74)	(26.64)	(-3.32)	(19.62)	(-1.95)	(63.40)	(-23.46)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,476,649	9,497,732	9,723,237	9,778,537	9,717,304	9,757,464	9,750,818	9,809,069
adj. R^2	0.11	0.07	0.32	0.28	0.34	0.28	0.69	0.64

Panel (c): Lagged cumulated ten days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-25.5457***	20.0470***	-0.0080***	0.0013***	-0.0115***	0.0023***	-4.7346***	1.4230***
	(-52.75)	(42.91)	(-36.24)	(6.00)	(-34.19)	(6.56)	(-80.58)	(20.38)
Ln(mcap)	-2.4685***	2.3864***	-0.0007***	0.0002***	-0.0018***	0.0008***	-0.5217***	0.1441***
	(-46.59)	(47.32)	(-30.68)	(10.90)	(-42.66)	(17.20)	(-60.26)	(16.69)
Raw return*Ln(mcap)	1.7055***	-1.3225***	0.0005***	-0.0001***	0.0007***	-0.0001***	0.2999***	-0.0908***
	(47.19)	(-37.17)	(33.51)	(-4.76)	(28.34)	(-4.05)	(69.99)	(-17.68)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,508,375	9,528,948	9,755,062	9,809,779	9,749,103	9,788,661	9,782,680	9,840,365
adj. R^2	0.11	0.07	0.32	0.28	0.34	0.28	0.69	0.64

Table 6: **Price Impact Tests – Interaction with Analyst Coverage**

This table reports the results from regressing our four price impact measures on lagged raw returns, analyst coverage, the interaction of lagged returns and analyst coverage, and controls (unreported). Panel (a) shows the results for using lagged one-day returns as the key independent variable. Panel (b) shows the results for using lagged cumulated five-day returns as the key independent variable. Panel (c) shows the results for using lagged cumulated ten-day returns as the key independent variable. In each panel, columns (1) and (2) show results for the *price impact costs* measure (scaled by 10^6 for better visibility), columns (3) and (4) show results for the *lambda* measure (scaled by 10^2 for better visibility), columns (5) and (6) show results for the *quote-based price impact* measure, and columns (7) and (8) show results for the *ln(Amihud)* measure. Odd columns show results for the net-buy subsample (i.e., all stock-day obsvervations on which the trade imbalance is positive). Even columns show results for the net-sell subsample (i.e., all stock-day obsvervations on which the trade imbalance is negative). All regressions contain stock-level controls (past turnover, past bid-ask spread, the natural logarithm of market capitalization, and analyst coverage), stock-month and date fixed effects. (Because of the stock-month fixed effect, the coefficient on analyst coverage alone should not be interpreted.) t-statistics are based on standard errors adjusted for double-clustering by stock and day. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Panel (a): Lagged one day

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-13.9434***	12.3871***	-0.0021***	0.0002	-0.0025***	-0.0010***	-3.3960***	2.6674***
	(-48.90)	(44.10)	(-14.72)	(1.54)	(-10.59)	(-4.75)	(-75.06)	(48.78)
AnalCov	0.0002	-0.0257	-0.0000	0.0001	-0.0001	0.0001	-0.0049	0.0321*
	(0.00)	(-0.35)	(-0.99)	(1.31)	(-1.08)	(0.42)	(-0.29)	(1.77)
Raw return*AnalCov	4.5345***	-3.9051***	0.0008***	-0.0001	0.0006***	0.0004***	1.1366***	-0.9086***
	(35.93)	(-29.40)	(11.85)	(-0.94)	(5.77)	(3.55)	(58.38)	(-37.88)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,398,341	9,419,490	9,644,239	9,699,636	9,638,338	9,678,670	9,671,770	9,730,123
adj. R^2	0.11	0.07	0.31	0.27	0.33	0.27	0.70	0.64

Panel (b): Lagged cumulated five days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-8.1280***	6.4088***	-0.0025***	0.0004***	-0.0040***	0.0007***	-2.0554***	0.7691***
	(-52.99)	(46.80)	(-33.33)	(5.88)	(-32.29)	(6.40)	(-88.90)	(30.98)
AnalCov	-0.0113	-0.0983	-0.0001	0.0001	0.0011***	-0.0001	-0.0014	0.0153
	(-0.14)	(-1.03)	(-1.29)	(1.22)	(20.13)	(-1.46)	(-0.08)	(0.85)
Raw return*AnalCov	2.6738***	-1.9452***	0.0009***	-0.0001***	-0.0001	0.0001	0.6571***	-0.2269***
	(39.22)	(-30.90)	(27.37)	(-2.81)	(-1.09)	(0.31)	(66.41)	(-20.82)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,476,649	9,497,732	9,723,237	9,778,537	9,717,304	9,757,464	9,750,818	9,809,069
adj. R^2	0.11	0.07	0.32	0.28	0.34	0.28	0.69	0.64

Panel (c): Lagged cumulated ten days

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells	Net Buys	Net Sells
Raw return	-8.8193***	7.0346***	-0.0027***	0.0005***	-0.0048***	0.0013***	-1.8947***	0.5108***
	(-61.43)	(55.07)	(-40.63)	(8.59)	(-45.49)	(12.94)	(-98.13)	(23.81)
AnalCov	-0.0937	-0.0483	-0.0001*	0.0001	-0.0002	0.0001	-0.0285	0.0161
	(-0.97)	(-0.43)	(-1.84)	(1.39)	(-1.40)	(0.35)	(-1.29)	(0.89)
Raw return*AnalCov	2.8099***	-2.0457***	0.0009***	-0.0001***	0.0013***	-0.0002***	0.5595***	-0.1297***
	(46.92)	(-36.97)	(32.67)	(-3.82)	(27.59)	(-4.18)	(70.75)	(-14.93)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	9,508,375	9,528,948	9,755,062	9,809,779	9,749,103	9,788,661	9,782,680	9,840,365
adj. R^2	0.11	0.07	0.32	0.28	0.34	0.28	0.69	0.64

Table 7: **UWIP and Price Informativeness**

This table reports results from regressing the price jump ratio as constructed in Weller (2018) (see Section 4.1 for details) on our measure of uncertainty about what's in the price (*uwip*) and controls. To construct *uwip*, we first run the following regression for each earnings announcement event: price impact $\text{costs}_{i\tau} = \beta_0 + \beta_1\,\text{return}_{i\tau-1} + \beta_2\,\text{netsell}_{i\tau} + \beta_3\,\text{return}_{i\tau-1} \times \text{netsell}_{i\tau} + \epsilon_{i\tau}$, where i denotes the announcement stock, τ captures all trading days occurring in the calendar quarter before the announcement, and $\text{netsell}_{i\tau}$ is a dummy variable that takes a value of one on days with a negative trade imbalance. Our model with uncertainty about what's in the price predicts $\beta_3 > 0$. Panel (a) shows results using *uwip(raw)*, defined as the (winsorized) estimated coefficient $\hat{\beta}_3$, as the key independent variable. Panel (b) shows results using *uwip(dec)*, defined as the decile rank of $\hat{\beta}_3$, as the key independent variable. *Ln(mcap)* is the (natural logarithm of the) market capitalization at the beginning of the announcement quarter. *Anal coverage* is the (natural logarithm of one plus) the number of analysts following the stock at the beginning of the announcement quarter. *Avg return* is the average raw return in the prior quarter. *Stddev return* is the standard deviation of raw returns in the prior quarter. *Ln(turnover)* is the share turnover in the prior quarter. *Avg spread* is the average bid-ask spread in the prior quarter. All regressions contain day fixed effects; columns (1) and (3) contain industry fixed effects (based on the SIC-2 digit industry classification); columns (2), (3), (5), and (6) contain stock fixed effects, columns (5) and (6) further include industry-year fixed effects. t-statistics are based on standard errors adjusted for double-clustering by stock and day. ***, ** and * indicate statistical significance at the 1%, 5% and 10% level, respectively.

Panel (a): Uwip(raw) as key independent variable

	Price Jump Ratio					
	(1)	(2)	(3)	(4)	(5)	(6)
Uwip(raw)	106.721***	116.442***	117.694***	110.492***	102.156***	103.235***
	(3.21)	(3.25)	(3.29)	(3.32)	(2.85)	(2.88)
Ln(mcap)	0.0054***	0.0042	0.0077***	0.0037**	0.0061**	0.0105***
	(3.58)	(1.59)	(2.84)	(2.23)	(2.06)	(3.44)
Anal coverage	0.0238***	0.0223***	0.0228***	0.0223***	0.0234***	0.0238***
	(8.66)	(5.67)	(5.76)	(7.75)	(5.75)	(5.79)
Avg return				-0.4317	-0.1434	-0.2323
				(-1.20)	(-0.37)	(-0.59)
Stddev return				-0.8980***	-0.1143	-0.0011
				(-8.00)	(-0.87)	(-0.01)
Ln(turnover)				0.0067***	-0.0071***	-0.0074***
				(3.51)	(-2.71)	(-2.70)
Avg spread				0.2748**	0.0802	0.1421
				(2.33)	(0.52)	(0.93)
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	No	No	Yes	No	No
Stock FE	No	Yes	Yes	No	Yes	Yes
Ind*year FE	No	No	Yes	No	No	Yes
N	102,223	100,912	100,851	102,053	100,743	100,682
adj. R^2	0.06	0.09	0.09	0.06	0.09	0.09

Panel (b): Uwip(dec) as key independent variable

	Price Jump Ratio					
	(1)	(2)	(3)	(4)	(5)	(6)
Uwip(raw)	0.0012**	0.0014**	0.0014***	0.0013***	0.0012**	0.0013**
	(2.41)	(2.58)	(2.61)	(2.65)	(2.35)	(2.37)
Ln(mcap)	0.0053***	0.0040	0.0076***	0.0037**	0.0061**	0.0105***
	(3.53)	(1.52)	(2.77)	(2.22)	(2.05)	(3.44)
Anal coverage	0.0237***	0.0223***	0.0227***	0.0223***	0.0234***	0.0238***
	(8.62)	(5.65)	(5.74)	(7.77)	(5.74)	(5.79)
Avg return				-0.4394	-0.1526	-0.2421
				(-1.22)	(-0.39)	(-0.62)
Stddev return				-0.9002***	-0.1140	-0.0012
				(-8.02)	(-0.87)	(-0.01)
Ln(turnover)				0.0065***	-0.0074***	-0.0077***
				(3.40)	(-2.84)	(-2.83)
Avg spread				0.2908**	0.0893	0.1514
				(2.46)	(0.58)	(0.99)
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	No	No	Yes	No	No
Stock FE	No	Yes	Yes	No	Yes	Yes
Ind*year FE	No	No	Yes	No	No	Yes
N	102,223	100,912	100,851	102,053	100,743	100,682
adj. R^2	0.06	0.09	0.09	0.06	0.09	0.09

A Appendix - Proofs

A.1 Proof of Proposition 2 - Both m and s are common knowledge

The main steps of the proof are in the text. Here, we display the calculations for the order size $x_{(1)}$ when $m \neq s$ (if $m = s$, then S do not trade). In that case, the price conjecture in Equation (1) leads to the following:

- If $m \neq s$ and $\theta_s = \sigma$ (which occurs with probability $1/2 \times 1/2 = 1/4$), then S buy $x_{(1)}$ shares so $\omega_1 = x_{(1)} + n$ and

$$\theta - p_1 = (\theta_m + \sigma) - p_1 = \begin{cases} 0 & \text{with proba. } x_{(1)}/4 \text{ (i.e., for } -2x_{(1)} + 1 < n \leq 1) \\ \sigma & \text{with proba. } (1 - x_{(1)})/4 \text{ (i.e., for } -1 \leq n \leq -2x_{(1)} + 1) \\ 2\sigma & \text{with proba. } 0 \text{ (i.e., for } -2x_{(1)} - 1 \leq n < -1) \end{cases}$$

  As a result, $E\left[\theta - p_1 | p_0, \theta_s = \sigma, m \neq s\right] = \sigma(1 - x_{(1)})$ and $Var\left[\theta - p_1 | p_0, \theta_s = \sigma, m \neq s\right] = \sigma^2 x_{(1)}(1 - x_{(1)})$. Plugging these expressions into the first-order condition for S' profit maximization and imposing rational expectations ($x_i = x_{(1)}$ for all i) yields:

$$x_{(1)} = \frac{E\left[\theta - p_1 | p_0, \theta_s = \sigma, m \neq s\right]}{\gamma Var\left[\theta - p_1 | p_0, \theta_s = \sigma, m \neq s\right]} = \frac{\sigma(1 - x_{(1)})}{\gamma \sigma^2 x_{(1)}(1 - x_{(1)})}$$

  and hence $x_{(1)} = \sqrt{1/(\gamma\sigma)}$.

- If $m \neq s$ and $\theta_s = -\sigma$ (which occurs with probability $1/2 \times 1/2 = 1/4$), then S sell $x_{(1)}$ shares so $\omega_1 = -x_{(1)} + n$ and

$$\theta - p_1 = (\theta_m - \sigma) - p_1 = \begin{cases} -2\sigma & \text{with proba. } 0 \text{ (i.e., for } 1 < n \leq 2x_{(1)} + 1) \\ -\sigma & \text{with proba. } (1 - x_{(1)})/4 \text{ (i.e., for} 2x_{(1)} - 1 \leq n \leq +1) \\ 0 & \text{with proba. } x_{(1)}/4 \text{ (i.e., for } -1 \leq n < 2x_{(1)} - 1) \end{cases}$$

  Hence, $E\left[\theta - p_1 | p_0, \theta_s = -\sigma, m \neq s\right] = -\sigma(1 - x_{(1)})$ and $Var\left[\theta - p_1 | p_0, \theta_s = -\sigma, m \neq s\right] = \sigma^2 x_{(1)}(1 - x_{(1)})$. Plugging these expressions into the first-order condition and imposing rational expectations ($x_i = x_{(1)}$ for all i) yields:

$$-x_{(1)} = \frac{E\left[\theta - p_1 | p_0, \theta_s = -\sigma, m \neq s\right]}{\gamma Var\left[\theta - p_1 | p_0, \theta_s = -\sigma, m \neq s\right]} = \frac{-\sigma(1 - x_{(1)})}{\gamma \sigma^2 x_{(1)}(1 - x_{(1)})},$$

  which yields again $x_{(1)} = \sqrt{1/(\gamma\sigma)}$. Thus, the order size is identical for $\theta_s = +\sigma$ and$\theta_s = -\sigma$, which confirms our conjecture.

A.2 Proof of Proposition 3 - Only m is common knowledge

The main steps of proof are in the text. Here, we first display the calculations for Ms' expectation of θ_s conditional on observing an order flow $-x_{(2)} + 1 \leq \omega_1 \leq 1$ (for other values of the order flow, M either learn θ_s perfectly or nothing at all); in that case, M know that either $m = s$ (and S do not trade) or $m \neq s$ and $\theta_s = \sigma$ (and S buy $x_{(2)}$). The former occurs with a probability 1/2 and the latter with a probability $1/2 \times 1/2 = 1/4$. Hence, $E(\theta_s | -x_{(2)} + 1 \leq \omega_1 \leq 1) = \frac{1/2 \times 0 + 1/4\sigma}{1/2 + 1/4} = \frac{1}{3}\sigma$.
Next, we display the calculations for the order size $x_{(2)}$ when $m \neq s$ (if $m = s$, then S do not trade). In that case, the price conjecture in Equation (2) leads to the following:

- If $m \neq s$ and $\theta_s = \sigma$ (which occurs with probability $1/2 \times 1/2 = 1/4$), then $\omega_1 = x_{(2)} + n$ and

$$\theta - p_1 = (\theta_m + \sigma) - p_1 = \begin{cases} 0 & \text{with proba. } x_{(2)}/8 \text{ (i.e., for } 1 - x_{(2)} < n \leq 1) \\ \frac{2}{3}\sigma & \text{with proba. } x_{(2)}/8 \text{ (i.e., for } -2x_{(2)} + 1 < n \leq 1 - x_{(2)}) \\ \sigma & \text{with proba. } (1 - x_{(2)})/4 \text{ (i.e., for } -1 \leq n \leq -2x_{(2)} + 1) \\ \frac{4}{3}\sigma & \text{with proba. } 0 \text{ (i.e., for } -1 - x_{(2)} \leq n < -1) \\ 2\sigma & \text{with proba. } 0 \text{ (i.e., for } -2x_{(2)} - 1 \leq n < -x_{(2)} - 1) \end{cases}$$

  As a result, $E[\theta - p_1|p_0, \theta_s = \sigma, m \neq s] = \sigma\left(1 - \frac{2}{3}x_{(2)}\right)$ and $Var[\theta - p_1|p_0, \theta_s = \sigma, m \neq s] = \frac{1}{9}\sigma^2 x_{(2)}(5 - 4x_{(2)})$. Plugging these expressions into the first-order condition for S' profit maximization and imposing rational expectations ($x_i = x_{(1)}$ for all i) yields:

$$x_{(2)} = \frac{E[\theta - p_1|p_0, \theta_s = \sigma, m \neq s]}{\gamma Var[\theta - p_1|p_0, \theta_s = \sigma, m \neq s]} = \frac{\sigma\left(1 - \frac{2}{3}x_{(2)}\right)}{\gamma\frac{1}{9}\sigma^2 x_{(2)}(5 - 4x_{(2)})}.$$

  Rearranging leads to the cubic equation:

$$9 - 6x_{(2)} - 5\gamma\sigma x_{(2)}^2 + 4\gamma\sigma x_{(2)}^3 = 0 \tag{5}$$

- If $m \neq s$ and $\theta_s = -\sigma$ (which occurs with probability $1/2 \times 1/2 = 1/4$), then $\omega_1 = -x_{(2)} + n$ and

$$p_1 = (\theta_m - \sigma) - p_1 = \begin{cases} -2\sigma & \text{with proba. } 0 \text{ (i.e., for } 1 + x_{(2)} < n \leq 1) \\ -\frac{4}{3}\sigma & \text{with proba. } 0 \text{ (i.e., for } 1 < n \leq x_{(2)} + 1) \\ -\sigma & \text{with proba. } (1 - x_{(2)})/4 \text{ (i.e., for } 2x_{(2)} - 1 \leq n \leq 1) \\ -\frac{2}{3}\sigma & \text{with proba. } x_{(2)}/8 \text{ (i.e., for } -1 + x_{(2)} \leq n < 2x_{(2)} - 1) \\ 0 & \text{with proba. } x_{(2)}/8 \text{ (i.e., for } -1 \leq n < x_{(2)} - 1) \end{cases}$$

  Hence, $E[\theta - p_1|p_0, \theta_s = -\sigma, m \neq s] = -\sigma\left(1 - \frac{2}{3}x_{(2)}\right)$ and $Var[\theta - p_1|p_0, \theta_s = -\sigma, m \neq s] = \frac{1}{9}\sigma^2 x_{(2)}(5 - 4x_{(2)})$. Plugging these expressions into the first-order condition yields:

$$-x_{(2)} = \frac{E[\theta - p_1|p_0, \theta_s = -\sigma, m \neq s]}{\gamma Var[\theta - p_1|p_0, \theta_s = -\sigma, m \neq s]} = \frac{-\sigma\left(1 - \frac{2}{3}x_{(2)}\right)}{\gamma\frac{1}{9}\sigma^2 x_{(2)}(5 - 4x_{(2)})},$$

  which again leads to Equation 5. Thefore, the order size is identical for $\theta_s = +\sigma$ and $\theta_s = -\sigma$, which confirms our conjecture.

To prove the existence and unicity of $x_{(2)}$, let $f(x) \equiv 9 - 6x - 5\gamma\sigma x^2 + 4\gamma\sigma x^3$. Given that $f(0) = 9 > 0$ and $f(1) = 3 - \gamma\sigma < 0$ by Assumption 1, f admits at least one root over the interval [0,1]. Note that, if Assumption 1 does not hold, i.e., if $\gamma\sigma < 3$, then f admits no root over that interval, implying that there is no equilibrium in which speculators' trades are not fully revealing. To establish the unicity of x, differentiate f and observe that $f'(x) = -6 - 10\gamma\sigma x + 12\gamma\sigma x^2$ admits 2 roots, $x_{-/+} = (5\gamma\sigma \pm \sqrt{25\gamma^2\sigma^2 + 72\gamma\sigma})/12$ where $x_- < 0$ and $x_+ > 1$ given Assumption 1. As a result, f' is negative over the interval [0,1], implying that f is monotonically decreasing over that interval. We conclude that f admits at most one root, $x_{(2)}$, over [0,1].

A.3 Proof of Proposition 4 - Neither m nor s are common knowledge

Recall that we conjecture that S buy (sell) an amount $x_{(3)}$ ($-x_{(3)}$) when $\theta_m \neq \theta_s$ and that they buy (sell) an amount $y_{(3)}$ ($-y_{(3)}$) when $\theta_m = \theta_s$ with $x_{(3)} \geq y_{(3)}$. We label $\neg m$ the component of the

fundamental that is not observed by M; for instance, if $m = 1$ (i.e., M observe $\theta_m = \theta_1$), then $\neg m = 2$ (i.e., M do not observe $\theta_{\neg m} = \theta_2$).
The main steps of the proof are in the text. Here, we first display the calculations for Ms' expectation of $\theta_{\neg m}$. Suppose M observe $\theta_m = \sigma$ and an order flow $-x_{(3)} + 1 \leq \omega_1 \leq y_{(3)} + 1$. In that case, M infer that S bought $y_{(3)}$ and hence that $\theta_s = \sigma$. The configuration $\theta_m = \theta_s = \sigma$ occurs either if $m = s$ (probability $1/2 \times 1/2 = 1/4$) or if $m \neq s$ and $\theta_m = \theta_s = \sigma$ (probability $1/2 \times 1/2 \times 1/2 = 1/8$). Hence, $E(\theta_{\neg m}|\theta_m = \sigma, -x_{(3)} + 1 \leq \omega_1 \leq y_{(3)} + 1) = \frac{1/4 \times 0 + 1/8\sigma}{1/4 + 1/8} = \frac{1}{3}\sigma$. Likewise, $E(\theta_{\neg m}|\theta_m = -\sigma, -y_{(3)} - 1 \leq \omega_1 \leq x_{(3)} - 1) = -\frac{1}{3}\sigma$.
Next, we display the calculations for the order sizes, $x_{(3)}$ and $y_{(3)}$. Denote $z \equiv \frac{x_{(3)} + y_{(3)}}{2}$. The price conjecture in Equations (3) and (4) lead to the following, starting with the case $m = s$:

- If $m = s, \theta_m = \theta_s = \sigma$ and $\theta_{\neg m} = \sigma$ (which occurs with probability $1/2 \times 1/2 \times 1/2 = 1/8$), then S buy $y_{(3)}$ shares so $\omega_1 = y_{(3)} + n$ and

$$(\theta, p_1, \theta - p_1) = \begin{cases} (2\sigma, \frac{4}{3}\sigma, \frac{2}{3}\sigma) & \text{with proba. } z/8 \text{ (i.e., for } -x_{(3)} + 1 - y_{(3)} < n \leq 1) \\ (2\sigma, \sigma, \sigma) & \text{with proba. } (1-z)/8 \text{ (i.e., for } -1 < n \leq -x_{(3)} + 1 - y_{(3)}) \\ (2\sigma, 0, 2\sigma) & \text{with proba. } 0 \text{ (i.e., for } -x_{(3)} - 1 - y_{(3)} < n \leq -1) \end{cases}$$

- If $m = s, \theta_m = \theta_s = \sigma$ and $\theta_{\neg m} = -\sigma$ (which occurs with probability $1/2 \times 1/2 \times 1/2 = 1/8$), then again S buy $y_{(3)}$ shares so $\omega_1 = y_{(3)} + n$ and

$$(\theta, p_1, \theta - p_1) = \begin{cases} (0, \frac{4}{3}\sigma, -\frac{4}{3}\sigma) & \text{with proba. } z/8 \text{ (i.e., for } -x_{(3)} + 1 - y_{(3)} < n \leq 1) \\ (0, \sigma, -\sigma) & \text{with proba. } (1-z)/8 \text{ (i.e., for } -1 < n \leq -x_{(3)} + 1 - y_{(3)}) \\ (0, 0, 0) & \text{with proba. } 0 \text{ (i.e., for } -x_{(3)} - 1 - y_{(3)} < n \leq -1) \end{cases}$$

As a result, $E[\theta - p_1|\theta_m = \sigma, \theta_s = \sigma] = \frac{1}{3}\sigma(1 - z)$ and $Var[\theta - p_1|\theta_m = \sigma, \theta_s = \sigma] = \frac{1}{9}\sigma^2(8 + z - z^2)$. Plugging these expressions into the first-order condition for S' profit maximization and imposing rational expectations ($x_i = x_{(1)}$ for all i) yields:

$$y_{(3)} = \frac{E[\theta - p_1|\theta_m = \sigma, \theta_s = \sigma]}{\gamma Var[\theta - p_1|\theta_m = \sigma, \theta_s = \sigma]} = \frac{\frac{1}{3}\sigma(1-z)}{\gamma\frac{1}{9}\sigma^2(8 + z - z^2)} = \frac{3(1-z)}{\gamma\sigma(8 + z - z^2)}.$$

- If $m = s, \theta_m = \theta_s = -\sigma$ and $\theta_{\neg m} = \sigma$ (which occurs with probability $1/2 \times 1/2 \times 1/2 = 1/8$), then S sell $y_{(3)}$ shares so $\omega_1 = -y_{(3)} + n$ and

$$(\theta, p_1, \theta - p_1) = \begin{cases} (0, 0, 0) & \text{with proba. } 0 \text{ (i.e., for } 1 < n \leq 1 + x_{(3)} + y_{(3)}) \\ (0, -\sigma, \sigma) & \text{with proba. } (1-z)/8 \text{ (i.e., for } -1 + x_{(3)} + y_{(3)} < n \leq 1 \\ (0, -\frac{4}{3}\sigma, \frac{4}{3}\sigma) & \text{with proba. } z/8 \text{ (i.e., for } -1 < n \leq -1 + x_{(3)} + y_{(3)}) \end{cases}$$

- If $m = s, \theta_m = \theta_s = -\sigma$ and $\theta_{\neg m} = -\sigma$ (which occurs with probability $1/2 \times 1/2 \times 1/2 = 1/8$), then again S sell $y_{(3)}$ shares so $\omega_1 = -y_{(3)} + n$ and

$$(\theta, p_1, \theta - p_1) = \begin{cases} (-2\sigma, 0, -2\sigma) & \text{with proba. } z/8 \text{ (i.e., for } 1 < n \leq 1 + x_{(3)} + y_{(3)}) \\ (-2\sigma, -\sigma, -\sigma) & \text{with proba. } (1-z)/8 \text{ (i.e., for } -1 + x_{(3)} + y_{(3)} < n \leq 1 \\ (-2\sigma, -\frac{4}{3}\sigma, -\frac{2}{3}\sigma) & \text{with proba. } 0 \text{ (i.e., for } -1 < n \leq -1 + x_{(3)} + y_{(3)} \end{cases}$$

As a result, $E[\theta - p_1|\theta_m = -\sigma, \theta_s = -\sigma] = -\frac{1}{3}\sigma(1 - z)$ and $Var[\theta - p_1|\theta_m = \sigma, \theta_s = \sigma] = \frac{1}{9}\sigma^2(8 + z - z^2)$. Plugging these expressions into the first-order condition for S' profit maximization and imposing rational expectations ($x_i = x_{(1)}$ for all i) yields:

$$-y_{(3)} = \frac{E[\theta - p_1|\theta_m = \sigma, \theta_s = \sigma]}{\gamma Var[\theta - p_1|\theta_m = \sigma, \theta_s = \sigma]} = \frac{-\frac{1}{3}\sigma(1-z)}{\gamma\frac{1}{9}\sigma^2(8 + z - z^2)} = -\frac{3(1-z)}{\gamma\sigma(8 + z - z^2)},$$

which is the same equation as in the case $(\theta_m = \sigma, \theta_s = \sigma)$.

We consider next the case $m \neq s$:

- If $m \neq s, \theta_m = \sigma$ and $\theta_s = \sigma$ (which occurs with probability $1/2 \times 1/2 \times 1/2 = 1/8$), then S buy $y_{(3)}$ shares so $\omega_1 = y_{(3)} + n$ and

$$(\theta, p_1, \theta - p_1) = \begin{cases} (2\sigma, \frac{4}{3}\sigma, \frac{2}{3}\sigma) & \text{with proba. } z/8 \text{ (i.e., for } -x_{(3)} + 1 - y_{(3)} < n \leq 1) \\ (2\sigma, \sigma, \sigma) & \text{with proba. } (1-z)/8 \text{ (i.e., for } -1 < n \leq -x_{(3)} + 1 - y_{(3)}) \\ (2\sigma, 0, 2\sigma) & \text{with proba. } 0 \text{ (i.e., for } -x_{(3)} - 1 - y_{(3)} < n \leq -1) \end{cases}$$

This case yields identical moments and equilibrium condition as the case $(m = s, \theta_m = \sigma, \theta_s = \sigma)$.

- If $m \neq s, \theta_m = \sigma$ and $\theta_s = -\sigma$ (which occurs with probability $1/2 \times 1/2 \times 1/2 = 1/8$), then S sell $x_{(3)}$ shares so $\omega_1 = -x_{(3)} + n$ and

$$(\theta, p_1, \theta - p_1) = \begin{cases} (0, \frac{4}{3}\sigma, -\frac{4}{3}\sigma) & \text{with proba. } 0 \text{ (i.e., for } 1 < n \leq 1 + x_{(3)} + y_{(3)}) \\ (0, \sigma, -\sigma) & \text{with proba. } (1-z)/8 \text{ (i.e., for } -1 + x_{(3)} + y_{(3)} < n \leq 1) \\ (0, 0, 0) & \text{with proba. } z/8 \text{ (i.e., for } -1 < n \leq -1 + x_{(3)} + y_{(3)}) \end{cases}$$

As a result, $E\left[\theta - p_1 | \theta_m = \sigma, \theta_s = -\sigma\right] = -\sigma\left(1 - z\right)$ and $Var\left[\theta - p_1 | \theta_m = \sigma, \theta_s = -\sigma\right] = \sigma^2 z(1-z)$. Plugging these expressions into the first-order condition for S' profit maximization and imposing rational expectations ($x_i = x_{(1)}$ for all i) yields:

$$-x_{(3)} = \frac{E\left[\theta - p_1 | \theta_m = \sigma, \theta_s = -\sigma\right]}{\gamma Var\left[\theta - p_1 | \theta_m = \sigma, \theta_s = -\sigma\right]} = \frac{-\sigma\left(1-z\right)}{\gamma\sigma^2 z(1-z)} = -\frac{1}{\gamma\sigma z}.$$

- If $m \neq s, \theta_m = -\sigma$ and $\theta_s = \sigma$ (which occurs with probability $1/2 \times 1/2 \times 1/2 = 1/8$), then S buy $x_{(3)}$ shares so $\omega_1 = x_{(3)} + n$ and

$$(\theta, p_1, \theta - p_1) = \begin{cases} (0, 0, 0) & \text{with proba. } z/8 \text{ (i.e., for } 1 - x_{(3)} - y_{(3)} < n \leq 1) \\ (0, -\sigma, \sigma) & \text{with proba. } (1-z)/8 \text{ (i.e., for } -1 < n \leq 1 - x_{(3)} - y_{(3)}) \\ (0, -\frac{4}{3}\sigma, \frac{4}{3}\sigma) & \text{with proba. } 0 \text{ (i.e., for } -1 - x_{(3)} - y_{(3)} < n \leq -1) \end{cases}$$

As a result, $E\left[\theta - p_1 | \theta_m = -\sigma, \theta_s = \sigma\right] = \sigma\left(1 - z\right)$ and $Var\left[\theta - p_1 | \theta_m = -\sigma, \theta_s = \sigma\right] = \sigma^2 z(1-z)$. Plugging these expressions into the first-order condition for S' profit maximization and imposing rational expectations ($x_i = x_{(1)}$ for all i) yields:

$$x_{(3)} = \frac{E\left[\theta - p_1 | \theta_m = -\sigma, \theta_s = \sigma\right]}{\gamma Var\left[\theta - p_1 | \theta_m = -\sigma, \theta_s = \sigma\right]} = \frac{\sigma\left(1-z\right)}{\gamma\sigma^2 z(1-z)} = \frac{1}{\gamma\sigma z},$$

which is the same equation as in the case $(\theta_m = \sigma, \theta_s = -\sigma)$.

- If $m \neq s, \theta_m = -\sigma$ and $\theta_s = -\sigma$ (which occurs with probability $1/2 \times 1/2 \times 1/2 = 1/8$), then S sell $y_{(3)}$ shares so $\omega_1 = -y_{(3)} + n$ and

$$(\theta, p_1, \theta - p_1) = \begin{cases} (-2\sigma, 0, -2\sigma) & \text{with proba. } z/8 \text{ (i.e., for } 1 < n \leq 1 + x_{(3)} + y_{(3)}) \\ (-2\sigma, -\sigma, -\sigma) & \text{with proba. } (1-z)/8 \text{ (i.e., for } -1 + x_{(3)} + y_{(3)} < n \leq 1) \\ (-2\sigma, -\frac{4}{3}\sigma, -\frac{2}{3}\sigma) & \text{with proba. } 0 \text{ (i.e., for } -1 < n \leq -1 + x_{(3)} + y_{(3)}) \end{cases}$$

This case yields identical moments and equilibrium condition as the case $(m = s, \theta_m = -\sigma, \theta_s = -\sigma)$.

Collecting the different cases and substituting out $z \equiv \frac{x_{(3)}+y_{(3)}}{2}$, investors' first-order conditions yield a system of two equations in $x_{(3)}$ and $y_{(3)}$:

$$\begin{aligned} x_{(3)} &= \frac{1}{\gamma\sigma\frac{x_{(3)}+y_{(3)}}{2}} \\ y_{(3)} &= \frac{3\left(1-\frac{x_{(3)}+y_{(3)}}{2}\right)}{\gamma\sigma\left(8+\frac{x_{(3)}+y_{(3)}}{2}-\left(\frac{x_{(3)}+y_{(3)}}{2}\right)^2\right)} \end{aligned}$$

The first equation implies that $y_{(3)} = \frac{2-\gamma\sigma x_{(3)}^2}{\gamma\sigma x_{(3)}}$. Plugging this expression in the first equation and rearranging leads to the quartic equation:

$$1-\gamma\sigma x_{(3)} - 2\gamma\sigma(1+4\gamma\sigma)x_{(3)}^2 + 2(\gamma\sigma)^2 x_{(3)}^3 + 4(\gamma\sigma)^3 x_{(3)}^4 = 0. \tag{6}$$

The equilibrium is thus characterized by Equation (6), together with the requirement that $x_{(3)} \geq y_{(3)}$. Since $y_{(3)} = \frac{2-\gamma\sigma x_{(3)}^2}{\gamma\sigma x_{(3)}}$, $x_{(3)} \geq y_{(3)}$ is equivalent to $x \geq 1/\sqrt{\gamma\sigma}$. Let $g(x) \equiv 1-\gamma\sigma x - 2\gamma\sigma(1+4\gamma\sigma)x^2 + 2(\gamma\sigma)^2x^3 + 4(\gamma\sigma)^3x^4$. We show next that g admits exactly one root in the interval $\left[1/\sqrt{\gamma\sigma}, 1\right]$, implying that there exists a unique equilibrium. The second derivative of g, $g''(x) = -4\gamma\sigma(1+4\gamma\sigma)+12(\gamma\sigma)^2x+48(\gamma\sigma)^3x^2$, is a quadratic function which admits two roots: one root, $(-1-\sqrt{1+48(1+4\gamma\sigma)/9})/(8\gamma\sigma)$, is negative and the other, $x_+ \equiv (-1+\sqrt{1+48(1+4\gamma\sigma)/9})/(8\gamma\sigma)$, is between 0 and 1. It follows that $g''(x) \leq 0$ for x in $[0, x_+]$ and $g''(x) \geq 0$ for x in $[x_+, 1]$, and so that g' is decreasing over $[0, x_+]$ and increasing over $[x_+, 1]$, where $g'(x) = -\gamma\sigma - 4\gamma\sigma(1+4\gamma\sigma)x + 6(\gamma\sigma)^2x^2 + 16(\gamma\sigma)^3x^3$. Given that $g'(0) = -\gamma\sigma < 0$ and $g'(1) = \gamma\sigma(-5-10\gamma\sigma+16(\gamma\sigma)^2) = \gamma\sigma(5(-1+(\gamma\sigma)^2)+10\gamma\sigma(-1+\gamma\sigma)+(\gamma\sigma)^2) > 0$ (from Assumption 1, each term in brackets is positive), there exists a unique x_* in $[x_+, 1]$ such that $g'(x) \leq 0$ for x in $[0, x_*]$ and $g'(x) \geq 0$ for $[x_*, 1]$. This implies in turn that g decreases over $[0, x_*]$ and increases over $[x_*, 1]$. Finally, observing that $g(1/\sqrt{\gamma\sigma}) < 0$ and $g(1) > 0$, g admits a unique root, $x_{(3)}$, in the interval $\left[1/\sqrt{\gamma\sigma}, 1\right]$. Hence, there exists a unique equilibrium.

A.4 Corollary 5 - Trading aggressiveness

In the proof of Proposition 4, we establish that $x_{(3)} > 1/\sqrt{\gamma\sigma}$, which leads to $x_{(3)} > x_{(1)} = 1/\sqrt{\gamma\sigma}$. This inequality further implies that $2/(\gamma\sigma x_{(3)}) < 2/\sqrt{\gamma\sigma}$. It follows that $y_{(3)} = \frac{2}{\gamma\sigma x_{(3)}} - x_{(3)} < 2/\sqrt{\gamma\sigma} - 1/\sqrt{\gamma\sigma} = 1/\sqrt{\gamma\sigma} = x_{(1)}$. Hence $0 < y_{(3)} < x_{(1)} < x_{(3)} < 1$.

A.5 Corollary 7 - Price informativeness

When m and s are common knowledge, price informativeness is given by:

$$\begin{aligned} E\left(Var\left(\theta|p_1,p_0\right)\right) &= Pr(m=s)\,E\left(\left(\theta-p_1\right)^2|m=s\right) \\ &\quad +Pr(m\neq s)\,E\left(\left(\theta-p_1\right)^2|m\neq s\right) \\ &= \frac{1}{2}\sigma^2 + \frac{1}{2}\sigma^2(1-x_{(1)}) \\ &= \sigma^2\left(1-\frac{1}{2}x_{(1)}\right) \end{aligned}$$

The ex-ante uncertainty is $Var(\theta) = 2\sigma^2$. When $x = 0$, half of this uncertainty is resolved through the publication of θ_m in the t=0 price. When $x = 1$, all the uncertainty is resolved for the case $m \neq s$,

while only half of the uncertainty is resolved for the case $m = s$.
When only m is common knowledge, price informativeness is given by:

$$
\begin{aligned}
E\left(Var\left(\theta|p_1, p_0\right)\right) &= Pr(m=s)\, E\left(\left(\theta - p_1\right)^2 |m=s\right) \\
&\quad + Pr(m \neq s)\, E\left(\left(\theta - p_1\right)^2 |m \neq s\right) \\
&= \frac{1}{2}\left[\begin{array}{c} Pr(m=s, \theta_s = \sigma)\, E\left(\left(\theta - p_1\right)^2 |m=s, \theta_s = \sigma\right) \\ + Pr(m=s, \theta_s = -\sigma)\, E\left(\left(\theta - p_1\right)^2 |m=s, \theta_s = -\sigma\right) \end{array}\right] \\
&\quad + \frac{1}{2}\left[\begin{array}{c} Pr(m \neq s, \theta_s = \sigma)\, E\left(\left(\theta - p_1\right)^2 |m \neq s, \theta_s = \sigma\right) \\ + Pr(m \neq s, \theta_s = -\sigma)\, E\left(\left(\theta - p_1\right)^2 |m \neq s, \theta_s = -\sigma\right) \end{array}\right] \\
&= \frac{1}{2}\left[\frac{1}{2}\left(\left(\frac{2}{3}\sigma\right)^2 \frac{x_{(2)}}{2} + \sigma^2(1 - x_{(2)}) + \left(\frac{4}{3}\sigma\right)^2 \frac{x_{(2)}}{2}\right) + \frac{1}{2}\left(\left(\frac{4}{3}\sigma\right)^2 \frac{x_{(2)}}{2} + \sigma^2(1 - x_{(2)}) + \left(\frac{2}{3}\sigma\right)^2 \frac{x_{(}}{2}\right.\right. \\
&\quad + \frac{1}{2}\left[\frac{1}{2}\left(\sigma^2(1 - x_{(2)}) + \left(\frac{2}{3}\sigma\right)^2 \frac{x_{(2)}}{2}\right) + \frac{1}{2}\left(\left(\frac{2}{3}\sigma\right)^2 \frac{x_{(2)}x}{2} + \sigma^2(1 - x_{(2)})\right)\right] \\
&= \frac{1}{2}\left[\left(\left(\frac{2}{3}\sigma\right)^2 \frac{x_{(2)}}{2} + \sigma^2(1 - x_{(2)}) + \left(\frac{4}{3}\sigma\right)^2 \frac{x_{(2)}}{2}\right)\right] \\
&\quad + \frac{1}{2}\left[\left(\sigma^2(1 - x_{(2)}) + \left(\frac{2}{3}\sigma\right)^2 \frac{x_{(2)}}{2}\right)\right] \\
&= \left(\frac{2}{3}\sigma\right)^2 \frac{x_{(2)}}{2} + \sigma^2(1 - x_{(2)}) + \frac{1}{2}\left(\frac{4}{3}\sigma\right)^2 \frac{x_{(2)}}{2} \\
&= \sigma^2(1 - x_{(2)}) + x_{(2)}\sigma^2 \frac{2}{3} \\
&= \sigma^2\left(1 - \frac{1}{3}x_{(2)}\right)
\end{aligned}
$$

As before, when $x = 0$, half of the total uncertainty is resolved through the publication of θ_m in the t=0 price. When $x = 1$, an additional one sixth of the total uncertainty is resolved through the trading by S.

When neither m nor s are common knowledge, price informativeness is given by:

$$
\begin{aligned}
E\left(Var\left(\theta|p_1,p_0\right)\right) &= Pr(m=s)\,E\left(\left(\theta-p_1\right)^2|m=s\right)\\
&\quad +Pr(m\neq s)\,E\left(\left(\theta-p_1\right)^2|m\neq s\right)\\
&= \frac{1}{2}\left[\begin{array}{c} Pr(m=s,\theta_s=\sigma,\theta_{\neg m}=\sigma)\,E\left(\left(\theta-p_1\right)^2|m=s,\theta_s=\sigma,\theta_{\neg m}=\sigma\right)\\ +Pr(m=s,\theta_s=\sigma,\theta_{\neg m}=-\sigma)\,E\left(\left(\theta-p_1\right)^2|m=s,\theta_s=\sigma,\theta_{\neg m}=-\sigma\right)\\ +Pr(m=s,\theta_s=-\sigma,\theta_{\neg m}=\sigma)\,E\left(\left(\theta-p_1\right)^2|m=s,\theta_s=-\sigma,\theta_{\neg m}=\sigma\right)\\ +Pr(m=s,\theta_s=-\sigma,\theta_{\neg m}=-\sigma)\,E\left(\left(\theta-p_1\right)^2|m=s,\theta_s=-\sigma,\theta_{\neg m}=-\sigma\right)\end{array}\right]\\
&\quad +\frac{1}{2}\left[\begin{array}{c} Pr(m\neq s,\theta_s=\sigma,\theta_m=\sigma)\,E\left(\left(\theta-p_1\right)^2|m\neq s,\theta_s=\sigma,\theta_m=\sigma\right)\\ +Pr(m\neq s,\theta_s=\sigma,\theta_m=-\sigma)\,E\left(\left(\theta-p_1\right)^2|m\neq s,\theta_s=\sigma,\theta_m=-\sigma\right)\\ +Pr(m\neq s,\theta_s=-\sigma,\theta_m=\sigma)\,E\left(\left(\theta-p_1\right)^2|m\neq s,\theta_s=-\sigma,\theta_m=\sigma\right)\\ +Pr(m\neq s,\theta_s=-\sigma,\theta_m=-\sigma)\,E\left(\left(\theta-p_1\right)^2|m\neq s,\theta_s=-\sigma,\theta_m=-\sigma\right)\end{array}\right]\\
&= \frac{1}{2}\left[\begin{array}{c} \frac{1}{4}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{4}\left(\left(\frac{4}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)\\ +\frac{1}{4}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{4}\left(\left(\frac{4}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)\end{array}\right]\\
&\quad +\frac{1}{2}\left[\begin{array}{c} \frac{1}{4}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{4}\left(\sigma^2\left(1-\frac{x+y}{2}\right)\right)\\ \frac{1}{4}\left(\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{4}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)\end{array}\right]\\
&= \frac{1}{2}\left[\frac{1}{2}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{2}\left(\left(\frac{4}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)\right]\\
&\quad +\frac{1}{2}\left[\frac{1}{2}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{2}\left(\sigma^2\left(1-\frac{x+y}{2}\right)\right)\right]\\
&= \frac{1}{2}\left[\frac{10}{9}\sigma^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right]\\
&\quad +\frac{1}{2}\left[\frac{2}{9}\sigma^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right]\\
&= \frac{2}{3}\sigma^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\\
&= \sigma^2\left(1-\frac{1}{3}\frac{x+y}{2}\right)
\end{aligned}
$$

As before, when $x=y=0$, half of the total uncertainty is resolved through the publication of θ_m in the t=0 price. If $x=y=1$ were possible, then there would be a further reduction of total uncertainty of about 13.88%.

Alternative proof for PI that follows directly from the Proof of Proposition 4:

- $p_1=E\left(\theta|\omega_1,p_0\right)\Rightarrow E\left(p_1|p_1,p_0\right)=E\left(E\left(\theta|\omega_1,p_0\right)|p_1,p_0\right)\Rightarrow p_1=E\left(\theta|p_1,p_0\right)$
- $PI=Var\left(E\left(\theta|p_1,p_0\right)\right)=Var(p_1)$

for PI(3): Considering the values of $p_1$and their probabilities in the different cases listed in the proof of Proposition 4, we calculate that $E\left(p_1\right)=0$ and $Var(p_1)=E\left(p_1^2\right)=\sigma^2\left(1+\frac{1}{3}\frac{x_{(3)}+y_{(3)}}{2}\right)$. As a result, $PI_{(3)}=\sigma^2\left(1+\frac{1}{3}\frac{x_{(3)}+y_{(3)}}{2}\right)$.