Uncertainty about What's in the Price

Joël Peress and Daniel Schmidt*

February 14, 2019

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 uptick, buy volume—because it may result from speculators trading on stale news—has a lower price impact compared to sell volume (and vice versa after recent price downticks). 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

*PRELIMINARY and INCOMPLETE; comments very welcome! 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.

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, most of the theoretical literature on this subject arguably relies on an 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 very profitable to trade—or do they possess *stale* information that is already reflected in the price? Such type of uncertainty should be very common. After all, prices can move for a myriad of reasons

[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).

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 glean information that they contemplate trading on, they will not know whether other investors have seen and traded on this information before and thus how novel it truly is. In this paper, we put forth a parsimonious trading model in which investors face this type of uncertainty and describe the resulting trading equilibrium.

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 the most recent price change 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, that 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 a recent price downturn (upturn).

The second insight involves market makers' assessment of adverse selection risk and is, to the best of our knowledge, new to the literature: after a prior positive (negative) price update, market makers judge 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. In fact, we argue that this 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.

The microstructure literature has identified inventory risk concerns as an important

[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.

source of order book asymmetry (e.g., Ho and Stoll, 1981; Madhavan and Smidt, 1993; Hendershott and Menkveld, 2014). In particular, when market makers are loath to deviate from a given target inventory level, they will naturally require a bigger price concession for accommodating an order that pushes their inventory further away from target as compared to an order that allows them to move their inventory towards the target. Hence, with inventory risk concerns, buy (sell) orders that come after buy orders (i.e., when market makers' inventory is likely to be above target) are expected to come with a larger (lower) price impact. Our model—derived in the absence of inventory risk concerns—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, inventory risk concerns and uncertainty about what's in the price are likely to co-exist. Which of these forces prevails to explain observed liquidity dynamics is thus an empirical question.

We shed a first 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. We find strong evidence for asymmetric price impact costs in a manner consistent with our model. 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 when the prior-day return was 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 the prior-day return was negative. We show that this contrasting pattern is not an artifact of negative return autocorrelations, leaving our mechanism—uncertainty about what's in the price—as its most likely explanation. Indeed, we are not aware of any other theory that could explain this finding.

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 with and without controls, 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 updates 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—sets our model apart 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. They further 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. In contrast, we allow for the possibility that investors don't know which pieces of information were observed by others. Moreover, while our model is in many ways simpler than Abreu and Brunnermeier (2002) and Abreu and Brunnermeier (2003)—e.g., our model is static and there is no coordination motive among investors—it is richer in at least one important dimension: we endogenize the stock price, allowing us to focus on how investors update the extent of their information advantage based on past price movements.

Finally, our paper is related to Tetlock (2010), 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 often 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 preliminary empirical evidence in favor of the model's predictions.

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 is a single stock that pays a dividend of $\theta = \theta_1 + \theta_2$ at date t=2. We assume that θ_1 and θ_2 are independent and both pay off $+\sigma$ or $-\sigma$ with equal probability. Hence, θ is -2σ with probability 25%, 0 with probability 50%, or $+2\sigma$ with probability 25%.

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 informed speculators and no noise traders. Market makers observe a part of the fundamental θ_m where $m \in \{1, 2\}$ with equal probability and then set $p_0 = E(\theta|\theta_m) = \theta_m$. 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.[3] m and s are drawn independently, implying that $m = s$ and $m \neq s$ occur with 50% probability each. At date t=1, there are also noise traders that trade a random amount n distributed according to a uniform distribution, $n \sim U[-1, +1]$. Noise traders and S submit their (market) orders and M set prices conditional on p_0 (that is, θ_m) and the aggregate order flow at date 1.

[3] The assumption that all speculators observe the same part of the fundamental is not crucial for our argument. Indeed, the model's key intuition about investors' updating on the novelty of their signal based on past price movements remains intact 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 $j \neq i$.

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 by M and S, respectively, and whether they know what the others know are central elements to the model. Figure 1 summarizes the model setup.

2.2 Solution when 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 θ_m $(= \theta_s)$. S then have no information advantage over M and thus refrain from trading at t=1. We therefore have $p_0 = p_1 = \theta_m$.

Next, consider the case $m \neq s$. In this case, M do not know the exact realization of θ_s, but they know that S have an information advantage that S will trade on. M will then try to back out θ_s from the order flow. We conjecture that investors S trade in a symmetric fashion and will buy x (sell $-x$) with $|x| < 1$ when they know $\theta_s = \sigma$ $(\theta_s = -\sigma)$. Hence, the order flow is $\omega_1 = x + n$ when $\theta_s = \sigma$ and $\omega_1 = -x + n$ when $\theta_s = -\sigma$. When M observe $\omega_1 > -x + 1$ $(\omega_1 < x - 1)$, they infer that $\theta_s = \sigma$ $(\theta_s = -\sigma)$. When $x - 1 \leq \omega_1 \leq -x + 1$, it is both possible that S bought or sold and M learns nothing about θ_s. The equilibrium price function thus equals:

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

Each investor i from set S takes the price function as given and chooses the x_i that maximize expected utility. Imposing rational expectations (i.e., $x_i = x$ for all i in S) on the usual first-order condition gives the following equilibrium condition:

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

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

Proposition 2. *Assume that m and s are common knowledge. At t=0, market makers M set $p_0 = \theta_m$. When $m = s$, speculators S refrain from trading at t=1 and $p_1 = p_0 = \theta_m$. When $m \neq s$, speculators S buy (sell) an amount x ($-x$) when $\theta_s = \sigma$ ($\theta_s = -\sigma$) with*

$$x = \sqrt{\frac{1}{\gamma\sigma}}$$

and p_1 is given by (1).

Intuitively, speculators S trade less aggressively when they are more averse to risk (γ larger) and when the final payoff is more uncertain (σ larger). Note that $x < 1$ by Assumption 1 and hence the order flow ω_1 is not fully revealing about θ_s, implying that speculators S derive positive expected utility from trading. The solution resembles Vives (1995), who also model trading by a continuum of risk-averse speculators in the presence of competitive market makers (but with normally-distributed payoff and noise).

2.3 Solution when 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 known 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 about whether (better) informed speculators are present or not.

When $m = s$, S have no information advantage over M and thus refrains from trading ($\omega_1 = n$). When $m \neq s$, S have additional information about θ and are conjectured to buy (sell) an amount x ($-x$) when $\theta_s = \sigma$ ($\theta_s = -\sigma$). Market makers M do not know, however, which of these cases has occurred and just try to learn from the order flow. When M observe $\omega_1 > 1$ ($\omega_1 < -1$), they infer that $\theta_s = \sigma$ ($\theta_s = -\sigma$). When $-x + 1 \leq \omega_1 \leq 1$ ($-1 \leq \omega_1 < x - 1$), M know that S did not sell $-x$ (buy x). 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 $\frac{1}{3}\sigma$ ($-\frac{1}{3}\sigma$).[4] Finally, when $x - 1 \leq \omega_1 \leq -x + 1$, M learns nothing about θ_s. The equilibrium price function thus equals:

$$p_1 = \begin{cases} \theta_m + \sigma & \text{for} \quad 1 < \omega_1 \leq x + 1 \\ \theta_m + \frac{1}{3}\sigma & \text{for} \quad -x + 1 < \omega_1 \leq 1 \\ \theta_m & \text{for} \quad x - 1 \leq \omega_1 \leq -x + 1 \\ \theta_m - \frac{1}{3}\sigma & \text{for} \quad -1 \leq \omega_1 < x - 1 \\ \theta_m - \sigma & \text{for} \quad -x - 1 \leq \omega_1 < -1 \end{cases} \tag{2}$$

Given this price function, S choose the 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 $[x - 1, x + 1]$. It is then easy to calculate:

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

Plugging into the first-order condition and imposing rational expectations (i.e., $x_i =$

[4] This follows from noting that the probability of $\theta_s = \sigma$ and $m \neq s$ equals $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$, whereas the probability for $m = s$ is $\frac{1}{2}$. Hence, $E(\theta_s| -x + 1 \leq \omega_1 \leq 1) = \frac{1/4}{1/4+1/2}\sigma = \frac{1}{3}\sigma$.

x for all i in S), we can solve for speculators' trading aggressiveness x. The following proposition summarizes the resulting equilibrium.

Proposition 3. *Assume that only m is common knowledge. At t=0, market makers M set $p_0 = \theta_m$. At t=1, speculators S refrain from trading when $m = s$ and buy (sell) an amount x ($-x$) when $m \neq s$ and $\theta_s = \sigma$ ($\theta_s = -\sigma$), where x is the unique solution of the cubic equation*

$$9 - 6x - 5\gamma\sigma x^2 + 4\gamma\sigma x^3 = 0$$

in the range $0 < x < 1$.[5] *Market makers M don't know whether $m = s$ or $m \neq s$ and set p_1 according to (2).*

As before, the equilibrium trading aggressiveness 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.

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.

[5] It is in fact here where Assumption 1 becomes binding. That is, for $\gamma\sigma < 3$, there does not exist a solution to the equation with $0 < x < 1$, implying that there is no equilibrium in which the order flow is not fully revealing.

2.4 Solution when 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 2.3, 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' S signal coincides with Ms' signal as revealed by the t=0 price ($\theta_m = \theta_s$), then investors S are unsure whether they observed the same part of 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 ($-x$) when $\theta_m \neq \theta_s$ and that they buy (sell) an amount y ($-y$) 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 (when $\theta_s = \sigma$) or sell $-x$ (when $\theta_s = -\sigma$). When $-x+1 < \omega_1 \leq y+1$, M infer that S bought y. The conditional expected value of $\theta_{\neg m}$ is then $\frac{1}{3}\sigma$ (see footnote 1). When $-x-1 \leq \omega_1 < y-1$, M know that S sold x. This implies $\theta_{\neg m} = -\sigma$. When $y-1 \leq \omega_1 \leq -x+1$, it is both possible that S bought or sold and thus M learns nothing from the order flow. The price function therefore is:

$$p_1(\omega_1|\theta_m = +\sigma) = \begin{cases} \frac{4}{3}\sigma & \text{for} \quad -x+1 < \omega_1 \leq y+1 \\ \sigma & \text{for} \quad y-1 \leq \omega_1 \leq -x+1 \\ 0 & \text{for} \quad -x-1 \leq \omega_1 < y-1 \end{cases} \tag{3}$$

For the case $\theta_m = -\sigma$, the logic is reversed. S either buy x or sell $-y$ and M draw analogous inferences from the order flow. We obtain:

$$p_1(\omega_1|\theta_m = -\sigma) = \begin{cases} 0 & \text{for} \quad -y+1 < \omega_1 \leq x+1 \\ -\sigma & \text{for} \quad x-1 \leq \omega_1 \leq -y+1 \\ -\frac{4}{3}\sigma & \text{for} \quad -y-1 \leq \omega_1 < x-1 \end{cases} \tag{4}$$

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 suggests that the speculators have traded 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 x and y. 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, resulting in an order flow drawn at random from the interval $[x-1, x+1]$. The investor then calculates:

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

When $\theta_s = -\sigma$, investors expect other investors in S to sell an amount $-y$, resulting in an order flow drawn at random from the interval $[-y-1, -y+1]$. The investor then calculates:[6]

$$\begin{aligned} E(\theta - p_1|\theta_m = -\sigma, \theta_s = -\sigma) &= -\frac{4}{3}\sigma - \left(-\sigma - \frac{1}{3}\frac{x+y}{2}\sigma\right) = -\frac{1}{3}\sigma\left(1 - \frac{x+y}{2}\right) \\ Var(\theta - p_1|\theta_m = -\sigma, \theta_s = -\sigma) &= \sigma^2\left(\frac{8}{9}\frac{x+y}{2} + \left(1 - \frac{x+y}{2}\right)\left(1 - \frac{1}{9}\left(1 - \frac{x+y}{2}\right)\right)\right) \end{aligned}$$

Investors first-order condition together with requiring rational expectations for the two

[6]Details for the derivation of this conditional portfolio variance are provided in Appendix A.1.

cases $\theta_s = \sigma$ and $\theta_s = -\sigma$ (i.e., $x_i = x$ in the former case and $x_i = y$ in the latter) yields a system of two equations that can be solved for x and y. 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, speculators S buy (sell) an amount x $(-x)$ when $\theta_m \neq \theta_s$ and $\theta_s = \sigma$ $(\theta_s = -\sigma)$ and they buy (sell) an amount y $(-y)$ when $\theta_m = \theta_s$ and $\theta_s = \sigma$ $(\theta_s = -\sigma)$, where x is the unique root in the range $0 < x < 1$ of the quartic equation*

$$1 - \gamma\sigma x - 2\gamma\sigma(1 + 4\gamma\sigma)x^2 + 2(\gamma\sigma)^2 x^3 + 4(\gamma\sigma)^3 x^4 = 0$$

and $y = \frac{2-\gamma\sigma x^2}{\gamma\sigma x}$. The price function p_1 depends on the realization of θ_m and is given by either (3) or (4), respectively.

Both x and y 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, it is easy to show that we have $x > y$ in equilibrium. Hence, as conjectured, 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.

2.5 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:* $0 < y_{(3)} < x_{(1)} < x_{(3)} < x_{(2)} < 1$.[7]

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).[8] When $\theta_s = \theta_m$, speculators are 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 uptick (downtick), 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

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

[8]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).

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 for the US stock market.

Finally, we examine the price informativeness for the three different equilibria. Following Kyle (1985), we measure price informativeness by $PI \equiv E\left[Var\left(\theta|p_1, p_0\right)\right]$. The smaller this measure, the more information is incorporated into the price, 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:*

$$\begin{aligned} PI_{(1)} &= \sigma^2\left(1-\frac{1}{2}x_{(1)}\right) \\ PI_{(2)} &= \sigma^2\left(1-\frac{1}{3}x_{(2)}\right) \\ PI_{(3)} &= \sigma^2\left(1-\frac{5}{36}\left(x_{(3)}+y_{(3)}\right)\right) \end{aligned}$$

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 second 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 so that price informativeness decreases.

3 How important is UWIP?

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 calculate price impact costs in the following way. First, we sign all TAQ trades using the Lee and Ready (1991) algorithm. To obtain 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. Finally, we define the price impact costs for stock i on date t as

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

where the total return comes from CRSP. Note that this definition is reminiscent of the Amihud (2002) illiquidity measure, with the important difference that our measure is based on the *signed* trade imbalance (rather than dollar volume) and thus uses the total return (rather than its absolute value) in the nominator.[9] We winsorize price impact at

[9]With price impact, buys ($\text{trade imbalance}_{it} > 0$) are expected to result in a positive return, while

the 1% level on both sides in order to mitigate the effect of outliers (which can obtain when the trade imbalance is close to zero).

Our model predicts that the price impact costs for buys and sells depend on prior returns in opposite ways: whereas the price impact of buys should decrease when the prior-day return was positive (since these buys are more likely to be triggered by stale news), the price impact of sells should increase in this case (since these sells are then more likely to be triggered by genuine, i.e. non-stale, news). To test this prediction, we regress our price impact measure on the prior-day stock return separately for days with positive and negative net-buying activity:

$$\text{price impact}_{it} = \alpha_i + \alpha_t + \beta_{buys}\,\text{return}_{it-1} + \epsilon_{it} \quad \text{if} \quad \text{trade imbalance}_{it} > 0$$

$$\text{price impact}_{it} = \alpha_i + \alpha_t + \beta_{sells}\,\text{return}_{it-1} + \epsilon_{it} \quad \text{if} \quad \text{trade imbalance}_{it} < 0$$

where α_i and α_t are stock and date fixed effects, respectively. Based on our model, we expect $\beta_{buys} < 0$ and $\beta_{sells} > 0$.

One problem with our approach is that it is potentially confounded by the negative autocorrelation of returns observed in individual stock return data. Indeed, a negative return yesterday predicts a positive return today, which enters our price impact measure in the nominator. Since the denominator of this measure is by definition positive (negative) in the sample of days with positive (negative) trade imbalance, this alone can explain why one may find $\beta_{buys} < 0$ and $\beta_{sells} > 0$. To mitigate this concern, we refine our price impact measure by first running a regression of returns on lagged returns (while including stock and date fixed effects) and then using the regression residuals in the nominator of our price impact measure. As a robustness test, we repeat our price impact regressions for this autocorrelation-corrected price impact measure.

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

sells (trade imbalance$_{it} < 0$) result in a negative return. Hence, for both buys and sells, a more positive price impact measure indicates higher price impact costs.

activity. For better visibility, the price impact measures 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 0.69%. The table further shows that the price impact is on average slightly higher on days with net selling activity, as compared to days with net buying activity.

3.2 Test results

Panel (a) of Table 2 shows the results for the basic (i.e., not autocorrelation-corrected) price impact measure. In line with the model's prediction, we find that on days with a positive trade imbalance, price impact costs are significantly negatively related to the prior-day return (column 1); while on days with a negative trade imbalance, they are significantly positively related to the prior-day return (column 3). The inclusion of stock and date fixed effects, if anything, only increases the economical and statistical significance of these findings (columns 2 and 4). In terms of economic magnitude, a one-standard deviation increase in the lagged return decreases (increases) the price impact on days with a positive (negative) net trade imbalance by about 10% of its standard deviation.

In Panel (b), we show the results for the autocorrelation-corrected price impact measure. This is an important robustness check since our definition of price impact together with the existence of a negative autocorrelation in returns can mechanically lead to the observed correlation pattern. Indeed, the results become somewhat weaker after correcting for the impact of return autocorrelation, but they remain statistically highly significant. In terms of economic magnitude, a one-standard deviation increase in the lagged return decreases (increases) the autocorrelation-corrected price impact on days with a positive (negative) net trade imbalance by about 6% of its standard deviation.

Overall, these results 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 (apart from return autocorrelation for which we control) that could explain why price impact costs for buys and sells are asymmetrically related to past price movements as we find.

4 UWIP and price informativeness

In this section, we check 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 should be associated with less informative stock prices. We test this prediction in the context of earnings announcements. Specifically, we follow Weller (2018) and construct a measure of the price jump ratio around earnings announcement dates, which we then regress on a self-constructed proxy for the extent of uncertainty about what's in the price.

The price jump ratio is designed to capture the fraction of 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): by smoothly trading on their information, informed investors ensure that prices drift toward the post-announcement asset value. Competition among informed traders only accelerates this process, so that even more information is 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 thus not suitable for our purpose.

We construct the price jump ratio as described in Weller (2018). Here, we provide a brief summary of the approach; see his paper for more detail. We start from the sample of quarterly earnings announcement over years 1995 to 2012. 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 keep 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 for uncertainty about what's in the price builds on the intuition described in Section 2.5: sell volume should come with a larger price impact when the prior day's return was high since it is then unlikely that speculators are trading on stale news. To capture this effect, we estimate for each event stock the following interaction specification using daily data over the calendar quarter prior to the earnings announcement date:

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

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 trade $\text{imbalance}_{i\tau} < 0$) 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.[10] Our model predicts that $\beta_3 > 0$ and hence a higher coefficient estimate indicates a larger uncertainty about what's in the price. Because the β_3 estimates are fat-tailed and skewed, we transform it into percentiles. Thus, our measure of choice,uwip_{it}, takes on values from 1 to 100 depending on the corresponding percentile 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 past returns, volatility, and turnover; accounting variables; analyst coverage information; as well as institutional ownership data (all control variables are defined in the header of Table 3 below).

4.2 Test results

Panel (a) of Table 3 shows the results when UWIP is estimated using the basic (i.e., not autocorrelation-corrected) price impact measure. Looking at column (1), we find a significantly positive effect of UWIP on the price jump ratio, implying that uncertainty

[10] Clearly, our interaction specification is closely related to the sample split approach from Section 3. Indeed, when we run this specification for the overall sample (and including stock and date fixed effects), we obtain $\hat{\beta}_3 = -31.14$ (t-statistic of -66), which is almost exactly equal to $\beta_{sells} - \beta_{buys}$ as reported in Panel (a) of Table 2.

about what's in the price is associated with a lower stock price informativeness as predicted by our model. In terms of economic magnitude, moving from the 1st to the 100th percentile of UWIP increases the price jump ratio by about 2 percentage points, or about 6% relative to the unconditional mean (35%) of the price jump ration in our sample. Subsequently adding stock market controls, accounting controls, analyst coverage controls, institutional ownership controls, or even industry×year fixed effects does not alter this picture: the coefficient estimate for UWIP barely changes and always remains statistically significant at the 5% level.

Panel (a) of Table 3 shows the results when UWIP is estimated using the autocorrelation-corrected price impact measure (see Section 3.1 for details). The results using this refined UWIP measure are even stronger than those from before, in line with the idea that correcting for autocorrelation improves the signal-to-noise ratio in the UWIP measure. Statistically, the coefficient estimate for UWIP is now always significant at the 1% level. Economically, moving from the 1st to the 100th percentile of UWIP now increases the price jump ratio by about 3 percentage points, or about 9% relative to its unconditional sample mean. Overall, these results confirm our prediction that uncertainty about what's in the price 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 upticks (downticks), 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 the prior-day's stock return. These results are not explained by the autocorrelation in returns and therefore 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.

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), 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.

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., 2010. All the news that's fit to reprint: Do investors react to stale information?, wP.

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.

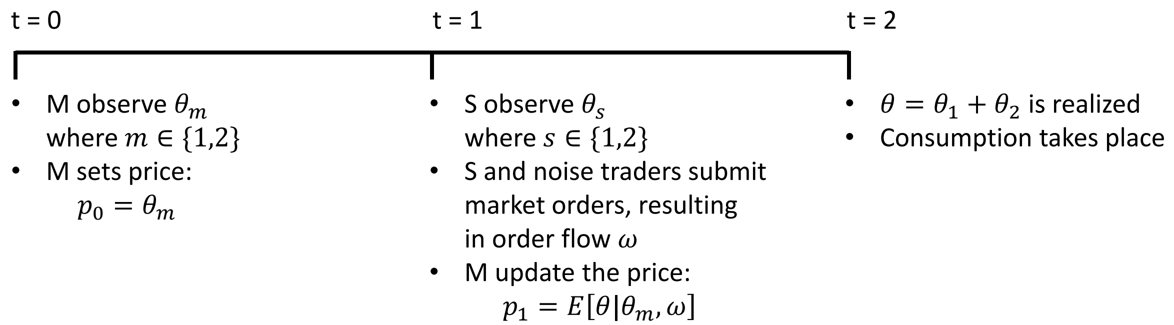
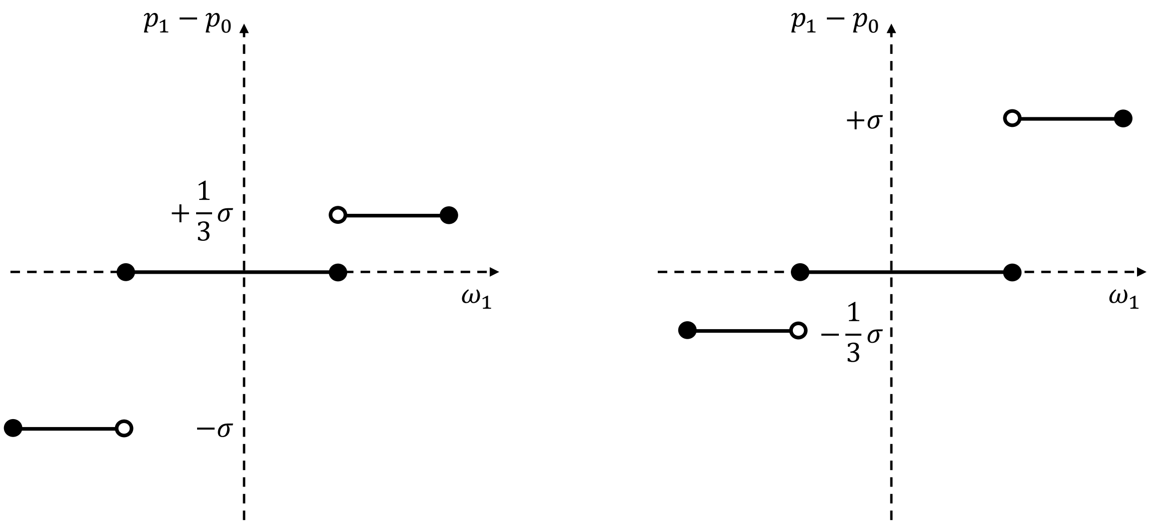


Figure 2: Equilibrium price function

This figure shows the equilibrium price function when the speculator faces uncertainty about whether his trading signal is already in the price. In Panel A, we show the price function for the case of prior positive news. In Panel B, we show the price function for prior negative news.



Panel A: For $\theta_m = +\sigma$

Panel B: For $\theta_m = -\sigma$

Table 1: **Price Impact Statistics**

This table reports descriptive statistics for the basic and autocorrelation-corrected price impact measures in the overall sample, as well as separately for days with positive and negative net trade imbalance. Price impact for a given stock-day is defined as the stock's total return over the net trade imbalance calculated from TAQ data. The autocorrelation-corrected price impact is calculated in two steps. First, we regress returns on lagged returns in a panel setting: $\text{return}_{it} = \alpha_i + \alpha_t + \beta\,\text{return}_{it-1} + \epsilon_{it}$. Second, autocorrelation-corrected price impact is defined as the residual $\hat{\epsilon}_{it}$ from this regression over the net trade imbalance. Both price impact measures are multiplied by 10^6 for better visibility. Lagged return is the stock's total return on the previous trading day. 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	0.6938	0.0055	6.0205
Price Impact (autocorrelation-corrected)	0.6839	0.0073	6.2622
Lagged Return	0.0007	0.0000	0.0356
Net Buys (N =11,928,063)			
Price Impact	0.6758	0.0040	6.0554
Price Impact (autocorrelation-corrected)	0.5766	0.0032	6.2558
Lagged Return	0.0023	0.0000	0.0352
Net Sells (N =12,532,851)			
Price Impact	0.7109	0.0078	5.9871
Price Impact (autocorrelation-corrected)	0.7860	0.0147	6.2666
Lagged Return	-0.0009	0.0000	0.0358

Table 2: **Price Impact Tests**

This table reports the results from regressing price impact on the lagged stock return as explained in Section 3.1. Panel (a) shows the results for the basic price impact measure as the dependent variable; Panel (b) shows the results for the autocorrelation-corrected price impact as the dependent variable. Price impact for a given stock-day is defined as the stock's total return over the net trade imbalance calculated from TAQ data. The autocorrelation-corrected price impact is calculated in two steps. First, we regress returns on lagged returns in a panel setting: $\text{return}_{it} = \alpha_i + \alpha_t + \beta\,\text{return}_{it-1} + \epsilon_{it}$. Second, autocorrelation-corrected price impact is defined as the residual $\hat{\epsilon}_{it}$ from this regression over the net trade imbalance. Both price impact measures are multiplied by 10^6 for better visibility. In columns (1) and (2), the regression of price impact on lagged return is run for the sample of stock-days in which the net trade imbalance is positive. In columns (3) and (4), the regression is run for the sample of stock-days in which the net trade imbalance is negative. Columns (2) and (4) contain stock and day 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): Price Impact as Dependent Variable

	Price Impact			
	Net Buys		Net Sells	
	(1)	(2)	(3)	(4)
Lagged Return	-15.4061***	-17.6882***	15.1528***	16.9014***
	(-61.52)	(-69.74)	(60.84)	(68.36)
Stock-FE	No	Yes	No	Yes
Date-FE	No	Yes	No	Yes
N	11,928,064	11,928,032	12,532,852	12,532,838
adj. R^2	0.01	0.06	0.01	0.06

Panel (b): Autocorrelation-corrected Price Impact as Dependent Variable

	Price Impact (corrected for return autocorrelation)			
	Net Buys		Net Sells	
	(1)	(2)	(3)	(4)
Lagged Return	-9.9776***	-10.9920***	9.7190***	10.1925***
	(-51.27)	(-51.27)	(49.64)	(48.70)
Stock-FE	No	Yes	No	Yes
Date-FE	No	Yes	No	Yes
N	11,928,063	11,928,032	12,532,851	12,532,838
adj. R^2	0.00	0.04	0.00	0.07

Table 3: **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: $\text{price impact}_{i\tau} = \beta_0 + \beta_1 \,\text{return}_{i\tau-1} + \beta_2 \,\text{netsell}_{i\tau} + \beta_3 \,\text{return}_{i\tau-1} \times \text{netsell}_{i\tau} + \epsilon_{i\tau}$, 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 $\beta_3 > 0$; we thus define UWIP as the percentile value of β_3 (implying that UWIP takes on values from 1 to 100). Panel (a) shows the results for estimating UWIP using the basic price impact measure; Panel (b) shows the results for estimating UWIP using the autocorrelation-corrected price impact measure (see Section 3.1 for details). Stock market controls include the average stock return, the standard deviation of returns, and the share turnover (all measured in the quarter prior to the announcement). Accounting controls include firm size (i.e., log of total assets), age (i.e., years since inception in Compustat), leverage, return on assets, and book-to-market ratio (all taken from the end of the business year preceding the earnings announcement). Coverage controls include the reporting lag, the number of analysts covering the firm, and the analyst forecase dispersion (measured at the end of the month prior to the announcement). Ownership controls include the fraction of institutional ownership and a institutional ownership concentration index (measured at the end of the quarter prior to the announcement). All regressions contain stock and day fixed effects; column (6) further includes industry×year fixed effects based on the SIC-2 digit industry classification. 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 as key Independent Variable

	Price Jump Ratio					
	(1)	(2)	(3)	(4)	(5)	(6)
UWIP	0.0002**	0.0002**	0.0002**	0.0002**	0.0002**	0.0002**
	(2.26)	(2.29)	(2.51)	(2.02)	(2.15)	(2.33)
Stock mkt contr.	No	Yes	Yes	Yes	Yes	Yes
Accounting contr.	No	No	Yes	Yes	Yes	Yes
Coverage contr.	No	No	No	Yes	Yes	Yes
Ownership contr.	No	No	No	No	Yes	Yes
Stock-FE	Yes	Yes	Yes	Yes	Yes	Yes
Date-FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year-FE	No	No	No	No	No	Yes
N	64,907	64,907	63,668	52,770	52,663	52,614
adj. R^2	0.09	0.09	0.09	0.08	0.08	0.09

Panel (b): Autocorrelation-corrected UWIP as key Independent Variable

	Price Jump Ratio					
	(1)	(2)	(3)	(4)	(5)	(6)
UWIP	0.0003***	0.0003***	0.0003***	0.0003***	0.0003***	0.0003***
(ac-corrected)	(3.56)	(3.59)	(3.64)	(3.07)	(3.08)	(3.36)
Stock mkt contr.	No	Yes	Yes	Yes	Yes	Yes
Accounting contr.	No	No	Yes	Yes	Yes	Yes
Coverage contr.	No	No	No	Yes	Yes	Yes
Ownership contr.	No	No	No	No	Yes	Yes
Stock-FE	Yes	Yes	Yes	Yes	Yes	Yes
Date-FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry×Year-FE	No	No	No	No	No	Yes
N	64,907	64,907	63,668	52,770	52,663	52,614
adj. R^2	0.09	0.09	0.09	0.08	0.08	0.09

A Appendix

A.1 Portfolio variance for the case $\theta_m = \theta_s = -\sigma$

Consider the case $\theta_m = \theta_s = -\sigma$. In that case, S are not sure whether $m = s$ (i.e., whether their signal is stale) or whether $m \neq s$ (i.e., whether their signal is novel). Based on the conjecture described in Section 2.4, investors S are expected to sell an amount $-y$, resulting in an order flow drawn at random from the interval $[-y-1, -y+1]$. The investor then calculates:

$$
\begin{aligned}
Var(\theta - p_1|\theta_m = \theta_s = -\sigma) &= E((\theta - p_1)^2|\theta_m = \theta_s = -\sigma) - [E(\theta - p_1|\theta_m = \theta_s = -\sigma)]^2 \\
&= Pr(m = s|\theta_m = \theta_s = -\sigma)\, E((\theta - p_1)^2|\theta_m = \theta_s = -\sigma, m = s) \\
&\quad + Pr(m \neq s|\theta_m = \theta_s = -\sigma)\, E((\theta - p_1)^2|\theta_m = \theta_s = -\sigma, m \neq s) \\
&\quad - [E(\theta|\theta_m = \theta_s = -\sigma) - E(p_1|\theta_m = \theta_s = -\sigma)]^2 \\
&= \frac{2}{3}\left[\begin{array}{l} \frac{1}{2}E((\theta - p_1)^2|\theta_m = \theta_s = -\sigma, m = s, \theta_{\neg m} = -\sigma) \\ +\frac{1}{2}E((\theta - p_1)^2|\theta_m = \theta_s = -\sigma, m = s, \theta_{\neg m} = -\sigma) \end{array}\right] \\
&\quad + \frac{1}{3}\left[E((\theta - p_1)^2|\theta_m = \theta_s = -\sigma, m \neq s)\right] \\
&\quad - [E(\theta|\theta_m = \theta_s = -\sigma) - E(p_1|\theta_m = \theta_s = -\sigma)]^2 \\
&= \frac{2}{3}\left[\begin{array}{l} \frac{1}{2}\left(\sigma^2\left(1 - \frac{x+y}{2}\right) + \left(\frac{4}{3}\sigma\right)^2 \frac{x+y}{2}\right) \\ \frac{1}{2}\left(\sigma^2\left(1 - \frac{x+y}{2}\right) + \left(\frac{2}{3}\sigma\right)^2 \frac{x+y}{2}\right) \end{array}\right] \\
&\quad + \frac{1}{3}\left(\sigma^2\left(1 - \frac{x+y}{2}\right) + \left(\frac{2}{3}\sigma\right)^2 \frac{x+y}{2}\right) \\
&\quad - \left(\frac{1}{3}\sigma\right)^2\left(1 - \frac{x+y}{2}\right)^2 \\
&= \frac{1}{3}\left(\sigma^2\left(1 - \frac{x+y}{2}\right) + \left(\frac{4}{3}\sigma\right)^2 \frac{x+y}{2}\right) \\
&\quad + \frac{2}{3}\left(\sigma^2\left(1 - \frac{x+y}{2}\right) + \left(\frac{2}{3}\sigma\right)^2 \frac{x+y}{2}\right) \\
&\quad - \left(\frac{1}{3}\sigma\right)^2\left(1 - \frac{x+y}{2}\right)^2 \\
&= \sigma^2\left(\frac{8}{9}\frac{x+y}{2} + \left(1 - \frac{x+y}{2}\right)\left(1 - \frac{1}{9}\left(1 - \frac{x+y}{2}\right)\right)\right)
\end{aligned}
$$

A.2 Price informativeness

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

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

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:

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

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{11}{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{13}{18}\sigma^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right) \\
&= \sigma^2\left(1-\frac{5}{18}\frac{x+y}{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. If $x=y=1$ were possible, then there would be a further reduction of total uncertainty of about 13.88%.