Journal of Financial Economics

Uncertainty about what is in the price[☆]

Joël Peress [a], Daniel Schmidt [b,*]

[a] INSEAD, Boulevard de Constance, Fontainebleau, 77300, France
[b] HEC Paris, 1 Rue de la Libération, Jouy-en-Josas, 78350, France

ARTICLE INFO

JEL classification:
G11
G14

Keywords:
Strategic trading
Learning from prices
Price impact

ABSTRACT

A critical question facing speculators contemplating to trade on private information is whether their signal has already been priced in by the market. In our model, speculators assess the novelty of their information based on recent price movements, and market makers are aware that speculators might be trading on stale news. An asymmetric response to past price movements ensues: after price increases, buy volume – because it may result from stale news trading – has a lower price impact than sell volume (and vice versa after price decreases). Consequently, return skewness is negatively related to lagged returns. We find strong support for these and other predictions using a comprehensive sample of US stocks.

1. Introduction

Asset prices reflect information. Yet, in a complex world it is seldom clear *whether* a given piece of information is already reflected in the price or 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 seems unrealistic 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 mostly looked at the asset pricing implications of the uncertainty that results when *uninformed* investors are not sure about the presence of informed investors (and thus about the importance of adverse selection). 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 is 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.

☆ Dimitris Papanikolaou was the editor for this article. We thank Dion Bongaerts, Pierre Collin-Dufresne, Xi Dong, Thierry Foucault, Xavier Gabaix, Sergei Glebkin, Jim Goldman, Brett Green, Johan Hombert, Augustin Landier, Jeongmin Lee (WFA discussant), Brandon Han (AFA discussant), Sophie Moinas (FutFinInfo discussant), Lin Peng (SFFS discussant), Shiyang Huang (EWFC discussant), Francesco Sangiorgi, Esad Smajlbegovic, Savitar Sundaresan, Bastian von Beschwitz, Bart Yueshen, Junyuan Zou, as well as two anonymous referees and seminar/conference participants at Aarhus University, Erasmus University, Frankfurst School of Finance & Management brownbag, HEC Paris brownbag, Imperial College Business School, University of Lausanne-EPFL, the 2020 European Winter Finance Conference, the 2023 FutFinInfo Conference, the 2020 Western Finance Association Meeting, the 2022 American Finance Association Meeting, and the 4th Shanghai Financial Forefront Symposium. This research is supported by a grant of the French National Research Agency (ANR), "Investissements d'Avenir" (LabEx Ecodec/ANR-11-LABX-0047).

[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).

Received 29 July 2022; Received in revised form 1 August 2024; Accepted 3 August 2024
Available online 24 August 2024

Hence, when investors contemplate trading on a signal, they will not know how many other investors have traded on this information before 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—which we dub *uncertainty about what is in the price* (henceforth UWIP)—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) (albeit not in an equilibrium framework): 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 price 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 price 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, and so moderate its price impact. A consequence is that the skewness of returns depends negatively on lagged returns. Indeed, after price increases, market makers lower prices more in response to sell orders than they raise them in response to buy orders. Moreover, because speculators might actually be trading on stale (positive) news, buy orders are more likely than sell orders after price increases. Both effects contribute to making returns more negatively skewed. Conversely, returns are more positively skewed after price decreases, both because market makers raise prices more in response to buy orders than they reduce them in response to sell orders, and because buy orders are less likely. Hence our model predicts that equilibrium prices, return skewness, and trading strategies are asymmetric across buy and sell orders and depend on prior price movements. In several extensions, we show that this asymmetry in the price response is robust to different trading protocols (it obtains with market and limit orders, in order- and quote-driven markets), learning from prices, and investor risk aversion. We also explore the persistence of UWIP in a dynamic (Glosten and Milgrom, 1985) setting. We find that the effects of UWIP dissipate over time when signals are assumed to be uncorrelated across speculators, while they persist if signals are correlated. Importantly, this version of our model underscores that UWIP always (i.e., regardless of the assumed signal distribution) matters right after speculators see a price move without knowing precisely which piece of news caused that move.

We argue that the type of asymmetric dependence which we describe here arises naturally when there is UWIP, making it a distinct "footprint" to look for in empirical data. Indeed, the literature to date has primarily focused on two different sources of asymmetry in 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 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 associated 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 predictions regarding UWIP prevail in the data.

We shed light on this question by examining an exhaustive sample of NYSE-traded stocks from 1993 to 2014. Starting with skewness, we find that the daily skewness of stock returns (estimated from intraday TAQ data) is negatively related to lagged returns, consistent with the model. This phenomenon is economically meaningful, as a one-standard deviation (1-SD) increase in lagged returns decreases skewness by about 9% of a SD. Moreover, in contrast to alternative mechanisms (discussed below) that rely on short sale constraints, this relationship holds for both positive and negative lagged returns, and is insensitive to short selling costs (which proxy for the tightness of short sale constraints). Turning to the price impact predictions, we compare price impact costs on days with net-buying and net-selling activity as a function of past returns. Using the Lee and Ready (1991) algorithm to infer trade direction, we compute daily measures of trade imbalances from intraday TAQ data. We document that, on days with net-buying activity, price impact costs (measured using four distinct proxies that reflect adverse selection) are negatively related with past returns, while on days with net-selling activity price impact costs are positively related with past returns.[3] Put differently, buys elicit a lower price impact 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. This phenomenon is again economically meaningful. For instance, a 1-SD increase in lagged returns decreases (increases) price impact costs on days with a positive (negative) trade imbalance by about 7%–8% of a SD, thus driving a wedge between buy- and sell-days of about 15% of a SD.

Next, we check whether our previous findings weaken when UWIP is arguably lower. We report four sets of results. First, we find that the asymmetric patterns in skewness and price impact are considerably less pronounced immediately after earnings announcements, when investors know better what information is already impounded in stock prices. Second, they are weaker for stocks with more public information; that is, for larger stocks, stocks with more analyst coverage, and stocks recently added to the S&P 500 index (which are known to receive more public scrutiny, see Denis et al., 2003). Third, we find that our results for skewness and price impact are more pronounced for "complicated" firms (Cohen and Lou, 2012), i.e., firms that operate in multiple segments (relative to their size). Fourth, the asymmetry in price impact between buys and sells is weaker for retail order flow (identified using the Boehmer et al., 2021 methodology), consistent with the intuition that market makers are less concerned with retail orders because they are less likely to possess novel private information. These tests confirm that non-public information (i.e., information whose degree of common knowledge is hard to ascertain), investor informedness, and information complexity play a central role in the phenomena we document.

Our model makes three additional predictions. First, when speculators are risk averse (as in our baseline model), UWIP makes them more reluctant to trade on their information, thereby reducing the information content of stock prices. The other predictions arise when risk is priced

---

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

[3] Our four price impact measures are: (i) a signed version of the Amihud (2002) illiquidity ratio, named *price impact costs*, defined as the ratio of a stock's daily return (adjusted for autocorrelation) over 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.

(i.e., when all agents are risk averse). Second, due to the ensuing risk discount in the price, UWIP makes return volatility depend negatively on past returns. Third, in the cross-section, stocks with higher UWIP earn higher expected returns (as compensation for risk). These predictions, however, are not unique to UWIP, and moreover, difficult to test. For instance, the volatility-return dependence—the so-called leverage effect—also arises in Banerjee and Green (2015). Empirically, the effect of UWIP on volatility and expected returns is hard to disentangle from adverse selection risk due to private information (e.g., Easley and O'hara, 2005). We leave this interesting but challenging task for future research, and focus here instead on testing the predictions that are unique to UWIP—i.e., those pertaining to how skewness and price impact respond asymmetrically to past returns.

Overall, our results indicate that uncertainty about what is in the price is a genuine concern for investors. Indeed, we are not aware of any other theory that *jointly* explains: (1) why return skewness depends negatively on past price movements; (2) why price impact costs depend on past price movements asymmetrically across buy and sell orders; and (3) why the dependence of both return skewness and price impact costs is consistently reduced after public announcements and for stocks in the limelight.

Related literature. Our paper contributes primarily to the theoretical literature on informed trading in financial markets, and more specifically, to the body of research relaxing the assumption that investors' information environment is common knowledge. Prior work finds that under multidimensional uncertainty—such as when the proportion of informed traders is unknown—traders might ignore their own information or delay acting on it, leading to herding (Avery and Zemsky, 1998) or to persistent mispricing (Abreu and Brunnermeier, 2002; Abreu and Brunnermeier, 2003). More recently, Gao et al. (2013), Banerjee and Green (2015), and Papadimitriou (2023) study traders who are uncertain about the proportion of informed traders. In Banerjee and Green (2015) in particular, learning about whether others are trading on informative signals or noise leads to prices that react asymmetrically to news about fundamentals and hence to returns that depend asymmetrically on lagged returns (as explained below, the asymmetry is different from what is predicted by our model). In Easley and O'Hara (1992), market makers are unsure whether speculators have observed a signal about the asset's fundamental. Importantly, the learning agents in all these papers are themselves uninformed; hence, they cannot use their own signal realization in combination with the price to update their beliefs about the information structure of the market. It is this interplay between an investor's own signal realization and recent price changes that lies at the heart of our model. Another line of research (Blume et al., 1994; Schneider, 2009) studies investors' use of trading volume data to learn about properties of other investors' private signals (i.e., their precision or correlation with other investors' signals). In contrast, our focus is on how investors use (endogenous) past price movements to determine the extent of their information advantage and their optimal trading intensity.

Our paper also relates to three other streams of research. The first is the literature on "technical analysis" (e.g., Brown and Jennings, 1989; Grundy and McNichols, 1989; Brunnermeier, 2005). In these models, past prices 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 our model, in contrast, investors use them to update on the probability that others have seen the same information before them (thereby rendering their signal 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. Saar (2001) makes a similar prediction in a model that preserves the common knowledge assumption but relies instead on a set of portfolio constraints that typically apply to mutual funds. Specifically, he assumes that informed investors cannot borrow, sell short, nor underdiversify (i.e., concentrate holdings on only a few stocks). After the price increases (decreases), informed traders are more (less) likely to own the stock—since they probably bought (sold) the stock on the past good (bad) news that increased (decreased) the price; as a result, their buys (sells) are more constrained by the diversification (short sale) constraint, reducing the information content of the buy (sell) order flow. Our mechanism does not rely on portfolio or short sale constraints (which are presumably less binding for informed traders such as hedge funds than for mutual funds); instead it follows from a relaxation of the (arguably unrealistic) common knowledge assumption. Empirically, we document that our findings are not driven by short sale constraints.[4]

Second, our paper contributes to the literature on stock return skewness. While this literature spans many aspects, our contribution is to shed light on the determinants of individual stock return skewness, and more specifically, on how it depends on past price movements. Prior research finds that return skewness is negatively related to lagged returns (e.g., Harvey and Siddique, 2000; Chen et al., 2001), but is less clear on the mechanism underlying this relationship. Prominent theories rely on the existence of bubbles (which build up and eventually burst) or on the combination of differences of opinions with short sale constraints (which temporarily prevent bearish information from being fully incorporated into prices; Chen et al., 2001). While these theories explain why skewness is more negative after high past returns, they fail to account for the symmetric phenomenon—which we find to be equally strong in the data—that skewness is more positive after low past returns.[5] Our model offers a parsimonious explanation for these patterns by merely requiring (fully rational) investors to be uncertain about how informed they really are. To be clear, our intention in this paper is to highlight that the dependence of skewness and price impact costs on past returns is multi-faceted, rather than to dismiss alternative mechanisms. More work is needed to ascertain the relative importance of each channel and the conditions under which they prevail.

Finally, our paper is also related to the empirical literature on stale news trading (e.g., Huberman and Regev, 2001; Tetlock, 2011; Gilbert et al., 2012). Most recently, Fedyk and Hodson (2023), in an experimental setting, show that finance professionals struggle with identifying (stale) information that recombines content from multiple sources. While these papers point to attention constraints, correlation neglect, or to 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 and may end up trading on stale information. We explore the ramifications of this idea in a model that is a entirely rational (apart from the usual assumption about noise trading) and shed a first light on its empirical predictions.

The paper proceeds as follows. Section 2 describes a simple trading game and solves it under different assumptions about the information structure, before discussing the distinct predictions that result from uncertainty about what is in the price. Section 3 presents empirical tests of our model predictions regarding return skewness, price impact costs, and stock price informativeness. Section 4 concludes.

[4] Specifically, we find that the dependence of return skewness and price impact costs on past returns does not strengthen when shorting fees (a proxy for the tightness of the constraint) are higher, see Section 3.6 for a detailed discussion.

[5] One exception is Xu (2007), who presents a model in which short sale-constrained investors disagree on the precision of a publicly observed signal. We find no evidence that the return-skewness relationship strengthens when shorting is more costly, contrary to what Xu (2007) implies. Moreover, his model predicts that the relationship should be stronger after public news announcements. We find the opposite in our data.

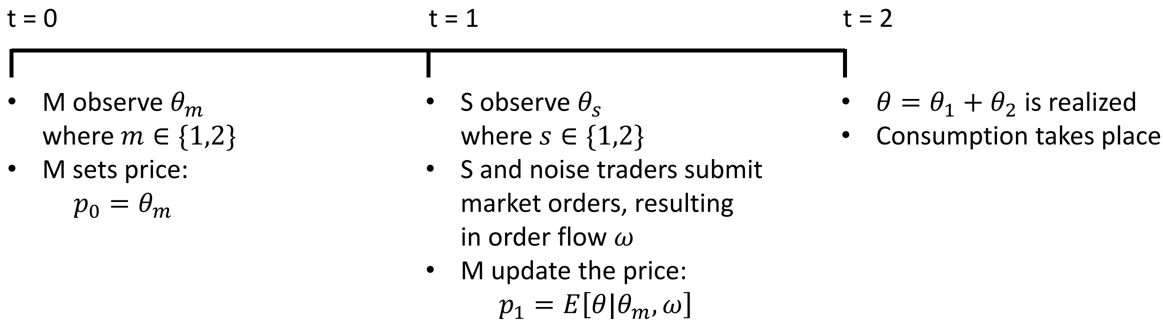

Fig. 1. Model setup. This figure summarizes the model setup. At $t = 0$, market makers M observe θ_m, where $m \in \{1,2\}$ with equal probability, and set $p_0 = \theta_m$. At $t = 1$, speculators S observe θ_s, where $s \in \{1,2\}$ with equal probability, and submit market order to maximize expected 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.

2. Model

We develop a parsimonious model in which investors face uncertainty about what is 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). 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.

We assume that speculators are sufficiently risk averse and/or that fundamental uncertainty is high enough. This assumption ensures that the aggregate order flow at $t = 1$ is not fully revealing (which would render the model uninteresting):

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. Fig. 1 summarizes the model setup.

2.2. Benchmark: No uncertainty about what is in the price

We describe a benchmark in which speculators face no uncertainty about what is 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 is in the price; that is, speculators know which component market makers observe, 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 also 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:

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

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

Proposition 1. *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:*

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

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 learn nothing about θ_s. Therefore, the equilibrium price function is given by Eq. (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):

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

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 $x_{(2)}$. The following proposition summarizes the resulting equilibrium.

Proposition 2. *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 do not 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 is 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 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 an asymmetry emerges: both high and low price signals lead investors to update upward the probability that informed traders are present and thus command a larger price discount, which attenuates (resp. amplifies) the market response to positive (resp. negative) news. Market makers are risk-neutral in our model, so there is no such risk discount effect and the price function remains symmetric despite the uncertainty about whether there are informed traders. Instead, as we will see next, there emerges another type of asymmetry (whose nature changes with p_0) when there is uncertainty about what is in the price.

2.3. Uncertainty about what is 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 S' signal coincides with the signal observed by M as revealed by the $t = 0$ price ($\theta_m = \theta_s$), then 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 $x_{(3)} \geq y_{(3)}$. 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. Their expectation of the component of θ that they do not observe is then $\frac{1}{3}\sigma$ (see Appendix A.3), 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 expectation 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 maintain the price at σ, the realization of their signal θ_m. As a result, the price function is given by Eq. (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 sell volume after a price increase is a more informative signal about the stock's fundamental compared to buy volume. The reason is that sell volume after a price increase indicates that speculators trade on genuine new information, whereas buy volume can also come from speculators trading on stale information (i.e., information already reflected in p_0). Hence, price impact is larger for sell (buy) 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):

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

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

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):

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

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

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 3. *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*

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

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

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

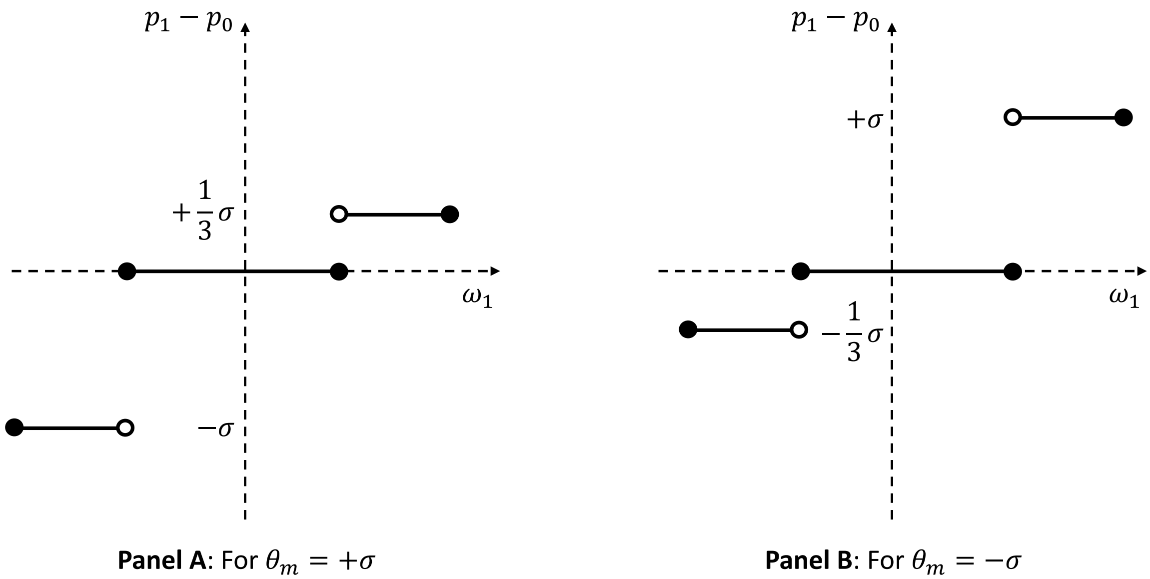

Fig. 2. Equilibrium price function. This figure shows the equilibrium price function when speculators face uncertainty about whether their trading signal is already in the price. In Panel A, we show the price function for the case of prior positive news ($\theta_m = +\sigma$). In Panel B, we show the price function for prior negative news ($\theta_m = -\sigma$).

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 $x_{(3)} \geq y_{(3)}$ 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

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 2.2.1) is denoted by (1), the case with only m being common knowledge (Section 2.2.2) with (2), and the case of interest with neither m nor s being common knowledge (Section 2.3) with (3). We highlight that only equilibrium (3) displays uncertainty about what is in the price.

We begin by comparing the price functions that obtain in the three equilibria.

Corollary 1. *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 is 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 is in the price: price impact costs differ for buys and sells as a function of past price movements, as illustrated in Fig. 2. Intuitively, when buy (sell) volume follows 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.

The asymmetric price impact of Corollary 1 naturally leads to return skewness, with a sign that depends on the prior price change, as described next.[6]

Corollary 2. *In equilibria (1) and (2), return skewness is unrelated to the lagged return. In equilibrium (3)—i.e., under uncertainty about what is in the price—return skewness is negatively related to the lagged return.*

After price increases (decreases), market makers update prices more aggressively in response to incoming sell (buy) orders than to incoming buy (sell) orders, resulting in negatively (positively) skewed returns. Moreover, after price increases (decreases), speculators are more likely to buy (sell) as they may be trading on positive (negative) stale news. This further contributes to negatively (positively) skewed returns after price increases (decreases).

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., 2019)—promotes real efficiency.

[6] Our prediction on the skewness of price changes, $p_1 - p_0$, can be expressed using the skewness of returns, $(p_1 - p_0)/p_0$, since we compare prices across buy and sell orders starting from the same initial price, p_0.

[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: $E\left(Var\left(\theta|p_1, p_0\right)\right) = Var(\theta) - Var\left(E\left(\theta|p_1, p_0\right)\right) = 2\sigma^2 - Var\left(E\left(\theta|p_1, p_0\right)\right)$.

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

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

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 is 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 is 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. Extensions and discussion of model assumptions

Our model is deliberately kept as simple as possible. Nonetheless, we show in Internet Appendices 1.1–1.3 and discuss here that the main intuition is robust to alternative assumptions. Before we describe those extensions, it is worth emphasizing that our model implications derive exclusively from the economy's information structure and, as such, are particularly robust. Specifically, agents, whether speculators or market makers, act on the component of their private signal (i.e., signal about the fundamental and order flow) that they perceive as novel. They identify that component by comparing their private signal to the past price. For instance, following a recent price appreciation, a positive signal is deemed less novel than a negative signal. This belief updating is what gives rise to asymmetric price movements and skewed returns—regardless of the trading environment.

2.5.1. Model with demand schedules

In Internet Appendix 1.1, speculators submit limit orders, i.e., demand schedules that are conditional on the price, rather than market orders as in the baseline. We show that the predictions of our model on return skewness and price informativeness are largely unchanged.[8] Specifically, there are no market makers but two groups of risk averse speculators. Speculators in group M (group S) observe, in period 0 (period 1), a part θ_m (θ_s) of the fundamental where m (s) $\in 1, 2$ with equal probability. The variables m and s are drawn independently so m can be equal to or different from s. Thus, speculators in group M play the role the market makers played in the baseline: they capitalize θ_m into the period-0 price, thereby making it public; but unlike market makers, they are risk averse.

In equilibrium, prices reveal θ_m and θ_s. Indeed, p_0 reveals θ_m to S-speculators (since, as in the baseline, there are no noise traders in that period), while $p_1 - p_0$ reveals θ_s to M-speculators.[9]

But importantly, prices do not always reveal whether θ_m and θ_s correspond to identical or different components of the fundamental, i.e., whether $m = s$ or $m \neq s$: while $m \neq s$ is the only possibility if $\theta_m \neq \theta_s$, both $m \neq s$ and $m = s$ are possible if $\theta_m = \theta_s$. In the case in which $\theta_m = \theta_s$, speculators cannot tell whether their information is truly novel or stale and thus face UWIP. As in the baseline, the equilibrium is characterized by an asymmetric dependence of the price change, $p_1 - p_0$ (and hence of skewness), on the lagged price p_0 (this asymmetry is again distinctive of UWIP since it vanishes when m is common knowledge). Thus, this property does not hinge on the type of order used by traders. The model with demand schedules differs from the baseline because of the existence of a risk discount (due to the absence of risk neutral agents). We discuss this point below.

2.5.2. Price manipulation

A key feature of our model is that speculators know more than market makers about how informed they really are.[10] Does this open the door to price manipulation—e.g., can a speculator pretend to be informed when she really is not? In our baseline model, speculators are competitive and thus an individual speculator cannot hope to manipulate the price with her action. In Internet Appendix 1.3, we relax the competitive-market assumption and study the case of a single (strategic) speculator who accounts for the effect of her order on market makers' pricing behavior. We solve the model for the case when m is common knowledge; i.e., when the speculator has a clear information advantage in that only she knows whether $m = s$. We show that there is no scope for price manipulation. Intuitively, the speculator trades only once. Hence, while she can move prices, she cannot benefit from the price movement because she has no time to flip her position before the fundamental is revealed.[11]

---

[8] This model makes no predictions on (the adverse-selection component of) price impact.

[9] The reason $p_1 - p_0$ is perfectly revealing, despite the presence of noise traders in period 1, is that $p_1 - p_0$ does not depend on noise trading if $\theta_m \neq \theta_s$. Indeed, if $\theta_m \neq \theta_s$, then speculators infer that $m \neq s$, and hence learn the final payoff with certainty; therefore they arbitrage away any deviation caused by noise trading, leading to a price change, $p_1 - p_0 = -\theta_m + \frac{1}{3}\gamma\sigma^2 S$, independent of n. If instead $\theta_m = \theta_s$, then $p_1 - p_0 = \frac{1}{3}\theta_m - \frac{1}{9}\gamma\sigma^2 S + \frac{4}{9}\gamma\sigma^2 n$ which depends on noise. The sharp divergence in price across those two cases makes it possible to perfectly learn θ_m and θ_s. One could prevent the price from being perfectly revealing for instance by introducing residual uncertainty (e.g., assume that θ depends on a third component, θ_3, which is only revealed at the final date $t = 2$), or by assuming that speculators can learn θ_m only imperfectly from p_0. We do not go down this route because we think it is instructive that UWIP remains *despite* the price being fully informative. This relates to our earlier comment that UWIP is an inherent feature of economy's information structure and thus insensitive to assumptions about the trading protocol.

[10] In the case in which m is common knowledge, speculators know whether or not $m = s$. In the case with UWIP, they know whether $\theta_m \neq \theta_s$ (in which case, they infer that their information is novel) or $\theta_m = \theta_s$ (in which case, their information may or may not be novel).

[11] That may no longer be true in a dynamic version of the model (see, e.g., the discussion in Banerjee and Breon-Drish, 2020). We leave this interesting question to future work as manipulation is not the main focus of our paper.

2.5.3. Model à la Glosten and Milgrom (1985) and dynamic extension

In Internet Appendix 1.2, we consider a Glosten and Milgrom (1985) setup in which the bid and ask prices are explicitly modeled. We find that a two-period version of the model closely resembles the baseline in that price changes depend asymmetrically on lagged prices. Specifically, following positive news, market makers raise the asset's ask price by less than they reduce its bid price. Conversely, following negative news, they decrease the bid price more than they increase the ask price. Again, the asymmetry arises exclusively as a result of UWIP (there is no such asymmetry when only m is common knowledge). This version of the model demonstrates that our findings do not hinge on the specifics of the trading protocol (e.g., whether the market is quote- or order-driven).

In addition, we extend this framework to multiple periods in order to study how UWIP evolves over time, i.e., whether agents learn about the staleness of signals. We find that this depends on how signals are distributed among speculators. If they are uncorrelated, then the effect of UWIP dissipates over time, whereas it persists if they are perfectly correlated. To see why, consider what happens as the number of orders grows to infinity. With i.i.d. signals, market makers can perfectly infer the fundamental from the proportion of buy and sell orders, and there is no more UWIP. In contrast, when speculators all observe the same signal, market makers recover that signal in the limit, but they remain unable to ascertain whether it corresponds to the same component of the fundamental as their own signal when $\theta_m = \theta_s$, and therefore UWIP remains unresolved. These findings show that the *persistence* of UWIP depends on the distribution of information across speculators. Importantly, they also underscore that UWIP always (i.e., regardless of that distribution) matters in the beginning: it is high just after speculators see market makers update prices without knowing which piece of news caused the update (and then it may or not diminish).[12]

2.5.4. Implications of a risk discount

An important feature of the baseline model is the presence of risk neutral agents (specifically, market makers). This leads to return distributions that are mirror images following good and bad news. As a result, (i) volatility does not depend on the news, despite UWIP, and (ii) (UWIP-induced) skewness is symmetric. To be precise, returns are left-skewed following good news and right-skewed following bad news (due to UWIP), but they are so to the same extent. Removing risk neutral agents breaks the symmetry, as the model with demand schedules in Internet Appendix 1.1 demonstrates (see also Banerjee and Green, 2015). In this model, prices are determined not only by the asset's expected cashflow (the only force at work in the baseline) but also by a risk discount. While the former operates symmetrically across good and bad news, the latter does not. This affects both the volatility and skewness of returns.

To see how volatility is impacted, observe that if $\theta_m \neq \theta_s$, then uncertainty is entirely resolved in period 1, thereby eliminating the risk discount and pushing the price up. If instead $\theta_m = \theta_s$, then the uncertainty persists so the risk discount is amplified, thereby pushing the price down. The resulting price changes dampen the cashflow effect if $\theta_m = \sigma$, but reinforce it if $\theta_m = -\sigma$. As a result, volatility depends on the lagged price p_0: returns are less volatile after good news, and more volatile after bad news, compared to the baseline. This pattern is not distinctive of UWIP; indeed, Banerjee and Green (2015) show that it emerges in a model in which uninformed investors are uncertain about the presence of informed investors (but informed investors do not face UWIP).[13] In other words, while our model predicts a negative dependence of volatility on past returns, a.k.a. a "leverage effect", this is not a novel prediction.

Novel to our model is that the existence of a risk discount in the price affects return skewness asymmetrically. Compared to the baseline, returns are less left-skewed after good news (i.e., they go up by less after further positive news than they go down after negative news) and more right-skewed after bad news. Put differently, the risk discount weakens the UWIP-induced left-skewness after good news and strengthens the right-skewness after bad news.[14] Finally, the model with priced risk also predicts that the risk premium is higher under UWIP than under common knowledge. This property echoes results from theories in which adverse selection or information risk are priced (e.g., Easley and O'hara, 2005).

To summarize, with a risk discount, UWIP makes returns (i) less (more) volatile after positive (negative) news, (ii) less left-skewed after positive news than they are right-skewed after negative news, and (iii) higher on average.

2.5.5. Other features of the model

We discuss here four final features of the model. 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 (as, e.g., in the models solved in Internet Appendix 1.1 and 1.2). Such a model yields similar insights. The price at $t = 0$ 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.

Second, 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 particularly 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.[15] This noise might stem from market makers observing a noisy signal of θ_s or from trading for reasons unrelated to the stock's fundamentals such as inventory concerns.[16] As before, 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

[12] This version of the model clarifies how to think of UWIP in a model in which speculators arrive sequentially and observe *different* signals (as opposed to having perfectly correlated signals as in our baseline model). When there are multiple sources of uncertainty that speculators could be informed about, a speculator that comes second (or third...) always faces uncertainty about whether her signal is stale or novel. Thus, UWIP arises naturally following price moves caused by private information.

[13] In our model with demand schedules and risk-averse agents, return volatility is independent of p_0 when m is common knowledge. However, this is due to the fact that, in our model with binary fundamentals, prices reveal all the information.

[14] We find support for this auxiliary prediction, as we discuss below.

[15] 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 is in the price disappears).

[16] Alternatively, and as noted above, noise in the $t = 0$ price could come from another group of speculators trading with noise traders at $t = 0$.

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.

Third, trading on stale news occurs in our model despite all investors (speculators and market makers) being *rational*. Indeed, we think of this as the key contribution of our model: in a world with multidimensional uncertainty, even rational investors are unsure what news is priced in and, hence, may end up trading on stale news. In practice, some stale news trading may be due to (irrational) noise traders or feedback traders/trend chasers (e.g., Barber and Odean, 2007; Tetlock, 2011). Since price impact in our theory is caused by adverse selection, our model has nothing to say on the effect of stale news trading on other sources of illiquidity, such as inventory or noise trader risk (e.g., Grossman and Miller, 1988; Foucault et al., 2011; Hendershott and Menkveld, 2014; Peress and Schmidt, 2020).[17] Still, its intuitions about price impact due to adverse selection are robust to the presence of naïve feedback traders. Indeed, investors who indiscriminately buy (sell) the stock at $t = 1$ after observing a price increase (decrease) at $t = 0$ do not affect our equilibrium price functions to the extent that they do not change the informativeness of order flow.[18]

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 is 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 is in the price?

3.1. Testable hypotheses

In this section, we present a first evaluation of the relevance of uncertainty about what is in the price (UWIP). Specifically, we derive from Corollaries 1–2 in Section 2.4 distinct predictions that help assess whether UWIP is an actual concern for stock market participants.

Our first two hypotheses follow from the asymmetry in equilibrium price functions that arises when investors face UWIP (see Corollaries 1 and 2).

Hypothesis 1. Stock return skewness depends negatively on past returns.

Hypothesis 2. Price impact costs depend asymmetrically on past returns: they decrease in past returns for buys and increase in past returns for sells.

We note that Hypothesis 2 predicts an asymmetry in price impact that is different from what models of inventory risk predict (Ho and Stoll, 1981; Madhavan and Smidt, 1993; Hendershott and Menkveld, 2014). According to those models, after a flow of sell orders, the market maker has an undesirably large inventory and thus posts a low price to discourage further sells.[19] Our model makes the opposite prediction for the *information*-related component of price impact: sells following earlier sells may come from stale news trading and thus require less price updating.

We further 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. For instance, immediately after earnings announcements, investors understand that recent price movements are driven by the public earnings news (i.e., in the language of our model, m is common knowledge), making it easier for them to assess whether their own information is already priced in. We therefore posit that UWIP, and the associated effects of past returns on return skewness and price impact costs, are weaker after earnings announcements. In a similar vein, to the extent that large stocks and stocks with high analyst coverage have more transparent prices, we expect these stocks to exhibit a weaker dependence of return skewness and price impact costs on past returns. In contrast, for stocks that are complicated and hard to analyze (e.g., stocks that operate in multiple segments; see Cohen and Lou, 2012), we expect a stronger dependence of return skewness and price impact costs on past returns.

Hypothesis 1′. The dependence of return skewness on past returns weakens after earnings announcements, for large stocks and for stocks with high analyst coverage, while it strengthens for complicated stocks (i.e., stocks that operate in multiple segments).

Hypothesis 2′. The asymmetric dependence of the price impact on past returns weakens after earnings announcements, for large stocks, and for stocks with high analyst coverage, while it strengthens for complicated stocks (i.e., stocks that operate in multiple segments).

Our final hypothesis is about the source of the order flow. Recall that, in our model, the effect of UWIP is inherently tied to the presence of informed traders. If there are no such traders, then price impact is no longer asymmetric.[20] A large body of empirical research finds institutional orders to be more informed than retail orders (e.g., Easley et al., 1996; Hendershott et al., 2015). We use this insight, together with a recent methodology for identifying retail order flow (Boehmer et al., 2021), to test the following hypothesis:

Hypothesis 2″. The asymmetric dependence of the price impact on past returns is weaker for retail order flow compared to institutional order flow.

[17] In our empirical analysis below, we therefore focus on liquidity measures related to adverse selection risk.

[18] Specifically, if the amount of feedback trading could be perfectly anticipated by market makers, Proposition 3 would remain unchanged except that the order flow would then be centered on this amount of feedback trading (instead of zero). If the amount of feedback trading were uncertain, it would essentially add noise to the order flow without altering the model's key intuitions; namely that speculators learn about the novelty of their signals from past price movements and market makers take this into account by charging a higher price impact for order flow that goes against recent price movements.

[19] For instance, in Hendershott and Menkveld (2014), price pressure decreases in the market maker's inventory; see their Equation (21) on p. 416.

[20] Viewed through the lens of our model, when there are no speculators in period 1, the order flow at that date only comes from noise traders and is therefore completely uninformative. Competitive risk-neutral market makers then charge no price impact, regardless of the direction of the order flow.

Corollary 3 yields the further prediction that when UWIP is higher, prices are less informative as risk-averse investors trade more cautiously and impound less information into stock prices. Additional predictions emerge when prices are set by market clearing in a model with risk-averse investors (instead of by risk-neutral market makers) and hence include a risk discount (see Section 2.5.4 and Internet Appendix 1.1). In that setting, UWIP raises the risk premium and leads return volatility to depend negatively on past returns. Empirical studies linking opacity in the information environment or adverse selection risk to volatility and expected returns, and those documenting the leverage effect, can thus be considered broadly consistent with our model (see, e.g., Easley et al., 2002; Fang and Peress, 2009). In addition, the negative relationship between return skewness and past returns (found in our baseline model) is now stronger for negative past returns than for positive past returns; in other words, returns are less left-skewed after positive news than they are right-skewed after negative news.

We acknowledge, however, that—with respect to volatility, expected returns, and stock price informativeness—identifying the *incremental* effect of UWIP is challenging as it requires finding a convincing empirical proxy that isolates variations in UWIP. A natural idea is to use the strength of the skewness-return dependence or of the price impact asymmetry to measure UWIP. Yet, as our previous discussion clarifies, such a measure simultaneously captures information uncertainty *and* informed trading (see footnote 20). To the extent that informed trading itself affects volatility, expected returns, and stock price informativeness (see, e.g., Easley and O'hara, 2005; Weller, 2018), any correlation between, e.g., price impact asymmetry and one of these outcomes may not solely be driven by variation in UWIP. We therefore leave a careful investigation of these additional predictions to future research, while focusing below on testing the predictions that are more directly and distinctly related to our mechanism—i.e., UWIP causing skewness-return dependence and an asymmetric price impact between buys and sells.[21]

3.2. 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 a stock's daily trade imbalance, we sum all signed trades in numbers of shares traded over the course of the day. We use this measure to split our sample into days with net-buy or net-sell activity. To estimate price impact costs, we also compute the daily trade imbalance in dollars by summing signed trades multiplied with the prevailing mid-quote at the end of the 5-min interval in which the trade occurred.

3.2.1. Skewness and price impact measures

We measure a stock's daily return skewness as the realized daily skewness based on intraday returns standardized by the realized variance:

$$\text{skewness}_{it} = \frac{\sqrt{K}\sum_{k=1}^{K}\left(\text{return}_{itk} - \overline{\text{return}}_{it}\right)^3}{\left[\sum_{k=1}^{K}\left(\text{return}_{itk} - \overline{\text{return}}_{it}\right)^2\right]^{3/2}},$$

where return_{itk} is the return (calculated from bid–ask midpoints) over 5-min interval k for stock i and day t, $\overline{\text{return}}_{it}$ is the return of stock i averaged over all 5-min intervals comprising day t, and K denotes the number of such intervals on day t. Negative (positive) values indicate that the stock's return distribution has a left tail that is fatter (thinner) than the right tail.

We employ four different price impact measures that are designed to capture adverse selection risk (as faced by the market makers in our model).[22] We do not study generic bid–ask spreads or realized spreads as they are influenced by, or designed to capture, liquidity costs arising from inventory risk (see Foucault et al., 2013, Chapter 2), which is outside of our model. The first three price impact measures 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 day 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.[23] The difference with the Amihud ratio is that we use the signed trade imbalance (in dollars) 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.

[21] Note that these tests do not suffer from the "joint hypothesis" problem identified above. Indeed, as the comparison between the different model benchmarks in the paper demonstrates, the presence of informed speculators alone does not give rise to an asymmetric price response. Rather, the asymmetric price response only emerges when we break down the common knowledge assumption (i.e., when neither m nor s are common knowledge). Hence, evidence for an asymmetric price response (that we document in our paper) proves that this type of uncertainty is important.

[22] These price impact measures capture (the adverse selection-related part of) the trading costs for the "average" trader in the stock market. Of course, the effective trading costs faced by different types of trades can vary. For instance, analyzing proprietary data of executed trades by a large money manager, Frazzini et al. (2018) find substantially lower trading costs for what they describe as a *patient* trader. Our model, however, is about how competitive market makers respond to order flow in a world in which investors face uncertainty about what is in the price. As such, price impact in our model should be viewed as pertaining to the "average" trader. Moreover, to the extent that the trading costs of different traders are correlated, uncertainty about what is in the price should affect the trading costs of different traders in a qualitatively similar way.

[23] We find even stronger results if we do not adjust returns for autocorrelation (available upon request). The adjustment for autocorrelation is done as follows: for a daily return of stock i in month $\tau(t)$, we run a regression of stock i's return on lagged returns over the past one to five days, denoted by subscript j, using the previous twelve month ($\tau-12$ to $\tau-1$) and record the autocorrelation coefficients $\hat{\beta}_{i\tau(t)j}$. The return adjusted for autocorrelation is defined as $\text{return}_{it} - \sum_{j=1}^{5}\hat{\beta}_{i\tau(t)j}\times\text{return}_{it-j}$.

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 (see Hasbrouck, 2009). The third measure is *quote-based price impact*, defined as the dollar-weighted daily average of 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 absolute return divided by its dollar volume.[24] Goyenko et al. (2009) show that it does a good job of capturing adverse selection. We winsorize skewness and price impact measures, as well as other continuous variables used in this study, at the 1% level on both sides.

3.2.2. Methodology

Our model's key predictions are that a stock's daily return skewness depends negatively on past returns (H1) and that price impact costs for buys and sells depend asymmetrically on past returns: they decrease in past returns for buys and increase in past returns for sells (H2). The model is agnostic about the horizon over which past returns should be measured; they may be measured intraday, over one day, or over multiple days. Accordingly, we consider in our empirical tests time windows spanning one, five, and ten trading days (to which we refer as the "lookback window").

Specifically, for skewness, we run the following regression:

$$\text{skewness}_{it} = \alpha_{i\tau} + \alpha_t + \beta \,\text{past return}_{it} + \gamma \, X_{it-1} + \epsilon_{it},$$

where skewness_{it} is the return skewness 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 over the lookback window (that is, on the prior trading day ($t-1$), cumulated over the previous five trading days ($t-5$ to $t-1$), or cumulated over the previous ten trading days ($t-10$ to $t-1$)), and X_{it-1} is a vector of controls, which includes past turnover and past squared return (as a proxy for volatility) measured over the same lookback window. Our theory predicts $\beta < 0$.

For price impact, we run the following pair of regressions separately for days with positive and negative net-buying activity:

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

where price impact_{it} is one of our four price impact proxies. Our theory predicts $\beta^b < 0$ and $\beta^s > 0$. When the past return is simply the stock's prior-day return, our price impact regressions are 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, one may mechanically find $\beta^b < 0$ and $\beta^s > 0$. This is why we used autocorrelation-adjusted returns in the construction of this price impact measure.

Note that, because our models are saturated with stock-month and date fixed effects, we are controlling for stock-specific variations in skewness and illiquidity costs and for any persistent firm characteristics (e.g., analyst coverage and market capitalization). For price impact, for instance, our identification comes from the incremental effect of past returns on price impact costs, separately for buys and sells, while controlling for the average level of price impact costs for the stock in the same month and for the average level of price impact costs across stocks on the same day. In the paper, we carry out our tests with raw returns. In Internet Appendix 2.1, we show that all our results are robust to using Fama–French 3-factor alphas.

3.2.3. Descriptive statistics

Table 1 Panel (a) reports summary statistics for our dependent and independent variables in the overall sample. For better visibility, *price impact costs*, *lambda*, and *quote-based price impact* are scaled by 10^6, 10^4, and 10^2, respectively. For instance, the median of the *price impact costs* variable implies that a one million USD net buy would be expected to push up the price by 0.87%. For the *quote-based price impact*, the median is slightly lower at 0.14%. The mean and median of our *return skewness* measure have opposite signs, but are both close to zero. The table also shows the standard deviation for each of our dependent and independent variables, which we use below to assess the economic significance of our findings.

An old literature finds that the price impact for institutional block purchases is larger than the price impact for block sales (Kraus and Stoll, 1972; Keim and Madhavan, 1996; this empirical fact motivates the Saar, 2001, model). Table 1 Panel (b) reports that, for all 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 activity. Hence, if anything, the price impact for sells is larger than the price impact for buys in our (more recent) sample.

3.3. Baseline results

Table 2 shows the results of the skewness tests for the different lookback windows. Consistent with Hypothesis H1, we find a strong negative relation between return skewness and past returns regardless of whether we focus on lagged 1-day returns (Column 1), lagged 5-days returns (Column 4), or lagged 10-days returns (Column 7). The economic magnitude of the effect is meaningful. For example, a 1-SD increase in lagged 10-days returns decreases daily return skewness by about 9% ($-2.32 \times 0.11/2.90$) of its SD.

Earlier work documents that return skewness is negatively related to lagged returns at *low* frequency (e.g., Harvey and Siddique, 2000; Chen et al., 2001) and attributes this phenomenon to the gradual build-up and eventual burst of stock price bubbles (Chen et al., 2001). Our results show that the skewness-return relationship also exists at *high* (i.e. daily) frequency—where bubbles are a less plausible explanation—and that it is a distinct phenomenon. Indeed, when we split the sample according to whether the past return is negative or positive (confer Columns 2–3, 5–6, and 8–9), we find that the negative skewness-return relationship is not confined to positive returns as the bubble explanation posits. Instead, we find that this relationship is even more pronounced after negative returns. While this fact is hard to reconcile with an explanation based on bubbles, it naturally follows from our model.[25]

---

[24] 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.

[25] Specifically, our baseline model predicts that the dependence of skewness on past returns is equally strong for positive and negative returns. Upon extending the model to include a risk discount in prices (e.g., when market makers are risk averse), we can also account for the heightened skewness-return relationship following negative returns (see the discussion in Section 2.5.4).

Table 1
Descriptive statistics.

Panel (a): Descriptive statistics for overall sample

	Mean	Median	Standard deviation
Dependent variables			
Return skewness	-0.0152	0.0309	2.8967
Price impact costs	0.6063	0.0087	5.8523
Lamdba	0.1413	0.0224	0.4777
Quote-based price impact	0.3966	0.1443	0.7834
Ln(Amihud)	-16.9706	-17.8151	1.8021
Independent variables			
Past 1-day return	0.0008	0.0000	0.0348
Past 5-days return	0.0045	0.0016	0.0752
Past 10-days return	0.0090	0.0051	0.1055
Past 1-day turnover	0.0068	0.0036	0.0095
Past 5-days turnover	0.0070	0.0040	0.0087
Past 10-days turnover	0.0070	0.0042	0.0084
Past 1-day volatility	0.0013	0.0002	0.0035
Past 5-days volatility	0.0015	0.0005	0.0029
Past 10-days volatility	0.0015	0.0006	0.0027
N	22,433,401		

Panel (b): Descriptive statistics for *net buy*-days

	Mean	Median	Standard deviation
Return skewness	0.4756	0.2233	2.6732
Price impact costs	0.4923	0.0045	5.0265
Lamdba	0.1254	0.0194	0.4469
Quote-based price impact	0.3739	0.1373	0.7434
Ln(Amihud)	-17.1294	-17.9773	1.7198
N	10,908,396		

Panel (c): Descriptive statistics for *net sell*-days

	Mean	Median	Standard deviation
Return skewness	-0.4843	-0.1869	3.0054
Price impact costs	0.6671	0.0157	5.9867
Lamdba	0.1542	0.0260	0.4980
Quote-based price impact	0.4158	0.1514	0.8154
Ln(Amihud)	-16.8346	-17.6295	1.8532
N	11,429,860		

This table reports descriptive statistics. Panel (a) shows statistics for our dependent and independent variables in the overall (stock-day) sample. Panels (b) and (c) show descriptive statistics separately for days with positive net trade imbalance (*net-buy* subsample) and negative net trade imbalance (*net-sell* subsample). ***Return skewness*** is defined as realized skewness standardized by realized variance of 5-min relative price changes on a given day. ***Price impact costs*** is defined as the ratio of the (autocorrelation-adjusted) return over the net trade imbalance (in dollar). 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. The measure is multiplied by 10^4 for better visibility. ***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. The measure is multiplied by 10^2 for better visibility. ***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. Past X-day return is the cumulated raw return over the previous X trading days. Past X-day turnover is the average share turnover over the previous X trading days. Past X-day volatility is the average of squared raw return over the previous X trading days. All dependent and independent variables are winsorized at the 1% level on both sides.

Table 3 shows the results of the price impact tests. Each panel focuses on one price impact measure. Regardless of whether we look at ***price impact costs*** (Panel (a)), ***lambda*** (Panel (b)), ***quote-based price impact*** (Panel (c)), or ***ln(Amihud)*** (Panel (d)) and regardless of the lookback window,

Table 2
Return skewness and past returns.

	Lagged 1-day			Lagged 5-days			Lagged 10-days		
	All days	≤ 0	> 0	All days	≤ 0	> 0	All days	≤ 0	> 0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past return	-0.8531***	-1.4637***	0.6213***	-2.0598***	-4.1087***	-3.1597***	-2.3196***	-3.9835***	-3.4973***
	(-24.61)	(-15.19)	(6.34)	(-74.27)	(-71.29)	(-71.72)	(-74.27)	(-77.09)	(-79.71)
Past turnover	-0.2307	1.8332***	-3.8769***	-2.2566***	-3.0702***	-3.2762***	-2.9110***	-4.5620***	-2.0980***
	(-1.63)	(9.92)	(-21.71)	(-10.83)	(-10.14)	(-11.16)	(-10.36)	(-10.39)	(-5.47)
Past volatility	12.5636***	18.2397***	-2.5886***	20.4112***	6.3265***	30.1649***	31.3325***	4.5639***	52.5382***
	(34.66)	(24.08)	(-3.96)	(32.94)	(6.72)	(36.53)	(36.02)	(3.34)	(42.48)
t-stat of difference in past return coef.		(15.52)***			(17.53)***			(10.46)***	
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	23,925,441	13,011,833	10,862,693	23,916,920	11,255,620	12,586,601	23,900,236	10,962,216	12,812,359
adj. R^2	0.06	0.06	0.08	0.06	0.08	0.08	0.07	0.08	0.08

This table reports the results from regressing daily *return skewness* on lagged raw returns as explained in Section 3.2. Columns (1)–(3) show the results for using lagged 1-day returns as the key independent variable. Columns (4)–(6) show the results for using lagged cumulated 5-day returns as the key independent variable. Columns (7)–(9) show the results for using lagged cumulated 10-day returns as the key independent variable. The past turnover and past volatility controls are chosen accordingly. Columns (1), (4), and (7) show results for the overall sample; Columns (2), (3), (5), (6), (8), and (9) break up the sample into days with a negative or positive past return, respectively (at the bottom of the table, we show *t*-statistics for whether the past return coefficients for these subsamples are statistically significantly different). 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.

we consistently obtain results in line with the model's prediction: on days with a positive trade imbalance (net-buys), price impact is significantly negatively related to past returns; whereas on days with a negative trade imbalance (net-sells), price impact is significantly positively related to past returns.[26] The only exception occurs for the *quote-based price impact* at the 1-day lookback window, for which the regression coefficient on the lagged return for net sells is negative (Panel (c), Column 2), but the coefficient estimate and its statistical significance are an order of magnitude smaller than for net buys (Panel (c), Column 1).

Results are highly statistically significant and appear to grow stronger with the length of the lookback window. For instance, a 1-SD increase in the lagged 1-day return decreases (increases) the *price impact costs* on days with a positive (negative) net trade imbalance by about 5% of its SD, thus driving a wedge between the *price impact costs* on buy- and sell-days of about 10% of its SD. The wedge equals about 15% of a SD for the 10-days lookback window. A similar pattern is observable for the *lambda* and the *quote-based price impact* measures, although the magnitudes are weaker. These results indicate that uncertainty about what is in the price is not only a short-term concern for market participants but one that extends over many days.

We emphasize that we control in our regressions for stock-specific trends in liquidity by including stock-month fixed effects. Moreover, our comprehensive panel dataset—covering all NYSE stocks for a 12-year period—yields strong statistical power as indicated by the large *t*-statistics (despite of double-clustering standard errors by stock and date). In conclusion, the results in Table 3 strongly support Hypothesis H2.

3.4. The role of the information environment

In this subsection, we conduct powerful auxiliary tests of our theory. If, as we argue, the dependence of skewness and price impact on past returns is caused by uncertainty about what is in the price, then it should be less pronounced (i) at times when this uncertainty is lower, and (ii) for stocks with lower information asymmetry (H1' and H2'). For brevity, we display here results for the 10-days lookback window.[27]

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 skewness regression as follows:

$$\text{skewness}_{it} = \alpha_{i\tau} + \alpha_t + \beta_1 \text{ past return}_{it} + \beta_2 \text{ EA}_{it} + \beta_3 \text{ past return}_{it} \times \text{EA}_{it} + \gamma X_{it-1} + \epsilon_{it},$$

where EA_{it} is a dummy variable that takes the value of one when stock i on date t had an earnings announcement over the past 10 trading days.[28] Likewise, we modify our price impact regressions as follows:

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

The variable past $\text{return}_{it} \times \text{EA}_{it}$ denotes the interaction of the earnings announcement dummy with the past return. All other variables and fixed effects are as in our baseline regression.[29] Based on our theory, we expect $\beta_1 < 0$, $\beta_1^b < 0$, and $\beta_1^s > 0$, but $\beta_3 > 0$, $\beta_3^b > 0$, and $\beta_3^s < 0$. In words,

[26] In Internet Appendix 3.1, we show that we obtain consistent results when we run a similar test using an intraday version of our *price impact costs* measure and 5-min intraday returns.

[27] In Internet Appendix 2.2, we report similar results for the 1-day and the 5-days lookback windows.

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

[29] One concern is that our results may be confounded by the earnings announcement itself, which is known to coincide with heightened price impact. In Internet Appendix 2.5, we show that our results are robust to excluding the earnings announcement day and the following day from our analysis.

Table 3
Price impact and past returns.

Panel (a): Dependent variable: Price impact costs

	Lagged 1-day		Lagged 5-days		Lagged 10-days	
	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)
Past return	-8.4501***	7.9972***	-4.5860***	3.9176***	-4.5070***	3.9062***
	(-62.51)	(58.74)	(-63.34)	(57.95)	(-64.92)	(60.71)
Past turnover	-3.6916***	-7.2845***	-2.1964***	1.5703***	3.8177***	6.5108***
	(-12.11)	(-19.02)	(-4.48)	(2.66)	(5.60)	(8.47)
Past volatility	46.8435***	-51.2077***	35.4212***	-96.9425***	15.8986***	-130.2595***
	(34.32)	(-35.83)	(16.72)	(-40.37)	(5.65)	(-42.02)
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
N	11,256,384	11,999,824	11,256,394	11,999,835	11,254,789	11,997,894
adj. R^2	0.17	0.12	0.17	0.12	0.17	0.12

Panel (b): Dependent variable: Lambda

	Lagged 1-day		Lagged 5-days		Lagged 10-days	
	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)
Past return	-0.1789***	0.0315***	-0.1658***	0.0468***	-0.1563***	0.0517***
	(-27.30)	(5.31)	(-45.13)	(15.15)	(-50.97)	(19.04)
Past turnover	-0.6391***	-0.7546***	-0.6653***	-0.9743***	-0.2762***	-0.7182***
	(-33.05)	(-36.16)	(-20.69)	(-27.08)	(-6.49)	(-15.29)
Past volatility	2.0497***	1.7077***	1.6103***	1.2327***	-0.0791	-1.0507***
	(26.13)	(23.44)	(12.42)	(9.48)	(-0.45)	(-6.10)
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
N	11,643,212	12,481,578	11,639,822	12,476,825	11,632,858	12,466,883
adj. R^2	0.35	0.30	0.35	0.30	0.35	0.30

Panel (c): Dependent variable: Quote-based price impact

	Lagged 1-day		Lagged 5-days		Lagged 10-days	
	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)
Past return	-0.2353***	-0.0318***	-0.2723***	0.0778***	-0.2752***	0.1085***
	(-23.32)	(-3.33)	(-49.77)	(15.28)	(-59.69)	(23.93)
Past turnover	-1.3414***	-1.2069***	-1.9172***	-1.9636***	-1.7660***	-1.9834***
	(-36.70)	(-32.36)	(-32.59)	(-32.63)	(-22.29)	(-24.08)
Past volatility	4.8702***	5.0263***	3.1059***	3.5527***	-0.5489*	-1.1362***
	(37.80)	(39.82)	(14.31)	(16.95)	(-1.80)	(-3.81)
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
N	11,632,436	12,446,048	11,629,070	12,441,313	11,622,111	12,431,403
adj. R^2	0.36	0.30	0.36	0.30	0.36	0.30

Panel (d): Ln(Amihud)

	Lagged 1-day		Lagged 5-days		Lagged 10-days	
	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)
Past return	-2.2732***	1.8995***	-1.2229***	0.5658***	-1.0076***	0.3665***
	(-86.84)	(64.44)	(-94.90)	(45.06)	(-96.74)	(34.90)
Past turnover	-7.3087***	-10.0111***	-6.4596***	-8.4929***	-3.6193***	-5.2103***
	(-72.03)	(-86.49)	(-46.66)	(-51.85)	(-20.10)	(-24.02)
Past volatility	1.4582***	-4.4702***	-9.8627***	-17.7004***	-21.1014***	-31.8585***
	(6.93)	(-18.28)	(-27.07)	(-41.81)	(-39.52)	(-50.58)
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes
N	11,678,977	12,524,311	11,675,574	12,519,533	11,668,565	12,509,524
adj. R^2	0.69	0.63	0.70	0.63	0.70	0.63

This table reports the results from regressing our four price impact measures on lagged raw returns as explained in Section 3.2. Panel (a) shows the results for the *price impact costs* measure (scaled by 10^6 for better visibility). Panel (b) shows results for the *lambda* measure (scaled by 10^4 for better visibility). Panel (c) shows results for the *quote-based price impact* measure (scaled by 10^2 for better visibility). Panel (d) shows results for the *ln(Amihud)* measure. In each panel, Columns (1)–(2) show results for using lagged 1-day returns as the key independent variable, Columns (3)–(4) show results for using lagged cumulated 5-day returns as the key independent variable, and Columns (5)–(6) show results for using lagged cumulated 10-day returns as the key independent variable. The past turnover and past volatility controls are chosen accordingly. 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.

return skewness should be negatively related to past returns, but this relation should be weakened just after earnings announcements. Likewise, on buy-days (sell-days), price impact should be negatively (positively) related to past returns, but less so just after earnings announcements.

Table 4 Panel (a) presents the results.[30] For return skewness (Column 1), the coefficient estimate on past returns interacted with the earnings announcement dummy (β_3) is significantly positive, indicating that the negative effect of past returns on skewness weakens after earnings announcements. In terms of magnitude, earnings announcements reduce the skewness-return relation by more than one quarter ($0.55/2.03 = 27\%$). Similarly, 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* (Columns 2–3), the results indicate that the effect of past returns on price impact is about a third lower ($1.02/2.97 = 34\%$) when an earnings announcement occurred over the previous 10 trading days.[31]

Our second set of tests exploits variations across stocks in the degree of uncertainty about what is in the price. We consider three stock characteristics that proxy for UWIP. First, we argue that stocks with more analyst coverage have a more transparent public information environment, implying that the scope for information asymmetry and thus UWIP is reduced. The tests are similar to the preceding ones, except that we now interact past returns with analyst coverage, instead of an earnings announcement dummy.[32] The findings, reported in Table 4 Panel (b), lend support to our mechanism: the dependence of skewness on past returns and the asymmetry in price impact on buy- and sell-days is reduced for stocks with more analyst coverage.

Second, we interact past returns with the stock's market capitalization (at the end of the previous month). Here, the prediction is less clear cut. On the one hand, larger stocks receive more media (and analyst) coverage and thus presumably have superior public information (Fang and Peress, 2009). On the other hand, larger stocks tend to operate in multiple segments and may thus be more "complicated" (Cohen and Lou, 2012). As reported in 4 Panel (c), we find that the first effect dominates: the dependence of skewness on past returns and the asymmetry in price impact is significantly muted for larger stocks. To see whether the second effect is also at play, we follow Cohen and Lou (2012) and retrieve from Compustat Segment files the number of industry segments (based on sales generated in 2-digit SIC industries) a firm operates in during a given year. Since there is a strong positive correlation between the number of segments and firm size (which we know works in the opposite direction), we orthogonalize the (natural logarithm of the) number of segments by regressing it on the (natural logarithm of) market capitalization. We then use the residuals from this regression, denoted $\#Segm^{orth}$, in our interaction tests. Intuitively, $\#Segm^{orth}$ will be large for firms that operate in many segments relative to their size and we expect such "complicated" firms to display higher UWIP.[33] As reported in 4 Panel (d), this is precisely what we find: past returns have a more negative effect on return skewness and lead to a bigger asymmetry in price impact for buys and sells.[34]

3.5. Institutional vs. retail trade imbalance

In our model, the asymmetry in price impact arises from market makers' concerns about adverse selection: after negative (positive) returns, they become more (less) worried that buy orders might stem from informed speculators with access to novel information. In practice, we expect the strength of these concerns to depend on the *source* of the order flow. To the extent that retail investors are less well-informed than institutional investors, the price impact asymmetry should be weaker for retail orders than for institutional orders. In this subsection, we test this prediction by analyzing data on retail investors' order flow estimated using the Boehmer et al. (2021) methodology.

Specifically, Boehmer et al. (2021) exploit a regulatory restriction by which only retail order flow, but not institutional order flow, can receive subpenny price improvement. As such, they identify (marketable) retail buys (sells) in TAQ as transactions with a price slightly below (above) the round penny. We have access to thus-identified retail buy and sell volume for the 2005–2014 subperiod. For each stock and day, we define retail trade imbalance as the difference between retail buy and sell volumes, and institutional trade imbalance as the difference between total trade imbalance and the retail trade imbalance. We then repeat our price impact regressions except that we now split the sample based on whether the institutional (retail) trade imbalance (instead of the total trade imbalance) is positive or negative.[35] Table 5 shows the results. Consistent with the idea that institutional order flow is perceived as more informative and thus more affected by UWIP, Panel (a) finds a strong price impact asymmetry for institutional buys and sells. Indeed, price impact is consistently and significantly negatively associated with past returns on institutional net-buy days, and significantly positively associated with past returns on institutional net-sell days. As shown in Panel (b), this asymmetry is less discernible for retail trades. Now, for the net-sell subsample, the coefficient on past returns is negative for three out of the four price impact proxies. For those, the difference in coefficients between the buy- and the sell-columns is an order of magnitude smaller for retail trades compared to institutional trades. Hence, market makers appear to be less worried about the informational content of retail order flow that goes against recent returns, presumably because retail investors are perceived to be less informed to begin with.

[30] Since one might be concerned that our results are confounded by the announcement itself, we show in Internet Appendix 2.5 that we obtain similar results when we exclude the earnings announcement day and the following day from the analysis.

[31] This finding speaks against an alternative explanation whereby the negative relation between return skewness and lagged returns stems from a combination of short sale constraints and disagreement about the precision of public news (Xu, 2007). More specifically, in the Xu (2007) model, short sale-constrained investors disagree on the precision of a publicly observed signal, leading to an overreaction (underreaction) to positive (negative) realizations of that signal. As a consequence, return skewness is positively correlated with contemporaneous returns, but negatively correlated with lagged returns. Under this explanation, one would have expected the return-skewness relation to be more pronounced in the immediate aftermath of (public) earnings announcements. We find the opposite.

[32] As we measure analyst coverage at the end of the previous month, the level effect of these variables is subsumed by the stock-month fixed effects.

[33] To map this intuition to our model, one can think of a standalone (i.e., one-segment) firm as a firm whose final payoff is uni-dimensional ($\theta = \theta_1$) as opposed to two-dimensional ($\theta = \theta_1 + \theta_2$) as in our baseline model. Obviously, with a uni-dimensional payoff, there is no UWIP and our predictions for return skewness and price impact do not obtain.

[34] In Internet Appendix 2.6, we show that we obtain similar results when we use a dummy variable that identifies conglomerate instead of the number of segments.

[35] Given that there is no obvious analogue for the skewness test, we focus on the price impact regressions for this analysis.

Table 4
Effect of the information environment.

Panel (a): Interaction with EA

	Return skewness	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
		Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past return	-2.0291***	-2.9678***	2.8325***	-0.0980***	0.0254***	-0.2138***	0.0734***	-0.8276***	0.2658***
	(-66.88)	(-55.24)	(52.72)	(-40.09)	(11.00)	(-50.73)	(16.81)	(-85.04)	(26.70)
EA	-0.0296***	-0.0331***	0.0320***	-0.0019***	-0.0012***	-0.0005	0.0006	-0.0226***	-0.0177***
	(-11.07)	(-7.49)	(5.68)	(-6.44)	(-3.44)	(-0.85)	(0.87)	(-20.59)	(-13.23)
Past return*EA	0.5476***	1.0181***	-0.9680***	0.0317***	-0.0109***	0.0474***	-0.0292***	0.2302***	-0.1322***
	(21.54)	(20.81)	(-17.23)	(10.29)	(-3.39)	(8.14)	(-4.63)	(22.29)	(-11.99)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	19,386,514	9,497,462	9,655,532	9,754,434	9,956,693	9,749,027	9,935,083	9,781,451	9,987,895
adj. R^2	0.06	0.13	0.06	0.33	0.28	0.34	0.28	0.69	0.64

Panel (b): Interaction with AnalCov

	Return skewness	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
		Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past return	-3.1621***	-8.9682***	6.3611***	-0.3306***	0.0802***	-0.4888***	0.1579***	-1.7865***	0.5756***
	(-74.15)	(-73.67)	(58.33)	(-54.21)	(14.83)	(-55.46)	(19.20)	(-115.26)	(33.47)
Past return*AnalCov	0.6929***	3.3782***	-2.1464***	0.1334***	-0.0251***	0.1633***	-0.0435***	0.5969***	-0.1845***
	(41.69)	(59.87)	(-40.82)	(47.12)	(-9.23)	(40.26)	(-10.80)	(85.59)	(-24.09)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	23,900,236	11,254,789	11,997,894	11,632,858	12,466,883	11,622,111	12,431,403	11,668,565	12,509,524
adj. R^2	0.07	0.18	0.12	0.35	0.30	0.36	0.30	0.70	0.63

Panel (c): Interaction with Ln(mcap)

	Return skewness	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
		Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past return	-6.9808***	-29.9299***	19.9523***	-1.1031***	0.2691***	-1.3889***	0.4360***	-5.0139***	1.8203***
	(-52.69)	(-63.05)	(44.41)	(-47.85)	(11.75)	(-43.28)	(13.21)	(-93.04)	(28.90)
Past return*Ln(mcap)	0.3926***	2.1085***	-1.3747***	0.0785***	-0.0186***	0.0923***	-0.0280***	0.3321***	-0.1243***
	(40.05)	(57.31)	(-39.21)	(44.97)	(-10.44)	(38.27)	(-10.98)	(81.83)	(-26.05)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	23,814,464	11,235,102	11,975,676	11,589,146	12,411,244	11,578,495	12,376,175	11,623,981	12,452,720
adj. R^2	0.07	0.18	0.12	0.35	0.30	0.36	0.30	0.70	0.63

Panel (d): Interaction with $\#Segm^{orth}$

	Return skewness	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
		Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past return	-2.2308***	-4.3843***	3.8170***	-0.1544***	0.0520***	-0.2723***	0.1078***	-0.9683***	0.3550***
	(-71.15)	(-60.29)	(56.82)	(-46.90)	(17.97)	(-55.85)	(22.62)	(-88.87)	(32.42)
Past return*$\#Segm^{orth}$	-0.0781**	-1.3792***	1.0867***	-0.0527***	0.0273***	-0.0611***	0.0230***	-0.1378***	0.1386***
	(-2.28)	(-11.67)	(11.15)	(-9.10)	(5.32)	(-7.06)	(2.90)	(-7.92)	(8.94)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	19,570,638	9,397,769	9,799,682	9,669,922	10,131,991	9,664,368	10,109,367	9,697,134	10,164,345
adj. R^2	0.06	0.17	0.12	0.37	0.32	0.37	0.31	0.70	0.64

This table reports results of cross-sectional tests for our return skewness and price impact regressions (see Section 3.2). For brevity, we only show results using lagged cumulated 10-day returns as the key independent variable. Panel (a) shows results for interacting the past return with an *EA* dummy, which flags whether or not there was an earnings announcement for the stock over the previous 10 trading days. The *EA* dummy is also added separately to the regression. Panel (b) shows results for interacting the past return with *AnalCov*, defined as the natural logarithm of one plus the number of analysts following the stock at the end of the prior month. *AnalCov* is not added separately to the regression as it is subsumed by the stock-month fixed effect. Panel (c) shows results for interacting the past return with *Ln(mcap)*, defined as the natural logarithm of the stock's market capitalization at the end of the prior month. *Ln(mcap)* is not added separately to the regression as it is subsumed by the stock-month fixed effect. Panel (d) shows results for interacting the past return with $\#Segm^{orth}$, defined as the residual of regressing the logarithm of the number of business segments the firm operates in (based on sales generated in different 2-digit SIC industries as in Cohen and Lou, 2012) on *Ln(mcap)*. $\#Segm^{orth}$ is not added separately to the regression as it is subsumed by the stock-month fixed effect. In each panel, Column (1) shows results for *return skewness* as the dependent variable (using the overall sample), Columns (2)–(3) show results for *price impact costs* as the dependent variable, Columns (4)–(5) show results for *lambda* as the dependent variable, Columns (6)–(7) show results for *quote-based price impact* as the dependent variable, and Columns (8)–(9) show results for *ln(Amihud)* as the dependent variable. Columns (2), (4), (6), and (8) show results for the net-buy subsample (i.e., all stock-day observations on which the trade imbalance is positive). Columns (3), (5), (7), and (9) show results for the net-sell subsample (i.e., all stock-day observations on which the trade imbalance is negative). All regressions contain past turnover and past volatility controls, 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.

Table 5
Institutional vs. retail trade imbalance.
Panel (a): Split based on institutional order imbalance

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Past return	-3.3249***	3.1638***	-0.0732***	0.0338***	-0.1369***	0.0350***	-0.6850***	0.0900***
	(-31.56)	(29.97)	(-16.58)	(8.08)	(-19.96)	(5.14)	(-48.50)	(6.56)
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	3,752,931	4,010,100	3,861,532	4,127,163	3,859,456	4,125,076	3,858,335	4,123,826
adj. R^2	0.10	0.07	0.17	0.16	0.31	0.28	0.78	0.77

Panel (b): Split based on retail order imbalance

	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Past return	-0.4240***	0.3101***	-0.0189***	-0.0116***	-0.0546***	-0.0244***	-0.3366***	-0.2292***
	(-7.81)	(5.85)	(-6.27)	(-3.76)	(-9.52)	(-4.24)	(-27.07)	(-18.85)
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	3,245,395	3,772,896	3,340,000	3,872,364	3,339,477	3,871,691	3,339,856	3,872,166
adj. R^2	0.08	0.09	0.24	0.24	0.34	0.33	0.80	0.79

This table reports results of our price impact regressions, except that we now split the sample by institutional or retail trade imbalance instead of by (total) trade imbalance (see Section 3.2). For brevity, we only show results using lagged cumulated 10-day returns as the key independent variable. In Panel (a), stock-days are split into net buys and net sells using the institutional trade imbalance (which is defined as total trade imbalance minus retail trade imbalance). In Panel (b), stock-days are split into net buys and net sells using the retail trade imbalance (which is defined as retail buys minus retail sells). Retail buys and sells are identified using the Boehmer et al. (2021) methodology over the 2004–2014 subperiod. In each panel, Columns (1)–(2) show results for *price impact costs* as the dependent variable, Columns (3)–(4) show results for *lambda* as the dependent variable, Columns (5)–(6) show results for *quote-based price impact* as the dependent variable, and Columns (7)–(8) show results for *ln(Amihud)* as the dependent variable. Columns (1), (3), (5), and (7) show results for the net-buy subsample (i.e., all stock-day observations on which the institutional or retail trade imbalance is positive). Columns (2), (4), (6), and (8) show results for the net-sell subsample. All regressions contain past turnover and past volatility controls, 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.

3.6. Robustness and alternative explanations

3.6.1. Inventory risk channel

As argued before, the asymmetry in price impact due to UWIP contrasts with the asymmetry predicted by models of inventory risk (e.g., Hendershott and Menkveld, 2014). In Internet Appendix 2.3, we confirm that our findings regarding the dependence of skewness and price impact on past returns are robust to controlling for the past trade imbalance (measured over the same lookback window), which proxies for the importance of inventory considerations.[36] A related concern is that both price impact and past returns are endogenous to a stock's liquidity. For instance, improvements in liquidity may lead to positive returns and lower future price impact. Note that our baseline regressions already control for slow-moving liquidity changes by including stock-month fixed effects. In Internet Appendix 2.4, we further show that our results are largely unaffected if we control for past average realized spreads (or bid–ask spreads computed from CRSP data).

3.6.2. S&P 500 index additions and deletions

In Table 4, we have shown that the effect of UWIP weakens for stocks with higher analyst coverage and larger capitalization (especially those operating in fewer business segments), suggesting that public attention helps reduce UWIP. In further support for this interpretation, we analyze stocks added to and deleted from the S&P 500 index. Prior research finds that stocks in the S&P 500 receive heightened analyst following and public scrutiny (Denis et al., 2003; Chen et al., 2004). We thus expect stocks added to (dropped from) the S&P 500 to experience a reduction (an increase) in UWIP and in the associated link between skewness/price impact and past returns. As Table 6 Panels (a) and (b) demonstrate, this is indeed what we find.

Specifically, in Panel (a), we focus on stocks added to the S&P 500 during our sample period, and restrict our sample to ±5 years around the additions.[37] We run our baseline regressions of skewness and price impact on past returns, including the interaction of past returns with a S&P dummy that flags the post-addition period. The interaction's coefficient goes consistently in the opposite direction to the coefficient on past return, though the economic and statistical significance of the effect are weaker than in our baseline regressions (due to the smaller sample size). In Panel (b), we repeat this analysis for stocks dropped from the S&P 500. The coefficient of interest—the interaction of past returns with a dummy flagging the post-deletion period—now goes in the same direction as the coefficient on past returns (although for skewness the interaction coefficient is

[36] In models à la Kyle (1985), past returns and past trade imbalance are observationally equivalent, which is why we do not include trade imbalance as a control in our baseline specification.

[37] To ensure our results are not confounded by the well-known index inclusion effect, we drop the month of the actual index inclusion from the analysis. In Internet Appendix 2.7, we confirm that obtain similar (albeit slightly weaker) results when we restrict the sample to ±2 years around S&P 500 additions/deletions.

Table 6
Alternative explanations.

Panel (a): S&P 500 additions (+/- 5 years)

	Return skewness	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
		Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past return	-1.2654***	-0.1040***	0.1952***	-0.0123***	-0.0026	-0.0991***	0.0202**	-0.1599***	0.0164
	(-21.05)	(-6.10)	(3.77)	(-5.53)	(-0.91)	(-13.34)	(2.32)	(-14.65)	(1.41)
Past return*SP500	0.4777***	0.0212	-0.0252	0.0064***	0.0021	0.0203**	-0.0252**	0.0848***	-0.0077
	(7.85)	(0.79)	(-0.35)	(2.77)	(0.68)	(2.29)	(-2.35)	(6.26)	(-0.54)
t-statistic of diff. of interaction coef.		(0.58)		(1.21)		(3.25)***		(5.20)***	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	986,131	579,780	392,873	586,564	398,214	586,183	397,895	586,796	398,491
adj. R^2	0.05	0.08	0.02	0.29	0.22	0.46	0.41	0.77	0.79

Panel (b): S&P 500 deletions (+/- 5 years)

	Return skewness	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
		Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past return	-1.0299***	-0.1981***	0.3037***	-0.0131***	0.0003	-0.1118***	-0.0074	-0.1511***	0.1361***
	(-14.71)	(-4.86)	(3.59)	(-6.00)	(0.15)	(-9.16)	(-0.50)	(-8.32)	(5.70)
Past return*exSP500	-0.0018	-0.5051***	0.4857***	-0.0332***	-0.0042	-0.0409	0.0273	-0.3063***	-0.0107
	(-0.02)	(-4.21)	(2.90)	(-3.42)	(-0.67)	(-1.48)	(0.96)	(-7.10)	(-0.28)
t-statistic of diff. of interaction coef.		(-4.20)***		(-2.80)***		(-2.01)**		(-5.10)***	
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	445,566	247,540	199,932	248,138	200,519	248,138	200,524	248,232	200,640
adj. R^2	0.04	0.08	0.00	0.43	0.36	0.45	0.38	0.63	0.59

Panel (c): Interaction with ShortFee

	Return skewness	Price impact costs		Lambda		Quote-based price impact		Ln(Amihud)	
		Net buys	Net sells	Net buys	Net sells	Net buys	Net sells	Net buys	Net sells
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Past return	-0.9709***	-3.6755***	3.6487***	-0.0577***	0.0274***	-0.1210***	0.0267***	-0.6231***	0.1014***
	(-37.22)	(-33.63)	(32.16)	(-14.14)	(6.89)	(-17.84)	(4.05)	(-43.21)	(7.27)
Past return*ShortFee	-0.0050	-0.0820***	-0.0343**	-0.0026***	-0.0011*	-0.0027**	-0.0008	-0.0227***	-0.0131***
	(-1.54)	(-4.33)	(-2.55)	(-3.56)	(-1.79)	(-2.11)	(-0.73)	(-9.91)	(-6.63)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Stock-month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Date FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	7,833,642	3,613,547	3,902,991	3,711,747	4,007,682	3,710,114	4,006,141	3,709,722	4,005,653
adj. R^2	0.08	0.10	0.06	0.16	0.15	0.31	0.28	0.78	0.77

This table reports results of additional cross-sectional tests for our return skewness and price impact regressions (see Section 3.2). For brevity, we only show results using lagged cumulated 10-day returns as the key independent variable. In Panel (a), we focus on stocks that are added to the S&P 500 index during our sample period. For each added stock, we include 5 years before and after the addition event (while excluding the month of the addition event). We show results for interacting the past return with $SP500$, a dummy that takes the value one after the stock is added to the S&P 500 and zero before. $SP500$ is not added separately to the regression as it is subsumed by the stock-month fixed effect. In Panel (b), we focus on stocks that are dropped from the S&P 500 index during our sample period (including 5 years before and after the deletion event). We show results for interacting the past return with $exSP500$, a dummy that takes the value one after the stock is dropped from the S&P 500 and zero before. $exSP500$ is not added separately to the regression as it is subsumed by the stock-month fixed effect. At the bottom of Panels (a) and (b), we show t-statistics for whether the coefficients of the interaction of the past return with $SP500$ or $exSP500$ are statistically significantly different between the subsamples of days with net buys and net sells. In Panel (c), we show results for interacting the past return with *ShortFee*, defined as the average short selling fee (i.e., equity lending fee) for the stock over the prior month (data comes from Markit and is only available to us for the period from July 2006 onward). *ShortFee* is not added separately to the regression as it is subsumed by the stock-month fixed effect. In each panel, Column (1) shows results for *return skewness* as the dependent variable (using the overall sample), Columns (2)–(3) show results for *price impact costs* as the dependent variable, Columns (4)–(5) show results for *lambda* as the dependent variable, Columns (6)–(7) show results for *quote-based price impact* as the dependent variable, and Columns (8)–(9) show results for *ln(Amihud)* as the dependent variable. Columns (2), (4), (6), and (8) show results for the net-buy subsample (i.e., all stock-day observations on which the trade imbalance is positive). Columns (3), (5), (7), and (9) show results for the net-sell subsample (i.e., all stock-day observations on which the trade imbalance is negative). All regressions contain past turnover and past volatility controls, 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.

not significant). Hence, as expected, we find that the empirical trace of UWIP diminishes after index additions, while it strengthens after index deletions.[38]

[38] In Internet Appendix 3.2, we show that we obtain consistent, albeit statistically weaker, results for stocks switching between the Russell 1000 and 2000 indices (see, e.g., Appel et al., 2016; Sammon, 2023; Coles et al., 2022). This is expected as, unlike S&P 500 inclusions, Russell reconstitution events induce primarily variations in passive ownership, and it is less clear how UWIP depends on the extent of passive ownership.

3.6.3. Short sale constraints

Finally, we test whether the dependence of skewness and price impact on past returns is mediated by short sale constraints. We do so as prior explanations for this dependence rely on short sale constraints (e.g., Saar, 2001; Hong and Stein, 2003; Xu, 2007). For instance, in Saar (2001), sells become relatively more informed after positive past returns because the stock is then more likely to be held by informed mutual funds (who bought the stock during the price run-up), implying that the short sale constraint binds less. This leads to an asymmetry in price impact between buys and sells that varies as a function of past returns. Similarly, in Hong and Stein (2003), short sale constraints prevent the views of bearish investors from being incorporated into prices. Their accumulated hidden information comes out during market declines, thus causing negative return skewness. In each case, short sale constraints are at the root of the asymmetry—thus, according to these theories, asymmetries in skewness and price impact should increase (decrease) as short sale constraints tighten (loosen).

To proxy for the tightness of short sale constraints, we use equity lending fees from IHS Markit, the leading provider of such data.[39] The higher the lending fee, the more expensive it is to borrow the stock and thus the more constrained is short selling. We then interact the stock's past returns with its average lending fee over the previous month. Table 6 Panel (c) shows the results. We find that short selling costs have no bearing on the dependence of return skewness on past returns (Column 1). This is inconsistent with the Hong and Stein (2003) and Xu (2007) explanations and suggests that—at least at the daily frequency—the observed skewness-return relation is more likely driven by a mechanism that does not rely on short sale constraints such as ours. For price impact (Columns 2–9), we find that short sale constraints do not increase the price impact asymmetry between buys and sells (as the interaction coefficients are negative on both net-buy and net-sell days). In particular, the negative interaction coefficients on net-sell days are inconsistent with the Saar (2001) model, as they imply that short sale constraints weaken, rather than strengthen, the price impact of sells observed after positive returns.[40] This suggest that our results are not driven by short sale constraints.

In summary, the results in this section are in strong agreement with Hypotheses H1 and H2, as well as with Hypotheses H1′, H2′, and H2′′, while being inconsistent with alternative explanations relying on short sale constraints. They thus lend support to the idea that UWIP is a real and important concern for investors.

4. Conclusion

This paper proposes a simple model in which speculators are unsure whether their signals are stale (i.e., already priced in) 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 their trading aggressiveness accordingly. Market makers, in turn, anticipate that speculators may be trading on stale news. The resulting price function is 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. As a results, return skewness is negatively related to past price changes. Moreover, by making speculators reluctant to trade, uncertainty about what is 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 these predictions. Specifically, we document that (1) return skewness is negatively associated with past returns and that (2) on days with a positive (negative) trade imbalance, price impact costs are negatively (positively) related to past stock returns. Moreover, we find that these dependencies are reduced after earnings announcements and for stocks with a large market capitalization and a high analyst coverage; i.e., when speculators know better whether their private signals are novel or stale. Overall, our results strongly suggest that uncertainty about what is in the price is a common and widespread concern for stock market participants.

CRediT authorship contribution statement

Joël Peress: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Daniel Schmidt:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors have nothing to disclose.

[39] Our equity lending data spans the period from July 2006 to the end of our sample period.

[40] We note that our baseline model cannot explain why we find a negative interaction coefficient between short sale constraints and past returns. However, we suspect that an extended model with short sale constraints can deliver this prediction. To see this, consider a version of our model in which the $t = 0$ price reflects the trading by another group of informed investors. If these investors are short sale constrained, $t = 0$ prices will be less informative and so less helpful for assessing the staleness of a signal. Hence, when determining trading aggressiveness and price impact, market participants put less weight on past returns.

Data availability

sample data (Reference data) (Mendeley Data)

Appendix A. Proofs

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

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

Rearranging leads to the cubic equation:

$$9 - 6x_{(2)} - 5\gamma\sigma x_{(2)}^2 + 4\gamma\sigma x_{(2)}^3 = 0 \tag{A.1}$$

- 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

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

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

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

which again leads to Eq. (A.1). Therefore, 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 3 - 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. 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, $E(\theta_{\neg m} | \theta_m = \sigma, -x_{(3)} + 1 \leq \omega_1 \leq y_{(3)} + 1) = \frac{1/4 \times 0 + 1/8\sigma}{1/4 + 1/8} = \frac{1}{3}\sigma$. 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 sold $x_{(3)}$ and hence that $\theta_s = -\sigma$. This tells them that $m \neq s$ and therefore that $E(\theta_{\neg m} | \theta_m = \sigma, -x_{(3)} - 1 \leq \omega_1 \leq y_{(3)} - 1) = -\sigma$. Finally, suppose M observes an order flow $y_{(3)} - 1 \leq \omega_1 \leq 1 - x_{(3)}$. Then M learn that S could have sold $x_{(3)}$ or bought $y_{(3)}$, and so cannot draw any inference on θ_s. In that case, $E(\theta_{\neg m} | \theta_m = \sigma, y_{(3)} - 1 \leq \omega_1 \leq 1 - x_{(3)}) = 0$. The analysis is similar if $\theta_m = -\sigma$. For example, $E(\theta_{\neg m} | \theta_m = -\sigma, -y_{(3)} - 1 \leq \omega_1 \leq x_{(3)} - 1) = -\frac{1}{3}\sigma$.

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 Eqs. (3) and (4) lead to the following, starting with the case $m = s$:

- Case 1. 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

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

- Case 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 again S buy $y_{(3)}$ shares so $\omega_1 = y_{(3)} + n$ and

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

- Case 3. 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

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

- Case 4. 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

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

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

- Case 5. 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

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

- Case 6. 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

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

- Case 7. 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

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

- Case 8. 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

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

Collecting the cases such that $\theta_m = \theta_s = \sigma$ (cases 1, 2 and 5) leads to $E\left[\theta - p_1 | \theta_m = \sigma, \theta_s = \sigma\right] = \frac{1}{3}\sigma(1-z)$ and $Var\left[\theta - p_1 | \theta_m = \sigma, \theta_s = \sigma\right] = \frac{1}{9}\sigma^2(8 + z - z^2)$. Plugging these expressions into the first-order condition for S' profit maximization and imposing rational expectations ($x_i = x_{(3)}$ and $y_i = y_{(3)}$ for all i) yields:

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

Likewise, collecting the cases such that $\theta_m = \theta_s = -\sigma$ (cases 3, 4 and 8) leads to $E\left[\theta - p_1 | \theta_m = -\sigma, \theta_s = -\sigma\right] = -\frac{1}{3}\sigma(1-z)$ and $Var\left[\theta - p_1 | \theta_m = \sigma, \theta_s = \sigma\right] = \frac{1}{9}\sigma^2(8 + z - z^2)$. Plugging these expressions into the first-order condition for S' profit maximization and imposing rational expectations ($x_i = x_{(3)}$ and $y_i = y_{(3)}$ for all i) yields:

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

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

Collecting the cases such that $\theta_m = \sigma$ and $\theta_s = -\sigma$ (case 6) leads to $E\left[\theta - p_1 | \theta_m = \sigma, \theta_s = -\sigma\right] = -\sigma(1-z)$ 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 ($x_i = x_{(3)}$ and $y_i = y_{(3)}$ for all i) yields:

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

Finally, collecting the cases such that $\theta_m = -\sigma$ and $\theta_s = \sigma$ (case 7) leads to $E\left[\theta - p_1 | \theta_m = -\sigma, \theta_s = \sigma\right] = \sigma(1-z)$ 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 ($x_i = x_{(3)}$ and $y_i = y_{(3)}$ for all i) yields:

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

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

Gathering 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 second equation and rearranging leads to the quartic equation:

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

The equilibrium is thus characterized by Eq. (A.2), together with the requirement that $x_{(3)} \geq y_{(3)}$. We show next that there exists a unique equilibrium.

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 $[1/\sqrt{\gamma\sigma}, 1]$, implying that there exists a unique equilibrium. The second derivative of g, $g''(x) = -4\gamma\sigma(1+4\gamma\sigma) + 12(\gamma\sigma)^2 x + 48(\gamma\sigma)^3 x^2$, 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 $[1/\sqrt{\gamma\sigma}, 1]$. Hence, there exists a unique equilibrium.

A.4. Corollary 7 - Price informativeness

1. **Both m and s are common knowledge**

   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 $Var(\theta) = 2\sigma^2$. It follows that $PI_{(1)} \equiv Var\left(E\left(\theta|p_1,p_0\right)\right) = Var(\theta) - E\left(Var\left(\theta|p_1,p_0\right)\right) = 2\sigma^2 - \sigma^2\left(1-\frac{1}{2}x_{(1)}\right) = \sigma^2\left(1+\frac{1}{2}x_{(1)}\right)$. 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 $m=s$.

2. **Only m is common knowledge**

   When only m is common knowledge, price informativeness is given by:

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

   Therefore $PI_{(2)} \equiv Var\left(E\left(\theta|p_1,p_0\right)\right) = Var(\theta) - E\left(Var\left(\theta|p_1,p_0\right)\right) = 2\sigma^2 - \sigma^2\left(1-\frac{1}{3}x_{(2)}\right) = \sigma^2\left(1+\frac{1}{3}x_{(2)}\right)$. 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 ($\frac{1}{3}\sigma^2/(2\sigma^2) = \frac{1}{6}$) of the total uncertainty is resolved through the trading by S.

3. **Neither m nor s are common knowledge**

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

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

$$
= \frac{1}{2}\left[\begin{array}{c} Pr(m=s,\theta_s=\sigma,\theta_{\neg m}=\sigma)\,E\left(\left(\theta-p_1\right)^2|m=s,\theta_s=\sigma,\theta_{\neg m}=\sigma\right) \\ +Pr(m=s,\theta_s=\sigma,\theta_{\neg m}=-\sigma)\,E\left(\left(\theta-p_1\right)^2|m=s,\theta_s=\sigma,\theta_{\neg m}=-\sigma\right) \\ +Pr(m=s,\theta_s=-\sigma,\theta_{\neg m}=\sigma)\,E\left(\left(\theta-p_1\right)^2|m=s,\theta_s=-\sigma,\theta_{\neg m}=\sigma\right) \\ +Pr(m=s,\theta_s=-\sigma,\theta_{\neg m}=-\sigma)\,E\left(\left(\theta-p_1\right)^2|m=s,\theta_s=-\sigma,\theta_{\neg m}=-\sigma\right) \end{array}\right]
$$

$$
+\frac{1}{2}\left[\begin{array}{c} Pr(m\neq s,\theta_s=\sigma,\theta_m=\sigma)\,E\left(\left(\theta-p_1\right)^2|m\neq s,\theta_s=\sigma,\theta_m=\sigma\right) \\ +Pr(m\neq s,\theta_s=\sigma,\theta_m=-\sigma)\,E\left(\left(\theta-p_1\right)^2|m\neq s,\theta_s=\sigma,\theta_m=-\sigma\right) \\ +Pr(m\neq s,\theta_s=-\sigma,\theta_m=\sigma)\,E\left(\left(\theta-p_1\right)^2|m\neq s,\theta_s=-\sigma,\theta_m=\sigma\right) \\ +Pr(m\neq s,\theta_s=-\sigma,\theta_m=-\sigma)\,E\left(\left(\theta-p_1\right)^2|m\neq s,\theta_s=-\sigma,\theta_m=-\sigma\right) \end{array}\right]
$$

$$
= \frac{1}{2}\left[\begin{array}{c} \frac{1}{4}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{4}\left(\left(\frac{4}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right) \\ +\frac{1}{4}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{4}\left(\left(\frac{4}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right) \end{array}\right]
$$

$$
+\frac{1}{2}\left[\begin{array}{c} \frac{1}{4}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{4}\left(\sigma^2\left(1-\frac{x+y}{2}\right)\right) \\ \frac{1}{4}\left(\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{4}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right) \end{array}\right]
$$

$$
= \frac{1}{2}\left[\frac{1}{2}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{2}\left(\left(\frac{4}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)\right]
$$

$$
+\frac{1}{2}\left[\frac{1}{2}\left(\left(\frac{2}{3}\sigma\right)^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right)+\frac{1}{2}\left(\sigma^2\left(1-\frac{x+y}{2}\right)\right)\right]
$$

$$
= \frac{1}{2}\left[\frac{10}{9}\sigma^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right]
$$

$$
+\frac{1}{2}\left[\frac{2}{9}\sigma^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)\right]
$$

$$
= \frac{2}{3}\sigma^2\frac{x+y}{2}+\sigma^2\left(1-\frac{x+y}{2}\right)
$$

$$
= \sigma^2\left(1-\frac{1}{3}\frac{x+y}{2}\right)
$$

Hence $PI_{(3)} \equiv Var\left(E\left(\theta|p_1,p_0\right)\right) = Var(\theta) - E\left(Var\left(\theta|p_1,p_0\right)\right) = 2\sigma^2 - \sigma^2\left(1-\frac{1}{3}\frac{x+y}{2}\right) = \sigma^2\left(1+\frac{1}{3}\frac{x+y}{2}\right)$. 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 total uncertainty would be further reduced by one sixth.

Ranking of price informativeness

We show that $PI_{(1)} > PI_{(3)}$. In the proof of Proposition 3, we establish that $x_{(3)} > 1/\sqrt{\gamma\sigma} = x_{(1)}$. This inequality implies that $1/(\gamma\sigma x_{(3)}) < 1/\sqrt{\gamma\sigma} = x_{(1)}$. Moreover, we show, also in the proof of Proposition 3, that $(y_{(3)} + x_{(3)})/2 = 1/(\gamma\sigma x_{(3)})$. Combining both expressions leads to $(y_{(3)} + x_{(3)})/2 < x_{(1)}$ and so to $x_{(1)} > (y_{(3)} + x_{(3)})/3$. Hence, $PI_{(1)} > PI_{(3)}$. Note also that $y_{(3)} = \frac{2}{\gamma\sigma x_{(3)}} - x_{(3)} < 2/\sqrt{\gamma\sigma} - 1/\sqrt{\gamma\sigma} = 1/\sqrt{\gamma\sigma} = x_{(1)}$ so $0 < y_{(3)} < x_{(1)} < x_{(3)} < 1$.

We show next that $PI_{(2)} > PI_{(3)}$. Given the expressions for $PI_{(2)}$ and $PI_{(3)}$, this inequality is equivalent to $\frac{x_{(3)}+y_{(3)}}{2} < x_{(2)}$. We employ again two results established in the proof of Proposition 3: $\frac{x_{(3)}+y_{(3)}}{2} = \frac{1}{\gamma\sigma x_{(3)}}$ and $x_{(3)} > 1/\sqrt{\gamma\sigma} = x_{(1)}$. They imply that $\frac{x_{(3)}+y_{(3)}}{2} < 1/\sqrt{\gamma\sigma} = x_{(1)}$, so it suffices to show that $x_{(1)} < x_{(2)}$. To do so, we show that $f(x_{(1)}) > 0$ where f is the decreasing function defined in the proof of Proposition 1, of which $x_{(2)}$ is a root: $f(x_{(1)}) = 9 - 6x_{(1)} - 5\gamma\sigma x_{(1)}^2 + 4\gamma\sigma x_{(1)}^3 = 9 - 6(1/\sqrt{\gamma\sigma}) - 5\gamma\sigma(1/\sqrt{\gamma\sigma})^2 + 4\gamma\sigma(1/\sqrt{\gamma\sigma})^3 = 4 - 2/\sqrt{\gamma\sigma} > 0$ for $\gamma\sigma > 3$. Hence, $f(x_{(1)}) > 0 = f(x_{(2)})$, which in turn implies $x_{(1)} < x_{(2)}$. Thus, $PI_{(2)} > PI_{(3)}$.

Intuition

S are most aggressive in case (2) when $m \neq s$; that is, when they have an information advantage but when market makers do not 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 is 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).[41] 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 is in the price.

[41] 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).

References

Abreu, D., Brunnermeier, M.K., 2002. Synchronization risk and delayed arbitrage. J. Financ. Econ. 66, 341–360.
Abreu, D., Brunnermeier, M.K., 2003. Bubbles and crashes. Econometrica 71, 173–204.
Amihud, Y., 2002. Illiquidity and stock returns. J. Financial Mark. 5, 31–56.
Appel, I., Gormley, T.A., Keim, D.B., 2016. Passive investors, not passive owners. J. Financ. Econ. 121, 111–141.
Avery, C., Zemsky, P., 1998. Multidimensional uncertainty and herd behavior in financial markets. Amer. Econ. Rev. 88, 725–748.
Banerjee, S., Breon-Drish, B., 2020. Strategic trading and unobservable information acquisition. J. Financ. Econ. 138, 458–482.
Banerjee, S., Green, B., 2015. Signal or noise? Uncertainty and learning about whether other traders are informed. J. Financ. Econ. 117, 398–423.
Barber, B.M., Odean, T., 2007. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21, 785–818.
Blume, L., Easley, D., O'Hara, M., 1994. Market statistics and technical analysis: The role of volume. J. Finance 49, 153–181.
Boehmer, E., Jones, C.M., Zhang, X., Zhang, X., 2021. Tracking retail investor activity. J. Finance 76, 2249–2305.
Brown, D.P., Jennings, R.H., 1989. On technical analysis. Rev. Financ. Stud. 2, 527–551.
Brunnermeier, M.K., 2005. Information leakage and information efficiency. Rev. Financ. Stud. 18, 417–457.
Chen, Q., Goldstein, I., Jiang, W., 2007. Price informativeness and investment sensitivity to stock price. Rev. Financ. Stud. 20, 619–650.
Chen, J., Hong, H., Stein, J.C., 2001. Forecasting crashes: Trading volume, past returns, and conditional skewness in stock prices. J. Financ. Econ. 61, 345–381.
Chen, H., Noronha, G., Singal, V., 2004. The price response to S&P 500 index additions and deletions: Evidence of asymmetry and a new explanation. J. Finance 59, 1901–1930.
Cohen, L., Lou, D., 2012. Complicated firms. J. Financ. Econ. 104, 383–400.
Coles, J.L., Heath, D., Ringgenberg, M.C., 2022. On index investing. J. Financ. Econ. 145, 665–683.
Denis, D.K., Mcconnell, J.J., Ovtchinnikov, A., Yu, Y., 2003. S&P 500 index additions and earnings expectations. J. Finance 58, 1821–1840.
Dessaint, O., Foucault, T., Fresard, L., Matray, A., 2019. Noisy stock prices and corporate investment. Rev. Financ. Stud. 32, 2625–2672.
Easley, D., Hvidkjaer, S., O'Hara, M., 2002. Is information risk a determinant of asset returns? J. Finance 57, 2185–2221.
Easley, D., Kiefer, N., O'Hara, M., 1996. Cream-skimming or profit-sharing? The curious role of purchased order flow. J. Finance 51, 811–833.
Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment. J. Finance 47, 577–605.
Easley, D., O'hara, M., 2005. Information and the cost of capital. J. Finance 59, 1553–1583.
Fang, L., Peress, J., 2009. Media coverage and the cross-section of stock returns. J. Finance 64, 2023–2052.
Fedyk, A., Hodson, J., 2023. When can the market identify old news? J. Financ. Econ. 149, 92–113.
Foucault, T., Fresard, L., 2012. Cross-listing, investment sensitivity to stock price, and the learning hypothesis. Rev. Financ. Stud. 25, 3305–3350.
Foucault, T., Pagano, M., Roell, A., 2013. Market Liquidity. Oxford University Press.
Foucault, T., Sraer, D., Thesmar, D., 2011. Individual investors and volatility. J. Finance 66, 1369–1406.
Frazzini, A., Israel, R., Moskowitz, T.J., 2018. Trading costs. Working Paper.
Gao, F., Song, F., Wang, J., 2013. Rational expectations equilibrium with uncertain proportion of informed traders. J. Financial Mark. 16, 387–413.
Gilbert, T., Kogan, S., Lochstoer, L., Ozyildirim, A., 2012. Investor inattention and the market impact of summary statistics. Manage. Sci. 58, 336–350.
Glosten, L.A., Milgrom, P.R., 1985. Bid, ask and transaction prices in a specialist market with heterogenously informed traders. J. Financ. Econ. 14, 71–100.
Goyenko, R.Y., Holden, C.W., Trzcinka, C.A., 2009. Do liquidity measures measure liquidity? J. Financ. Econ. 153–181.
Grossman, S., 1976. On the efficiency of competitive stock markets where trades have diverse information. J. Finance 31, 573–585.
Grossman, S.J., Miller, M.H., 1988. Liquidity and market structure. J. Finance 43, 617–633.
Grossman, S.J., Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. Amer. Econ. Rev. 70, 393–408.
Grundy, B.D., McNichols, M., 1989. Trade and the revelation of information through prices and direct disclosure. Rev. Financ. Stud. 2, 495–526.
Harvey, C.R., Siddique, A., 2000. Conditional skewness in asset pricing tests. J. Finance 55, 1263–1295.
Hasbrouck, J., 2009. Trading costs and returns for U.S. equities: Estimating effective costs from daily data. J. Finance 64, 1445–1477.
Hellwig, M.F., 1980. On the aggregation of information in competitive markets. J. Econom. Theory 22, 477–498.
Hendershott, T., Livdan, D., Schuerhoff, N., 2015. Are institutions informed about news? J. Financ. Econ. 117, 249–287.
Hendershott, T., Menkveld, A.J., 2014. Price pressures. J. Financ. Econ. 114, 405–423.
Ho, T., Stoll, H., 1981. Optimal dealer pricing under transaction and return uncertainty. J. Financ. Econ. 9 (47–73), 47–73.
Holden, C.W., Jacobsen, S., 2014. Liquidity measurement problems in fast, competitive markets: Expensive and cheap solutions. J. Finance 69, 1747–1785.
Hong, H., Stein, J.C., 2003. Differences of opinions, shore-sales constraints, and market crashes. Rev. Financ. Stud. 16, 487–525.
Huberman, G., Regev, T., 2001. Contagious speculation and a cure for cancer: A nonevent that made stock prices soar. J. Finance 56, 387–396.
Keim, D.B., Madhavan, A., 1996. The upstairs market for large-block transactions: Analysis and measurement of price effects. Rev. Financ. Stud. 9, 1–36.
Kraus, A., Stoll, H.R., 1972. Price impacts of block trading on the new york stock exchange. J. Finance 27, 569–588.
Kyle, A., 1985. Continuous auctions and insider trading. Econometrica 53, 1461–1484.
Lee, C.M.C., Ready, M.J., 1991. Inferring trade direction from intraday data. J. Finance 46, 733–746.
Llorente, G., Michaely, R., Saar, G., Wang, J., 2002. Dynamic volume-return relation of individual stocks. Rev. Financ. Stud. 15, 1005–1047.
Luo, Y., 2005. Do insiders learn from outsiders? Evidence from mergers and acquisitions. J. Finance 60, 1951–1982.
Madhavan, A., Smidt, S., 1993. An analysis of changes in specialist inventories and quotations. J. Finance 48, 1595–1628.
Papadimitriou, D., 2023. Trading under uncertainty about other market participants. Financial Review 58, 343–367.
Peress, J., Schmidt, D., 2020. Glued to the TV: Distracted noise traders and stock market liquidity. J. Finance 75, 1083–1133.
Saar, G., 2001. Price impact asymmetry of block trades: An institutional trading explanation. Rev. Financ. Stud. 14, 1153–1181.
Sammon, M., 2023. Passive Ownership and Price Informativeness. Working Paper.
Schneider, J., 2009. A rational expectations equilibrium with informative trading volume. J. Finance 64, 2783–2805.
Tetlock, P.C., 2011. All the news that's fit to reprint: Do investors react to stale information? Rev. Financ. Stud. 24, 1481–1512.
Treynor, J.L., Ferguson, R., 1985. In defense of technical analysis. J. Finance 40, 757–775.
Vives, X., 1995. Short-term investment and the informational efficiency of the market. Rev. Financ. Stud. 8, 125–170.
Weller, B.M., 2018. Does algorithmic trading reduce information acquisition? Rev. Financ. Stud. 31, 2184–2226.
Xu, J., 2007. Price convexity and skewness. J. Finance 62, 2521–2552.