Identifying Price Informativeness*

Eduardo Dávila[†] Cecilia Parlatore[‡]

September 2020

Abstract

This paper shows how to identify and estimate price informativeness. Starting from i) an asset pricing equation and ii) a stochastic process for asset payoffs, we show how to use regressions of changes in asset prices on changes in asset payoffs to recover an exact measure of price informativeness in a large class of environments. We implement our methodology empirically computing a panel of stock-specific measures of price informativeness for U.S. stocks between 1980 and 2017. In the cross section, we find that large stocks, stocks that with high turnover, and stocks with high institutional ownership have higher price informativeness. In the time series, we find that the median, mean, and standard deviation of the distribution of price informativeness have steadily increased since the mid 1980's.

JEL Classification: D82, D83, G14

Keywords: price informativeness, informational efficiency, information aggregation

*This a substantially revised version of earlier drafts. We would like to thank our discussants Xavier Vives and Harry Mamaysky, Joe Altonji, Yakov Amihud, Philip Bond, Ricardo Caballero, Jennifer Carpenter, John Campbell, John Cochrane, Ian Dew-Becker, Darrell Duffie, Emmanuel Farhi, Itay Goldstein, Gary Gorton, Joel Hasbrouck, Greg Kaplan, Ralph Koijen, Arvind Krishnamurthy, Ben Lester, Guido Lorenzoni, Stephen Morris, Giuseppe Moscarini, Stefan Nagel, Guillermo Ordoñez, Monika Piazzesi, David Sraer, Martin Schneider, David Thesmar, Aleh Tsyvinski, Harald Uhlig, Moto Yogo, Kathy Yuan, Wei Xiong, Laura Veldkamp, Annette Vissing-Jørgensen, and Toni Whited, as well as seminar participants at SITE, Michigan Ross, BYU, University of Luxembourg, University of Frankfurt, NYU Five Star, UC Berkeley Haas, Yale, University of Virginia, University of Chicago, Sciences Po, Adam Smith Conference, Stockholm School of Economics, University of Washington Foster, FIRS, Wharton Liquidity Conference, University of Rochester, Q Group Fall Seminar, Stanford, and the Virtual Finance Theory Seminar for helpful comments and discussions. We are especially thankful to Alexi Savov for extended conversations on the topic of this paper, as well as for sharing his code. Luke Min and Jack Kelly provided outstanding research assistance.

[†]Yale University and NBER. Email: eduardo.davila@yale.edu

[‡]New York University, Stern School of Business and NBER. Email: cparlato@stern.nyu.edu

1 Introduction

Financial markets play an important role by aggregating dispersed information about the fundamentals of the economy. By pooling different sources of information, asset prices act as a public signal to any external observer in the economy, potentially influencing individual decisions. This view, uncontested within economics and traced back to Hayek (1945), has faced significant challenges when translating its theoretical findings to more applied settings because measuring the informational content of prices is not an easy task.[1] In particular, one may be interested in understanding whether different markets aggregate dispersed information to different degrees. Or more specifically, one may want to find which stocks are more informative at a particular moment or whether prices have become more or less informative over time. These questions can only be answered if price informativeness can be correctly measured.

In this paper, we develop a methodology that allows us to identify and estimate exact stock-specific measures of price informativeness. To derive our results, we only need to postulate i) an asset pricing equation and ii) a stochastic process for asset payoffs. Our main result shows that a specific combination of outcomes (R-squareds) of linear regressions of changes in asset prices on changes in asset payoffs exactly identify price informativeness within a large class of models that may feature rich heterogeneity across investors — in terms of private signals, private trading needs, and preferences — and satisfy minimal distributional assumptions.

We begin by formally defining two price informativeness measures: absolute and relative price informativeness. Absolute price informativeness, which corresponds to the precision of the unbiased signal about the innovation to the asset payoff contained in the asset price, measures the precision of the public signal revealed by the asset price. Relative price informativeness, which corrects absolute price informativeness to account for the variability of the asset payoff, measures how much can be learned from the price relative to the total amount that can be learned. Relative price informativeness takes values between 0 and 1, which makes it easily interpretable and comparable across stocks. Moreover, in a Gaussian environment, relative price informativeness exactly corresponds to the Kalman gain in the updating process of a Bayesian external observer who only learns from the price. For instance, finding that relative price informativeness is 0.2 implies i) that the uncertainty faced by the external observer about the asset payoff is reduced by 20% after observing the price, and ii) that an external observer puts a weight of 20% in the price signal (and a weight of 80% in the prior) when forming a posterior belief over the future payoff.

We succinctly describe here our approach to identifying and estimating (relative) price informativeness. Consider estimating the following two regressions relating log-price changes, Δp_t, to the

[1] Hayek (1945) highlights the relevance of price informativeness as follows: "The economic problem of society is (...) rather a problem of how to secure the best use of resources known to any of the members of society, for ends whose relative importance only these individuals know. Or, to put it briefly, it is a problem of the utilization of knowledge which is not given to anyone in its totality."

contemporary and future differences in log-asset payoffs, respectively denoted by Δx_t and Δx_{t+1}:

$$\Delta p_t = \overline{\beta} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t+1} + e_t \tag{R1}$$

$$\Delta p_t = \overline{\zeta} + \zeta_0 \Delta x_t + e_t^{\zeta}, \tag{R2}$$

where we respectively denote the R-squareds of Regressions R1 and R2 by $R^2_{\Delta x, \Delta x'}$ and $R^2_{\Delta x}$. We show that the normalized difference in R-squareds

$$\frac{R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x}}$$

exactly corresponds to relative price informativeness. In addition to this identification result, we show that estimating these two regressions using ordinary least squares yields a consistent estimate of relative price informativeness. An important implication of our results is that it is possible to recover price informativeness by relying exclusively on price and payoff information, without having to observe the sources of noise in asset prices – subsumed in the error terms e_t and e_t^{ζ}. Our procedure to identify price informativeness is therefore agnostic about the sources of noise.

Our identification results do not require us to fully specify the model primitives. However, a fully microfounded model is necessary to understand the link between the primitives in the economy and price informativeness. For this reason, we explicitly develop several microfounded dynamic models of trading that map to the general asset pricing and payoff equations that we specify to present our identification results. First, we study a model in which investors have private signals about future payoffs and orthogonal trading motives in the form of random priors (sentiment). Next, we illustrate how to interpret price informativeness in two canonical models: i) a representative agent model similar to those used in the macro-finance literature, and ii) a model with informed and uniformed investors, as in the classic literature on information and learning. Finally, we present conditions on investors' asset demands that are sufficient to generate the assumed asset pricing equation. These applications show that our identification results apply to economies i) with or without dispersed information among investors, ii) with time-varying risk aversion, iii) in which investors may or may not learn from prices, and iv) in which noise may arise from different sources.

Even though price informativeness and price/return predictability may seem closely connected, these are conceptually different notions. We relate our results to the well-established literature in return predictability, which is based on running predictive regressions of future returns on current variables, which is the opposite from our approach. We explain that predictive regressions are the appropriate tool if one is interested in predicting future returns, but that our approach is the correct one if one wants to recover price informativeness. We also discuss in detail how our results relate to alternative measures of informativeness, like the posterior variance of a Bayesian external observer, or the forecasting price efficiency measure defined in Bond, Edmans and Goldstein (2012).

Before empirically implementing our results, we extend our methodology in several dimensions. First, we describe how to account for a non-zero correlation between payoff innovations and noise. Second, we allow for the possibility of public signals about future payoffs. Third, we augment the payoff process to have an unlearnable component. Finally, we adapt our methodology to identify informativeness about future prices, instead of future payoffs.[2] These extensions show that our results apply to more general and empirically relevant scenarios, allowing us also to refine the interpretation of our empirical findings.

Finally, we make use of our identification results to construct and analyze measures of stock-specific relative price informativeness. We recover a panel of stock-specific measures of price informativeness between 1980 and 2017 by running rolling time-series regressions of the form implied by Proposition 1 at the stock level using quarterly data. We find that the distribution of informativeness across stocks is right-skewed, with time-series averages of the median and mean levels of price informativeness across all stocks and years respectively given by 1.84% and 4.31%.

Our approach allows us to uncover both cross-sectional and time-series patterns regarding the behavior of price informativeness. In the cross-section, we find that that stocks that i) are larger, ii) turn over more quickly, and iii) have a higher institutional ownership share have higher price informativeness. In the time series, we find that the median and mean price informativeness have steadily increased over time since the mid-1980s. The standard deviation of price informativeness has also increased over this period. In the Appendix, we include additional results that show the robustness of our cross-sectional and time-series findings.

Our theoretical framework builds on the literature that studies the role played by financial markets in aggregating dispersed information, following Grossman and Stiglitz (1980), Hellwig (1980), Diamond and Verrecchia (1981), De Long et al. (1990), among others. Vives (2008) and Veldkamp (2011) provide thorough reviews of this well-developed and growing body of work. To our knowledge, we provide the first identification results of price informativeness within the literature of learning in financial markets.

While the role of financial markets aggregating information has been the subject of a substantial theoretical literature, the development of empirical measures of price informativeness is more recent. An earlier body of work proposes ad-hoc variables to study the informational content of prices. Influenced by the predictions of the CAPM/APT frameworks and following the prominent Roll (1988), Morck, Yeung and Yu (2000) study regressions of asset returns on a single or multiple factors and informally argue that the R^2 of such regressions can be used to capture whether asset prices are informative/predictive about firm-specific fundamentals. This ad-hoc measure, sometimes referred to as price nonsynchronicity, has been used in several empirical studies that link price informativeness to capital allocation. In particular, Wurgler (2000) finds that countries with higher price nonsynchronicity display a better allocation of capital. Durnev, Morck and Yeung (2004) document a positive correlation between price nonsynchronicity and corporate investment. Chen, Goldstein and Jiang (2006) establish that there exists a positive relation between the sensitivity of corporate investment to stock prices and two measures of the information contained in prices, price nonsynchronicity and the probability of informed trading (PIN), concluding that managers learn from the price when making corporate investment decisions. The PIN, developed

[2]In earlier versions of this paper, we also extended the results to environments with multiple risky assets and strategic investors.

in Easley, O'Hara and Paperman (1998), measures the probability of an informed trade using high frequency data through the lens of a model with noise and informed traders.

While some of the existing empirical work uncovers interesting empirical relations, Hou, Peng and Xiong (2013) forcefully highlight that a measure like Roll's R^2 lacks an structural interpretation. They question the link between return R^2 and price informativeness theoretically, in rational and behavioral settings, and empirically. In general, even if ad-hoc measures of price informativeness are correlated with true price informativeness, it is impossible to interpret the magnitude of these ad-hoc variables without a structural interpretation. By showing how to identify and recover exact stock-specific measures of price informativeness, we can reach precise quantitative conclusions.

More recently, Bai, Philippon and Savov (2015), have considered the question of whether financial markets have become more informative over time. Even though their empirical approach is motivated by a theoretical model, they do not provide identification results. There are several significant differences between our approach and theirs. First, they propose to measure the informational content of prices using forecasting price efficiency (FPE), a concept introduced in Bond, Edmans and Goldstein (2012). As we show in this paper, even though FPE may be the appropriate variable to compute social welfare in some environments, FPE does not separately identify price informativeness from payoff volatility. That is, FPE can be high because asset payoffs have low volatility or because asset prices are very informative. Second, they estimate FPE running cross-sectional regressions at specific points in time. This approach implicitly assumes that the data generating process (including the distribution of payoffs, signals, and noise) is the same for all stocks at a given point in time – an assumption that is easily falsifiable, as we show in this paper. We instead recover a panel of stock-specific measures of price informativeness by using rolling regressions. Finally, while they run predictive regressions of future fundamentals on current market values, we show that in order to recover consistent estimates of price informativeness, one must regress price changes (endogenous) on future payoffs (exogenous).

Our results and the recent work of Farboodi et al. (2019) and Kacperczyk, Sundaresan and Wang (forthcoming) complement each other. While our focus is to provide identification results for price informativeness (i.e., the signal-to-noise ratio in prices) in a general framework, Farboodi et al. (2019) seek to understand how changes in data processing over time have changed the amount of information (signal) incorporated in asset prices. Using a structural model of equilibrium asset prices, they find that the divergence in price informativeness across stocks is due to an increase in the amount of information incorporated in prices of large, high growth stocks driven by an increase in data processing capacity. Kacperczyk, Sundaresan and Wang (forthcoming) study the relation between price informativeness increase and the ownership share of foreign institutional investors, finding a positive relation.

As in any structural model, the measure of informativeness that we recover is linked to our assumptions on the behavior of investors and the market structure. While our framework is general along several dimensions, there is scope to think about how to identify price informativeness in alternative models of trading that depart from our linearity assumptions, like those of Albagli, Hellwig and Tsyvinski (2014, 2015, 2017). In particular, our analysis purposefully abstracts from feedback between prices and fundamentals, summarized in Bond, Edmans and Goldstein (2012) and tested in and Chen, Goldstein and Jiang (2006). Incorporating a two-way feedback between asset prices and payoffs introduces non-linearities that must be addressed using full-information methods.

Section 2 describes the general framework used to define price informativeness and presents our main results. Section 3 studies several microfounded models that are special cases of the general framework. Section 4 extends our results to more general environments. Section 5 empirically implements the methodology introduced in the paper, while Section 6 concludes. All proofs, derivations, and additional results are in the Appendix.

2 General Framework

In this section, we show how to formally identify and estimate price informativeness from an asset pricing equation and a stochastic process for asset payoffs. To stay as close as possible to the empirical implementation in Section 5, we derive the main results in the body of the paper in a log-difference-stationary environment, which is perceived to be a better representation of reality. In the Appendix, we re-derive the main results of the paper in a level-stationary environment, which is the benchmark environment in the literature on information and learning in financial markets (Vives, 2008; Veldkamp, 2011), and also in log-level-stationary and difference-stationary environments.

2.1 Environment

We consider a discrete time environment with dates $t = 0, 1, 2, \ldots, \infty$, in which investors trade a risky asset in fixed supply at a (log) price p_t at each date t. We assume that the (log) payoff of the risky asset at date $t+1$, x_{t+1}, follows a difference-stationary AR(1) process

$$\Delta x_{t+1} = \mu_{\Delta x} + \rho \Delta x_t + u_t, \tag{1}$$

where $\Delta x_t \equiv x_t - x_{t-1}$, $\mu_{\Delta x}$ is a scalar, $|\rho| < 1$, and where the innovations to the payoff difference, u_t, have mean zero, a finite variance denoted by $\mathbb{V}ar\left[u_t\right] = \sigma_u^2 = \tau_u^{-1}$, and are identically and independently distributed over time. Note that the innovation to the $t+1$ payoff difference, u_t, is indexed by t — instead of $t+1$ — because investors can potentially learn about the realization of u_t at date t.

We assume that the equilibrium (log) price difference is given by

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t, \tag{2}$$

where $\overline{\phi}$, ϕ_0, ϕ_1, and ϕ_n are parameters and where $\Delta n_t \equiv n_t - n_{t-1}$ represents the change in the aggregate component of investors' trading motives that are orthogonal to the asset payoff, given by $\Delta n_t = \mu_{\Delta n} + \varepsilon_t^{\Delta n}$, where $\mathbb{V}ar\left[\Delta n_t\right] = \sigma_{\Delta n}^2 = \tau_{\Delta n}^{-1}$. As shown in Section 3, the random variable n_t can be interpreted as a measure of investors' sentiment, risk-bearing capacity, or noise trading activity. Our timing assumes that date t variables, in particular Δx_t and u_t, are realized before the price p_t is determined. We assume that u_t and Δn_t are independent – this is without loss of generality, as we show in Section 4.

In Section 3, we show that Equation (2) emerges endogenously as the solution to a fully-specified dynamic model of trading. In that case, the parameters $\overline{\phi}$, ϕ_0, ϕ_1, and ϕ_n can be mapped to specific combinations of primitives. In Section 4, we extend our results to even more general environments.

2.2 Price Informativeness: Definition

From the perspective of understanding the informativeness of prices about future payoffs, the key variable of interest is the unbiased signal of the innovation to future payoffs u_t contained in the price. This endogenous signal, which we denote by π_t, is given by

$$\pi_t \equiv \frac{\Delta p_t - \left(\overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n} + (\phi_0 + \rho \phi_1) \Delta x_t\right)}{\phi_1}. \tag{3}$$

Given Equation (3), $\pi_t = u_t + \frac{\phi_n}{\phi_1} (\Delta n_t - \mu_{\Delta n})$ defines an endogenous unbiased signal about u_t, where $\frac{\phi_n}{\phi_1} (\Delta n_t - \mu_{\Delta n})$ acts as the noise contained in the price signal. This signal π_t is unbiased because $\mathbb{E}[\pi_t | u_t, \Delta x_t] = u_t$.

We formally define two related measures of informativeness: absolute and relative price informativeness. These are the relevant measures for an external observer who learns about the future payoff from the price.

Definition. (Price informativeness)
a) *Absolute price informativeness,* denoted by $\tau_\pi \in [0, \infty)$, is the precision of the unbiased signal about the innovation to the asset payoff contained in the asset price. Given Equation (2), it is formally given by

$$\tau_\pi \equiv (\mathbb{V}\text{ar}[\pi_t | x_{t+1}, \Delta x_t])^{-1} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_{\Delta n}, \tag{4}$$

where $\tau_{\Delta n} = \mathbb{V}\text{ar}[\Delta n_t]^{-1}$.
b) *Relative price informativeness*, denoted by $\tau_\pi^R \in [0, 1]$, is the ratio between absolute price informativeness and the sum of absolute price informativeness and the precision of the innovation to the asset payoff. Given Equation (2), it is formally given by

$$\tau_\pi^R \equiv \frac{\tau_\pi}{\tau_\pi + \tau_u}, \tag{5}$$

where $\tau_u = \mathbb{V}\text{ar}[u_t]^{-1}$.

The definition of *absolute price informativeness* connects with the large body of work that follows Blackwell (1953). According to Blackwell's informativeness criterion to rank experiments/signals, a signal is more informative than other when it is more valuable to a given decision-maker. According to that criterion, in the environment considered here, absolute price informativeness induces a complete order of price signals for a decision-maker with a quadratic objective around the value of the future asset payoff. This definition is widely used in the vast literature on information and learning in financial markets, see, e.g., Vives (2008) and Veldkamp (2011).

Intuitively, absolute price informativeness measures the signal-to-noise ratio of the signal contained in the asset price. If the price is very responsive to x_{t+1}, perhaps because investors trade with very precise information about the future payoff, ϕ_1 and price informativeness will be higher. Alternatively, if the price is mostly driven by trading motives that are orthogonal to future payoffs, perhaps reflecting investors' sentiment, $\phi_n^2 \tau_{\Delta n}^{-1}$ will be higher and price informativeness will be lower. When price informativeness is high, an external observer receives a very precise signal about future payoffs by observing the asset price p_t. On the contrary, when price informativeness is low, an external observer learns little about future payoffs by observing the asset price p_t.

The definition of *relative price informativeness* corrects absolute price informativeness to account for the variability of the fundamental, via τ_u. This measure captures the precision of the price signal, given by τ_π, relative to the sum of the prior and the signal precisions of an external observer, given by $\tau_\pi + \tau_u$. When uncertainty is Gaussian, relative price informativeness as defined in Equation (5) corresponds exactly to the Kalman gain of a Bayesian external observer who only learns from the price, as shown in Equation (7) below.

Relative price informativeness is an appealing object because it provides a bounded, unit-free measure of informativeness that facilitates making precise quantitative comparisons. The unit-free nature of this measure is particularly relevant when comparing informativeness across assets with different underlying payoff distributions (i.e., different τ_u), for which comparing absolute price informativeness is meaningless. In Remark 1 below, we further explain how absolute and relative price informativeness relate to other notions like posterior variances or forecasting price efficiency. In the body of the paper, we focus on the identification of relative price informativeness because it is easily interpretable and comparable across stocks. We include identification results for absolute price informativeness in the Appendix. Going forward, to simplify the exposition, we often refer to relative price informativeness simply as price informativeness.

2.3 Price Informativeness: Identification

Proposition 1 introduces the main result of the paper. It shows how to combine the R-squareds of regressions of prices on realized and future payoffs to recover price informativeness.

Proposition 1. (Identifying price informativeness) *Let $\overline{\beta}$, β_0, and β_1 denote the coefficients of the following regression of log-price differences on realized and future log-payoff differences,*

$$\Delta p_t = \overline{\beta} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t+1} + e_t, \tag{R1}$$

where $\Delta p_t = p_t - p_{t-1}$ denotes the date t change in log-price, $\Delta x_t = x_t - x_{t-1}$ and $\Delta x_{t+1} = x_{t+1} - x_t$ respectively denote the date t and $t+1$ log-payoff differences, and where $R^2_{\Delta x, \Delta x'}$ denotes the R-squared of Regression R1. Let $\overline{\zeta}$ and ζ_0 denote the coefficients of the following regression of log-price *differences on realized log-payoff differences,*

$$\Delta p_t = \overline{\zeta} + \zeta_0 \Delta x_t + e_t^{\zeta}, \tag{R2}$$

where $R^2_{\Delta x}$ denotes the R-squared of Regression R2. Then, relative price informativeness, τ_π^R, can be recovered as

$$\tau_\pi^R = \frac{R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x}}. \tag{6}$$

Estimating Regressions R1 and R2 via OLS yields consistent estimates of $R^2_{\Delta x, \Delta x'}$ and $R^2_{\Delta x}$.

The proof of Proposition 1 relies on identifying the right combination of parameters in the econometric specification defined by Regressions R1 and R2 that maps into the definition of relative price informativeness, τ_π^R. We show in the Appendix that a similar logic can be used to recover absolute price informativeness. It should be evident that if one could observe the non-payoff-related determinants of prices (n_t or Δn_t), that information could be used to directly recover all the relevant primitives in Equations (1) and (2). The non-trivial economic content of Proposition 1 is that if one is interested in recovering price informativeness, it is possible to do so by relying exclusively on price and payoff information, without having to observe n_t or Δn_t.

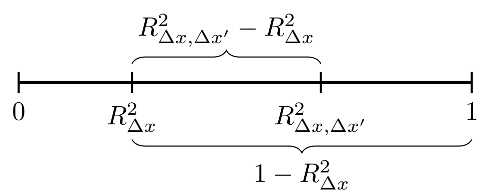

Figure 1: Interpreting relative price informativeness

Note: Relative price informativeness can be computed as the reduction in uncertainty, $R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}$, relative to the remaining residual uncertainty about future payoffs after conditioning on the realized date t payoff, $1 - R^2_{\Delta x}$.

Figure 1 illustrates how to interpret Equation (6). The denominator $1 - R^2_{\Delta x}$ can be interpreted as the remaining residual uncertainty about future payoffs after conditioning on the realized date t asset payoff. The numerator $R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}$ can be interpreted as the proportion of residual uncertainty about future payoffs that is revealed by observing the asset price at date t in addition to the realized payoff Δx_t. Because $R^2_{\Delta x, \Delta x'} \geq R^2_{\Delta x}$ and $R^2_{\Delta x, \Delta x'} \in [0, 1]$, it must be that $\tau_\pi^R \in [0, 1]$. As we show in the Appendix, if all random variables in the model are Gaussian, the posterior distribution over u_t of a Bayesian external observer who uses the asset price as a signal has the following normal distribution

$$u_t | \, p_t, \Delta x_t \sim N\left(\tau_\pi^R \pi_t, (\tau_\pi + \tau_u)^{-1}\right), \tag{7}$$

where π_t, τ_π, and τ_π^R are respectively defined in Equations (3), (4), and (5). Quantitatively, a relative price informativeness of, for instance, 0.15, implies that the initial uncertainty of an external observer about the innovation to the future payoff is reduced by 15% after learning from the price.

Even though we emphasize the economic identification of price informativeness, we also address how to recover consistent estimates. It is worth highlighting that estimating Regressions R1 and R2 through OLS (ordinary least-squares) yields consistent estimates of all the relevant parameters and R-squareds, since the error terms in both regressions are orthogonal to the regressors. Therefore, the estimates of price informativeness implied by Equations (3) and (4) will also be consistent.

We would like to conclude this section with three remarks.

Remark 1. *Alternative measures of informativeness.* The notion of price informativeness defined above can be related to other variables, in particular, i) the posterior variance of the future payoff conditional on the price, given by $\mathcal{V}_P \equiv \mathbb{V}\mathrm{ar}\left[u_t | \Delta x_t, \Delta p_t\right]$ and ii) forecasting price efficiency (FPE), given by $\mathcal{V}_{\mathrm{FPE}} \equiv \mathbb{V}\mathrm{ar}\left[\mathbb{E}\left[u_t | \Delta x_t, \Delta p_t\right]\right]$, as defined in Bond, Edmans and Goldstein (2012). Both variables are linked through the Law of Total Variance, as follows

$$\underbrace{\mathbb{V}\mathrm{ar}\left[u_t | \Delta x_t\right]}_{\tau_u^{-1}} = \mathbb{E}\left[\underbrace{\mathbb{V}\mathrm{ar}\left[u_t | \Delta p_t, \Delta x_t\right]}_{\mathcal{V}_P}\right] + \underbrace{\mathbb{V}\mathrm{ar}\left[\mathbb{E}\left[u_t | \Delta p_t, \Delta x_t\right]\right]}_{\mathcal{V}_{\mathrm{FPE}}}.$$

While $\mathcal{V}_P$ corresponds to the residual uncertainty about future payoffs after observing the price, $\mathcal{V}_{\mathrm{FPE}}$ measures the variation on the expectation of future payoffs after observing the price. When uncertainty is Gaussian, for a Bayesian external observer, both variables correspond to

$$\mathcal{V}_P = \frac{1}{\tau_\pi + \tau_u} \qquad \text{and} \qquad \mathcal{V}_{\mathrm{FPE}} = \frac{\tau_\pi}{\tau_\pi + \tau_u}\frac{1}{\tau_u}. \tag{8}$$

Equation (8) illustrates the challenge faced by both variables to identify price informativeness: they confound the effect of uncertainty about future payoffs (τ_u^{-1}) with price informativeness (τ_π). For instance, $\mathcal{V}_{\mathrm{FPE}}$ can be be high because the fundamental is easy to predict (high τ_u) or because asset prices are very informative about future payoffs (high τ_π). The same ambiguous inference applies to $\mathcal{V}_P$, which implies that neither of these measures are adequate to recover τ_π or τ_π^R.

Equation (8) also shows why $\left(\mathbb{V}\mathrm{ar}\left[\pi_t | u_t, \Delta x_t\right]\right)^{-1}$ is a desirable primitive notion of informativeness, since it can be derived without making assumptions on how an external observer updates. That is, finding $\mathbb{V}\mathrm{ar}\left[\pi_t | u_t, \Delta x_t\right]$ does not require to make distributional assumptions beyond the existence of second moments, while linking $\mathcal{V}_P$ or $\mathcal{V}_{FPE}$ to the precision of the information contained in prices requires making assumptions on distributions of priors, signals, and updating procedures.

Remark 2. *Informativeness vs. predictability.* Even though price informativeness and price/return predictability may seem closely connected, these are conceptually different notions. Given the assumptions made so far, Proposition 1 shows that running regressions of prices, which are endogenous, on future payoffs, which are exogenous, allows us to recover price informativeness. This entails running a regression of a date t variable, Δp_t, on a future explanatory variable, Δx_{t+1}, which contrasts with the well-established literature on return predictability (Cochrane, 2005; Campbell, 2017). One may wonder why not recover price informativeness using regressions of future payoffs on prices, since this type of regression can also be used for predictive purposes. This would imply reinterpreting Regression R1 as follows

$$\Delta x_{t+1} = \overline{\gamma} + \gamma_0 \Delta x_t + \gamma_1 \Delta p_t + \nu_t, \tag{R3}$$

where $\overline{\gamma} = -\frac{\overline{\beta}}{\beta_1}$, $\gamma_0 = -\frac{\beta_0}{\beta_1}$, $\gamma_1 = \frac{1}{\beta_1}$, and $\nu_t = -\frac{\beta_n}{\beta_1} n_t$. The main pitfall of this regression is that the OLS estimates of the coefficients and the residual variance will produce estimates of their structural counterparts that are not consistent as long as $\sigma_n^2 \neq 0$, since $\mathbb{C}\mathrm{ov}\left(\Delta p_t, \nu_t\right) = -\beta_n \frac{\beta_n}{\beta_1} \sigma_n^2 \neq 0$. In the Appendix, we illustrate how the OLS estimate of γ_1 in Regression R3 is downward biased and not consistent. We also describe the relation between our results and the literature on return predictability in detail.

To clarify, predictive regressions are the right tool if one is interested in forecasting future prices/returns using current fundamentals. As forcefully expressed by Cochrane (2005), the errors in predictive regressions are by construction orthogonal to the forecasts, so there is no scope for biases of lack of consistency in those cases. However, if one in interested in recovering price informativeness, which is a specific combination of structural variables, the approach developed in Proposition 1 is the adequate one.

Remark 3. *Payoff interpretation.* At the level of generality considered here, the payoff variable x_t could in principle represent any variable that satisfies Equations (1) and (2). That is, even though it may seem that, for instance, dividends are the most natural payoff measure, the results derived so far are agnostic about the exact nature of the payoff variable. We use this logic to justify the choice of earnings, instead of dividends, as the payoff measure in the empirical implementation of the results in Section 5.

3 Structural Models

We have shown in Section 2 that it is sufficient to specify an asset pricing equation and a stochastic process for asset payoffs to identify price informativeness. In this section, we explore several fully-specified environments that are consistent with Equations (1) and (2).

First, we study a model in which investors have private signals about future payoffs and orthogonal trading motives in the form of random priors (sentiment). Subsequently, we illustrate how to interpret price informativeness in two canonical models: i) a representative agent model similar to those used in the macro-finance literature, and ii) a model with informed and uniformed investors, as in the classic literature on information and learning. Finally, we present conditions on investors' asset demands that are sufficient to generate an asset pricing equation of the form assumed in Equation (2).

The applications below illustrate the generality of our results. In particular, they show that our identification results apply to economies i) with or without dispersed information among investors, ii) with time-varying risk aversion, iii) in which investors may or may not learn from prices, and iv) in which noise may arise from different sources. These applications also facilitate the interpretation of the empirical findings presented in Section 5 and highlight that our approach does not take a stance on the source of the aggregate noise.

3.1 Sentiment as Noise

We start by considering a model in which investors' sentiment is the source of noise in the price. Starting from primitives allows us to understand which assumptions on investors' behavior endogenously determine an equilibrium pricing equation of the form assumed in Section 2.

Environment We consider a tractable overlapping generations model. Time is discrete, with dates denoted by $t = 0, 1, 2, \ldots, \infty$. The economy is populated by a continuum of investors, indexed by $i \in I$, who live for two dates and are born with wealth w_t^i. We assume that the distribution of initial wealth is ergodic, bounded, and i.i.d. across investor types. An investor born at date t has well-behaved expected utility preferences over terminal wealth w_{t+1}^i, with flow utility given by $U_i\left(w_{t+1}^i\right)$, where $U_i'(\cdot) > 0$ and $U_i''(\cdot) < 0$.

There are two long-term assets in the economy: a risk-free asset in perfectly elastic supply, with gross return $R^f > 1$, and a risky asset in fixed supply Q, whose date t (log) payoff is $x_t = \ln(X_t)$ and which trades at a (log) price $p_t = \ln(P_t)$. The process followed by x_t is given by

$$\Delta x_{t+1} = \mu_{\Delta x} + u_t, \tag{9}$$

where $\Delta x_{t+1} = x_t - x_{t-1}$, $\mu_{\Delta x}$ is a scalar, and $x_0 = 0$. The realized payoff x_t is common knowledge to all investors before the price p_t is determined. The realized payoff at date $t+1$, x_{t+1}, is only only revealed to investors at date $t+1$. Note that Equation (9) is a special case of Equation (1) when $\rho = 0$. We focus on the $\rho = 0$ case only to simplify the exposition.

We assume that investors receive private signals about the innovation to the risky asset payoff. Formally, each investor receives a signal about the payoff innovation u_t given by

$$s_t^i = u_t + \varepsilon_{st}^i \quad \text{with} \quad \varepsilon_{st}^i \sim N\left(0, \tau_s^{-1}\right),$$

where $\varepsilon_{st}^i \perp \varepsilon_{st}^j$ for all $i \neq j$, and $u_t \perp \varepsilon_{st}^i$ for all t and all i.

We also assume that investors have additional private trading motives coming from random heterogeneous priors that are random in the aggregate. This is a particularly tractable formulation that sidesteps many of the issues associated with classic noise trading while still preventing full revelation of information – see Davila and Parlatore (2017) for a thorough analysis of this formulation, which can be seen as an extension of the classic DSSW model (De Long et al., 1990) that incorporates learning from prices. Formally, each investor i born at date t has a prior over the innovations to the payoff difference u_t given by

$$u_t \sim_i N\left(\overline{n}_t^i, \tau_u^{-1}\right),$$

where

$$\overline{n}_t^i = n_t + \varepsilon_{\overline{n}t}^i \quad \text{with} \quad \varepsilon_{\overline{n}t}^i \overset{\text{iid}}{\sim} N\left(0, \tau_{\overline{n}}^{-1}\right),$$

and

$$\Delta n_t = \mu_{\Delta n} + \varepsilon_t^n \quad \text{with} \quad \varepsilon_t^n \sim N\left(0, \tau_{\Delta n}^{-1}\right),$$

where $n_0 = 0$, $\mu_{\Delta n}$ is a scalar, and where $\varepsilon_t^n \perp \varepsilon_{\overline{n}t}^i$ for all t and all i. The variable n_t, which can be interpreted as the aggregate sentiment in the economy, is not observed and acts as a source of aggregate noise, preventing the asset price from being fully revealing.

Each investor i born at date t is endowed with wealth w_t^i and optimally chooses a portfolio share in the risky asset, denoted by θ_t^i, to solve

$$\max_{\theta_t^i} \mathbb{E}_t^i \left[U_i\left(w_{t+1}^i\right)\right] \tag{10}$$

subject to a wealth accumulation constraint

$$w_{t+1}^i = \left(R^f + \theta_t^i \left(\frac{X_{t+1} + P_{t+1}}{P_t} - R^f\right)\right) w_t^i, \tag{11}$$

where the information set of an investor i in period t is given by $\mathcal{I}_t^i = \left\{s_t^i, \overline{n}_t^i, \{x_s\}_{s \leq t}, \{p_s\}_{s \leq t}\right\}$.

Definition. (Equilibrium) A stationary rational expectations equilibrium in linear strategies is a set of portfolio shares θ_t^i for each investor i at date t and a price function P_t such that: i) θ_t^i maximizes the investor i's expected utility given his information set and ii) the price function P_t is such that the market for the risky asset clears at each date t, that is, $\int \theta_t^i w_t^i di = Q$.

In this class of models, it is well-known that it is not possible to characterize in closed-form the portfolio problem solved by investors and the equilibrium price – see e.g., Vives (2008). However, we show that it is possible to find a closed-form solution to the model in approximate form.

Equilibrium Characterization In the Appendix, we show that the risky asset demand of an investor i can be approximated as

$$\theta_t^i \approx \frac{1}{\gamma^i} \frac{k_0 + k_1 \mathbb{E}_t^i \left[p_{t+1} - x_{t+1}\right] + \mathbb{E}_t^i \left[\Delta x_{t+1}\right] - (p_t - x_t) - r^f}{\mathbb{V}ar_t^i \left[k_1 \left(p_{t+1} - x_{t+1}\right) + \Delta x_{t+1}\right]},$$

where $\gamma^i \equiv -\frac{w^i U''(w^i)}{U'(w^i)}$, and k_0 and k_1 are scalars, defined in in the Appendix.

As we show in the Appendix, taking a first order log-linear approximation of the first order condition, we have that the portfolio choice of investor i in period t can be approximated by

$$\theta_t^i \approx \alpha_x^i x_t + \alpha_s^i s_t^i + \alpha_n^i \overline{n}_t^i - \alpha_p^i p_t + \psi^i,$$

where the coefficients α_x^i, α_s^i, α_n^i, and α_p^i are positive scalars that represent the individual demand sensitivities to the contemporary payoff, the private signal, the private trading needs, and the asset price, respectively, and ψ^i can be positive or negative scalar that incorporates the risk premium. Using the market clearing condition with this approximation and the information structure described above yields a log-linear approximated price equal to

$$p_t = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} n_t + \frac{\overline{\psi}}{\overline{\alpha_p}},$$

where $\overline{\alpha_y} \equiv \int \alpha_y^i w^i di$ denotes the wealth weighted cross-sectional average of a given coefficient α_y^i and $\overline{\psi} = \int \psi^i w^i di - Q$. Using this expression, we can map the equilibrium price process in the model to the one assumed in the general framework.

First, we take a first order Taylor expansion of an investor's future marginal utility $U'\left(w_{t+1}^i\right)$ around the current date t wealth level w_t^i. Second, we impose that terms of order $(dt)^2$, that is, terms that involve the product of two or more net interest rates, are negligible. Third, as in Campbell and Shiller (1988), we take a log-linear approximation of returns around a predetermined dividend-price ratio. Finally, we assume that the joint distribution of demand sensitivities and risk aversion is time-invariant.

Lemma 1. *The price process assumed in Equation* (2) *in the general framework in Section 2 can be obtained endogenously as an approximation of the equilibrium price process in the model described in this section, i.e., the equilibrium price process is given by*

$$\Delta p_t \approx \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t,$$

where the coefficients $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, *and* $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$ *are determined in equilibrium.*

Lemma 1 and the payoff process assumed in this section imply that all the identification results we derived in the general framework in Section 2 can be applied in the context of the fully specified model derived in this section. This connection allows us to give a structural interpretation to the coefficients we recover from Regressions R1 and R2.

For example, the sensitivity of the price to the future payoff, ϕ_1, is given by ratio of the (wealth-weighted) averages of individual demand sensitivities to information and price $\frac{\overline{\alpha_s}}{\overline{\alpha_p}}$. Therefore, the more weight individual investors put on their private signals, the more sensitive the price will be to the future payoff and, everything else equal, the higher price informativeness (higher $R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}$). Analogously, when investors put more weight on their orthogonal trading motives, i.e., high $\overline{\alpha_n}$, the price will be more sensitive to the aggregate sentiment, and everything else equal, price informativeness will be lower (lower $R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}$).

3.2 Representative Agent

In this section, we show how to map the canonical representative agent model widely used in macro to the setting in Section 2. This application shows that our identification results do not rely on assuming dispersed information across investors and can accommodate time-varying risk aversion.

Environment Suppose there is one representative agent in the model introduced in the previous section 3.1. This is the same as having all investors $i \in I$ receive the exact same signal,

$$s_t^i = u_t + \varepsilon_{st} \quad \text{with} \quad \varepsilon_{st} \sim N\left(0, \tau_s^{-1}\right),$$

have the same prior, $u_t \sim_i N\left(\overline{n}_t, \tau_u^{-1}\right)$, where

$$\overline{n}_t = n_t + \varepsilon_{\overline{n}t} \quad \text{with} \quad \varepsilon_{\overline{n}t} \overset{\text{iid}}{\sim} N\left(0, \tau_{\overline{n}}^{-1}\right),$$

and have the same initial endowment wealth, $w_t^i = w_t$, and utility, $\gamma^i = \gamma$.

Equilibrium Characterization In this case, the log-linearly approximated price is equal to

$$p_t \approx \frac{\alpha_x}{\alpha_p} x_t + \frac{\alpha_s}{\alpha_p} s_t + \frac{\alpha_n}{\alpha_p} \overline{n}_t + \frac{\overline{\psi}}{\alpha_p}.$$

where the coefficients α_x, α_s, α_n, and α_p are demand sensitivity and $\overline{\psi}$ is a constant.

Since all investors receive the same signal s_t and have the same prior $\overline{n}_t$, there is no asymmetric information among investors in the model and, therefore, investors do not learn from the price. However, the price contains information about the innovation u_t for an external observer who only observes the price. The equilibrium price can be rewritten as

$$p_t \approx \frac{\alpha_x}{\alpha_p} x_t + \frac{\alpha_s}{\alpha_p} u_t + \frac{\alpha_s}{\alpha_p} \varepsilon_{st} + \frac{\alpha_n}{\alpha_p} \overline{n}_t + \frac{\psi}{\alpha_p}.$$

From the perspective of an external observer, there are two sources of noise that prevent the change in the price from being fully revealing. The noise in the signal ε_{st} and the investors' priors $\overline{n}_t$. It is easy to map the representative agent model into the framework we develop in Section 2, as the lemma below shows.

Lemma 2. *The price process assumed in Equation* (2) *in the general framework in Section 2 can be obtained endogenously as an approximation of the equilibrium price process in the model described in this section, i.e., the equilibrium price process is given by*

$$\Delta p_t \approx \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta \hat{n}_t,$$

where the coefficients $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\alpha_p} - \frac{\overline{\alpha_s}}{\alpha_p}$, $\phi_1 = \frac{\overline{\alpha_s}}{\alpha_p}$, *and* $\phi_n = \frac{\overline{\alpha_n}}{\alpha_p}$ *are equilibrium outcomes, and where* $\Delta \hat{n}_t \equiv \Delta \overline{n}_t + \frac{\alpha_s}{\alpha_n} \Delta \varepsilon_{st}$.

As in the previous section, Lemma 2 and the payoff process assumed allow us to apply all the identification results derived in Section 2 within the representative agent model. This shows that the price process in Equation (2) also encompasses models in which all investors share the same information and there is no learning form the price. In fact, our general framework does not require information to be dispersed in the economy and it can accommodate environments with and without learning.

3.3 Informed, Uninformed, and Noise Traders

Noise traders are a widely used modeling device in environments with dispersed information to avoid dealing with fully revealing equilibria. Our approach fully accommodates the presence of noise traders. This application highlights that our identification procedure accommodates different forms of noise, which allows us to remain agnostic about the source of noise in the economy.

Environment Suppose that we are in the same model developed in Section 3.1 with the only difference that there are three types of investors: informed, uninformed, and noise traders. Informed and uninformed investors share the same prior and only differ in the information they receive. Informed investors receive a perfectly informative signal of the innovation to the payoff. Uninformed investors and noise traders do not receive any signals. Mapping this to the model in Section 3.1 this implies that the prior distribution of the innovation u_t for informed and uninformed investors is

$$u_t \sim_i N\left(\overline{n}_t, \tau_u^{-1}\right),$$

where $\overline{n}_t \overset{\text{iid}}{\sim} N\left(0, \tau_{\overline{n}}^{-1}\right)$ and the precision of the signals for informed investors is $\tau_{si} = \infty$ and for uninformed investors is $\tau_{si} = 0$.

Finally, noise traders have private trading motives that are orthogonal to the innovation to the payoff that are the sole drivers of their demand. Formally, the demand of all noise traders in period t is random and given by $\delta_t \sim N\left(0, \tau_N^{-1}\right)$. The noise trader demand is only observed by noise traders.

Equilibrium Characterization In this case, the first order log-approximated price is

$$p_t \approx \frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \overline{n}_t + \frac{\overline{\psi}}{\overline{\alpha_p}} + \delta_t,$$

where $\overline{\alpha_y} \equiv \int_{I \cup U} \alpha_y^i w_t^i di$ denotes the wealth-weighted cross-sectional average of α_y^i over the set of informed and uninformed investors with $\alpha_s^i = 0$ for all uninformed investors, $\alpha_n^i = 0$ for all informed investors, and $\overline{\psi} \equiv \int_{I \cup U} \psi^i w_t^i di - Q$.

Lemma 3. *The price process assumed in Equation* (2) *in the general framework in Section 2, can be obtained endogenously as an approximation of the equilibrium price process in the model described in this section, i.e., the equilibrium price process is given by*

$$\Delta p_t \approx \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta \tilde{n}_t,$$

where the coefficients $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, *and* $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$ *are equilibrium outcomes and* $\Delta \tilde{n}_t \equiv \Delta \overline{n}_t + \frac{\overline{\alpha_p}}{\overline{\alpha_n}} \Delta \delta_t$.

Lemma 3 shows that all our identification results in Section 2 remain valid within the classic information model in Grossman and Stiglitz (1980) with inelastic noise traders. Within the model, only uninformed investors learn from the price and the only source of noise for them is the noise trader demand. However, for an external observer who only learns from the price there are two sources of noise embedded in the change in the price. The change in the noise trader demand $\Delta\delta_t$ and the change in the prior of the investors $\Delta\overline{n}_t$. Defining $\Delta\tilde{n}_t \equiv \Delta\overline{n}_t + \frac{\overline{\alpha_p}}{\overline{\alpha_n}}\Delta\delta_t$ allows us to clearly map this model into the framework we develop in the general framework in Section 2. This lemma together with the results in the previous two sections show that the price process assumed in our general framework can accommodate different sources of noise that prevent the price from being fully revealing for an external observer.

3.4 Asset Demand Model

Finally, we show that one can directly specify investors' asset demand and information structure, along with a payoff process, to identify price informativeness. This is an intermediate approach between using a fully-specified model, as in Sections 3.1 to (3.3), and directly postulating an equilibrium pricing equation, as in Section 2.

Consider a discrete time environment with a continuum of investors, indexed by $i \in I$, who trade a risky asset in fixed supply at a (log) price p_t each date $t = 0, 1, \ldots, \infty$. Assume that the (log) payoff of the risky asset in period $t+1$, x_{t+1}, is given by the following stationary AR(1) process in differences

$$\Delta x_{t+1} = \mu_{\Delta x} + \rho\Delta x_t + u_t,$$

where $\mu_{\Delta x}$ is a scalar, $|\rho| < 1$, and where the innovations to the payoff, u_t, have mean zero, finite variance, and are independently distributed. Investors trade in period t with imperfect information about the innovation to the payoff, u_t, which is realized at the end of the period. When trading in period t, the contemporaneous payoff u_t has already been realized and is common knowledge to all investors.

Each period t, an investor i observes a private signal s_t^i of the innovation to the payoff u_t.[3] Investors have an additional motive for trading the risky asset that is orthogonal to the asset payoff. We denote by $\overline{n}_t^i$ investor i's additional trading motive in period t. These additional trading motives are private information of each investor and are random in the aggregate.

We rederive the main result of the paper under two assumptions. The first assumption imposes an additive informational structure and guarantees the existence of second moments, while the second assumption imposes a linear structure for investors' equilibrium asset demands. In general, linear demands can be interpreted as a first-order approximation to other forms of asset demands, so one may expect our results to approximately hold in a larger class of models. Both assumptions facilitate the aggregation of individual demands in order to yield a linear equilibrium pricing function.

[3] Assuming that investors observe private signals about the payoff, x_{t+1}, or its innovation, u_t, is formally equivalent, since x_t is known to investors when trading in period t.

Assumption 1. (Additive noise) *Each period t, every investor i receives an unbiased private signal s_t^i about the innovation to the payoff, u_t, of the form*

$$s_t^i = u_t + \varepsilon_{st}^i, \tag{12}$$

where ε_{st}^i, $\forall i \in I$, $\forall t$, are random variables with mean zero and finite variances, whose realizations are independent across investors and over time. Each period t, every investor i has a private trading need $\overline{n}_t^i$, of the form

$$\overline{n}_t^i = n_t + \varepsilon_{\overline{n}t}^i, \tag{13}$$

where Δn_t is a random variable with finite mean, denoted by $\mu_{\Delta n}$, and finite variance, and where $\varepsilon_{\overline{n}t}^i$, $\forall i \in I$, $\forall t$, are random variables with mean zero and finite variances, whose realizations are independent across investors and over time.

Assumption 1 imposes restrictions on the noise structure in the signals about the innovation to the fundamental u_t and on all other sources of investors' private trading needs by making them additive and independent across investors. This assumption does not restrict the distribution of any random variable beyond the existence of finite first and second moments. Our second assumption describes the structure of the investors' demands for the risky asset θ_t^i.

Assumption 2. (Linear asset demands) *Investors' asset demands satisfy*

$$\theta_t^i = \alpha_s^i s_t^i + \alpha_x^i x_t + \alpha_n^i \overline{n}_t^i - \alpha_p^i p_t + \psi^i,$$

where α_s^i, α_x^i, α_n^i, α_p^i, and ψ^i are individual demand coefficients, determined in equilibrium.

Assumption 2 imposes a linear structure on the individual investors' net asset demand for the risky asset. More specifically, that an individual investor's demand is linear in his signal about the fundamental and his private trading needs, as well as in the asset price p_t and the current realization of the fundamental x_t. It also allows for an individual specific invariant component ψ^i. All the models explored in Section (3) are consistent with Assumptions 1 and 2.

Lemma 4. *The price process assumed in Equation* (2) *in the general framework in Section 2 can be obtained endogenously when Assumptions 1 and 2 are satisfied.*

Note that these assumptions allows for a rich cross-sectional heterogeneity among investors. In particular, it accommodates heterogeneity in investors' risk aversion, in the precision of their information, and in the distribution of their idiosyncratic trading motives. As the applications above show, our assumptions can accommodate models with informed and uninformed traders, which can be mapped to environments in which one set of agents does not observe any private signal, and those with classic noise traders, which can be mapped to environments in which one set of agents trades fixed amounts regardless of the price or other features of the environment. Given that linear asset demands can be interpreted as an approximation to more general models, Lemma 4 implies that our results should be valid more broadly in an approximate sense.

4 Extensions

Before empirically implementing our results, we extend the results derived in Section 2 in several dimensions. We describe i) how to account for a non-zero correlation between u_t and Δn_t, ii) how to allow for the possibility of public signals about future payoffs, iii) how to implement our results when the payoff process has an unlearnable component, and iv) how to adapt our methodology to identify informativeness about future prices.[4] This section has two goals. First, it shows that our results apply to more general and empirically relevant scenarios. Second, it allows us to refine the interpretation of our empirical findings.

4.1 Correlated Payoff and Noise

In Section 2, we assume that the innovation to asset payoff u_t is uncorrelated with the aggregate source of noise Δn_t. In this section, we allow for the aggregate source of noise to be correlated with the payoff. Formally, we consider the following process for the aggregate noise

$$\Delta n_t = \mu_{\Delta n} + \omega u_t + \varepsilon_t^n,$$

with $\omega \neq 0$, where ε_t^n has mean zero and finite variance and is i.i.d. across time and independent of the innovations u_t. In this case, the equilibrium price process

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t$$

can be written as

$$\Delta p_t = \overline{\phi} - \phi_n \omega \mu_{\Delta x} + \phi_n \mu_{\Delta n} + (\phi_0 - \phi_n \omega \rho) \Delta x_t + (\phi_1 + \phi_n \omega \rho) \Delta x_{t+1} + \phi_n \varepsilon_t^n. \tag{14}$$

and absolute price informativeness is given by

$$\tau_\pi \equiv \mathbb{V}ar\left[\pi_t | u_t, x_t\right]^{-1} = \left(\frac{\phi_1 + \phi_n \omega}{\phi_n}\right)^2 \tau_{\Delta n}. \tag{15}$$

The main difference between this extension and the baseline framework is that the sensitivity of the price to the future payoff, which determines how much information is impounded in the price, has an additional component that comes from the comovement between the payoff and noise, $\phi_n \omega$, as it can be seen from Eqs. (14) and (15). However, this sensitivity can still be recovered from the coefficient on Δx_{t+1} in Regression R1 and, as the proposition below shows, all identification results derived in the general framework remain valid.

Proposition 2. (Identifying price informativeness when payoff and noise are correlated) *The identification results in Proposition 1 remain valid when the future payoff x_{t+1} is correlated*

[4] In earlier versions of this paper, we also extended the results to environments with multiple risky assets and strategic investors.

with the aggregate noise n_t, i.e., relative price informativeness can be recovered from the R-squared of Regressions R1 and R2 as follows

$$\tau_\pi^R = \frac{R^2_{\Delta x,\Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x}}.$$

As the proposition above shows, even though the structural characterization of price informativeness is different when the payoff is correlated with the aggregate noise, price informativeness can be recovered in the same way as in the baseline general framework. Therefore, we can assume that the private trading motives are orthogonal to the asset payoff without loss of generality.

4.2 Public Signals

In our results until now, we have considered private signals as the only source of information in the economy. In this subsection, we consider the case in which investors also observe a public signal about the fundamental. We extend the environment in the general framework in Section 2 by considering an environment in which investors to observe a vector of N public signals

$$\chi_t = \omega u_t + \overline{\varepsilon}_t^{\chi},$$

where $\overline{\varepsilon}_t^{\chi}$ has mean zero and finite variance and is i.i.d. across time and independent of the innovations u_t. In this case, we augment the price process in Eq. (3) to include the private signals χ_t as follows

$$\Delta p_t = \overline{\phi} + \phi_{-1}\Delta x_{t-1} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_\chi \Delta \chi_t + \phi_n \Delta n_t. \tag{16}$$

There are two relevant notions of price informativeness that depend on the information available to the external observer. If the public signals are not available to the external observer, price informativeness is given by $\tau_\pi \equiv \mathbb{V}ar\left[\pi_t' \middle| u_t, x_t\right]^{-1}$. If public signals are observed by the external observer, price informativeness is given by $\tau_{\pi'} \equiv \mathbb{V}ar\left[\pi_t' \middle| u_t, x_t, \chi_t\right]^{-1}$. The proposition below provides identification results for each of these two cases.

Proposition 3. (Identifying price informativeness with public signals)

a) When the public signals available to investors are not available to the external observer, the identification results in Proposition 1 remain valid, i.e., relative price informativeness can be recovered from the R-squared of Regressions R1 and R2 as follows

$$\tau_\pi^R = \frac{R^2_{\Delta x,\Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x}}.$$

b) When the public signals available to investors are part of the information set of the external observer, relative price informativeness can be recovered from regressions of prices on payoffs and *the public signals*

$$\Delta p_t = \overline{\beta}' + \beta_0' \Delta x_t + \beta_1' \Delta x_{t+1} + \beta_2' \Delta \chi_t + e_t', \tag{R1–PS}$$

$$\Delta p_t = \overline{\zeta}' + \zeta_0' \Delta x_t + \zeta_2' \Delta \chi_t + e_t^{\zeta\prime}, \tag{R2–PS}$$

as

$$\tau_{\pi'}^R = \frac{R^2_{\Delta x, \Delta x', \Delta \chi} - R^2_{\Delta x, \Delta \chi}}{1 - R^2_{\Delta x, \Delta \chi}},$$

where $R^2_{\Delta x, \Delta x', \Delta \chi}$ *and* $R^2_{\Delta x, \Delta \chi}$ *are the R-squareds of Regression R1–PS and Regression R2–PS, respectively.*

Part a) of the proposition above follows directly from reinterpreting the unobserved terms in the price equation (16), i.e., $\phi_\chi \Delta \chi_t + \phi_n \Delta n_t$, as noise that is correlated with the payoff and using Proposition 2. This shows that our results are robust to investors having public information when this information is not available to the external observer.

Part b) shows how to extend our identification results when the public signals are part of the external observer's information set. The proof is analogous to the proof of Proposition 1 in the general framework, with the coefficients in Regressions R1–PS and R2–PS being matched to the expressions of the price process in Eqs. (16). A detailed proof of Proposition 3 can be found in the Appendix.

Intuitively, to capture the additional information contained in the price, the Regressions R1–PS and R2–PS need to condition for the information set of the external observer. In the general framework in Section 2 we assume that the external observer only observes the price and the contemporary payoff. If the information set of the external observer contains other information, then Proposition 3 shows that our identification results remain valid provided the regressions of prices on payoffs are augmented to include the information available to the external observer.

One can draw an analogy between the standard notions of market efficiency (weak, semi-strong, and strong) and the set of controls used in our regressions. For instance, regressions that include exclusively past payoffs as controls resemble weak-form efficiency notions. By expanding the set of controls to include public or private information, the recovered informativeness measures resemble semi-strong or strong notions of efficiency.

4.3 Learnable and Unlearnable Payoff

In our derivations so far, we have considered that all components of the payoff are learnable, that is, that there is no systematic component of the payoff that deviates from the signals received by the investors. However, it is plausible to think that investors can only learn about a part of the innovation and that the remainder is unlearnable. Formally, we assume that the innovation to the payoff is given by

$$\Delta x_{t+1} = \mu_{\Delta x} + u_t,$$

where $\Delta x_{t+1} = x_t - x_{t-1}$, μ_x is a scalar, and $x_0 = 0$. Moreover, the innovation to the payoff is given by

$$u_t = u_t^L + u_t^U,$$

where u_t^L and u_t^U are the learnable and unlearnable components of the innovation where

$$u_t^L \sim N\left(0, \left(\tau_u^L\right)^{-1}\right) \quad \text{and} \quad u_t^U \sim N\left(0, \left(\tau_u^U\right)^{-1}\right)$$

with $u_t^L \perp u_t^U$. The main difference between these two components is that investors only receive private signals about the learnable component, u_t^L. Formally, each investor receives a signal

$$s_t^i = u_t^L + \varepsilon_{st}^i \quad \text{with} \quad \varepsilon_{st}^i \sim N\left(0, \tau_s^{-1}\right),$$

where $\varepsilon_{st}^i \perp \varepsilon_{st}^j$ for all $i \neq j$, $u_t^L \perp \varepsilon_{st}^i$, and $u_t^U \perp \varepsilon_{st}^i$ for all t and all i.

In this case, the price process is given by

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t - \phi_1 \Delta u_t^U, \tag{17}$$

where $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$ and $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$. Using the process for the payoff, this price process can also be written as

$$\Delta p_t = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_0 \Delta x_t + \phi_1 u_t^L + \phi_1 u_{t-1}^U + \phi_n \Delta n_t.$$

Then, the unbiased signal about the change in the learnable component of the innovation contained in the price is

$$\pi_t^L \equiv \frac{1}{\phi_1}\left(\Delta p_t - \left(\phi + \phi_1 \mu_x + \phi_n \mu_{\Delta n} + \phi_0 \Delta x_t\right)\right) = u_t^L + u_{t-1}^U + \frac{\phi_1}{\phi_n} \Delta \varepsilon_t^n.$$

and absolute and relative price informativeness are respectively given by

$$\tau_\pi^L \equiv \mathbb{V}ar\left[u_{t-1}^U + \frac{\phi_n}{\phi_1}\Delta n_t \middle| u_t^L, \Delta x_t\right]^{-1} = \left(\left(\tau_u^L + \tau_u^U\right)^{-1} + \left(\frac{\phi_n}{\phi_1}\right)^2 \tau_{\Delta n}^{-1}\right)^{-1}$$

and

$$\tau_\pi^{LR} = \left(\frac{\tau_u^L}{\tau_u^L + \tau_u^U} + \left(\frac{\phi_n}{\phi_1}\right)^2 \left(\frac{\tau_{\Delta n}}{\tau_u^L}\right)^{-1}\right)^{-1}.$$

Note that the noise in the signal π_t^L is not independent of Δx_t since $\Delta x_t = \mu_{\Delta x} + u_{t-1}^L + u_{t-1}^L$. Then, conditioning on Δx_t becomes relevant in our definition of price informativeness because Δx_t contains information about the noise in the signal.

When there is an unlearnable component of the payoff, the error term in regression R1 is given by $e_t^{\Delta x, \Delta x'} = \phi_n \varepsilon_t^n - \phi_1 \Delta u_t^U$ and, therefore, is correlated with the regressors Δx_{t+1} and Δx_t. Moreover, the error term in Regression R2 is given by $e_t^{\Delta x} = \phi_n \varepsilon_t^n + \phi_1 u_t^L + \phi_1 u_{t-1}^U$, which is correlated with Δx_t

through u_{t-1}^{U}. Hence, to obtain consistent estimates of the regression parameters, we need to run Regressions R1 and R2 instrumenting for Δx_{t+1} and Δx_t with a variable that is correlated with the learnable component of the innovation but not with the unlearnable one. However, while using an instrument for Δx_{t+1} and Δx_t provides consistent estimates, the identification procedure needs to be adjusted to take into account the extra term in the expression for absolute price informativeness and in the error term e_t.[5] The following proposition characterizes the new identification result when there is an unlearnable component of the innovation to the payoff.

Proposition 4. (Identifying price informativeness about learnable component of payoffs) *When there is an unlearnable component of the payoff, relative price informativeness about the learnable component of the innovation can be recovered as follows*

$$\tau_{\pi}^{LR} = \frac{\mathbb{V}ar\left(e_t^{\Delta x}\right) - \mathbb{V}ar\left(e_t^{\Delta x, \Delta x'}\right)}{\mathbb{V}ar\left(e_t^{\Delta x}\middle|\, \Delta x_t\right)} \frac{1}{1 - \frac{\mathbb{V}ar\left(u_t^U\right)}{\mathbb{V}ar\left(u_t^L\right)}}, \tag{18}$$

where

$$\frac{\mathbb{V}ar\left(u_t^U\right)}{\mathbb{V}ar\left(u_t^L\right)} = \frac{1 - \frac{\mathbb{V}ar\left(e_t^{\Delta x}\right) - \mathbb{V}ar\left(e_t^{\Delta x, \Delta x'}\right)}{\phi_1^2 \mathbb{V}ar(u_t)}}{1 + \frac{\mathbb{V}ar\left(e_t^{\Delta x}\right) - \mathbb{V}ar\left(e_t^{\Delta x, \Delta x'}\right)}{\phi_1^2 \mathbb{V}ar(u_t)}} \tag{19}$$

and $\mathbb{V}ar\left(e_t^{\Delta x}\right)$ and $\mathbb{V}ar\left(e_t^{\Delta x, \Delta x'}\right)$ are the variances of the residuals in Regressions R2 and R1, respectively, $\mathbb{V}ar\left(e_t^{\Delta x}\middle|\, \Delta x_t\right)$ is the variance of the residual in a regression of the residuals in Regression R2 on Δx_t, and $\mathbb{V}ar(u_t)$ is variance of the residuals in a regression of Δx_{t+1} on Δx_t. Estimating Regressions R2 and R1 consistently requires instrumenting for Δx_{t+1} and Δx_t with a variable that is correlated with the learnable component of the innovation but not with the unlearnable one.

Proposition 4 shows how to identify relative price informativeness when there is an unlearnable component of the payoff. There are two terms that help identify price informativeness in this case. The first term is in Equation (18) is a ratio of residual variances, which is analogous to the ratio of normalized difference in R-squareds in our previous identification results, once adjusted for the correlation between Δx_t and the error term $e_t^{\Delta x}$. In the limit, when the variance of the unlearnable component is 0, as expected, we recover the result from Proposition 1, since

$$\lim_{\mathbb{V}\mathrm{ar}\left(u_t^U\right) \to 0} \frac{\mathbb{V}\mathrm{ar}\left(e_t^{\Delta x}\right) - \mathbb{V}\mathrm{ar}\left(e_t^{\Delta x, \Delta x'}\right)}{\mathbb{V}\mathrm{ar}\left(e_t^{\Delta x}\middle|\, \Delta x_t\right)} = \frac{R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x}}.$$

The second term in Equation (18) is a correction term that takes into account the presence of an unlearnable component in the innovation to the payoff. When there is no unlearnable component

[5] A good candidate for an instrument is the earnings forecasts. These forecasts are determined by the information available to analysts about the future payoff and, therefore, are only correlated with their learnable component.

in the innovation, i.e., when $\mathbb{V}\text{ar}\left(u_t^U\right) \to 0$, the correction term is one, which brings us back to our identification results in our benchmark model. On the other hand, when the unlearnable component is the one driving all the change in the payoff, i.e., when $\mathbb{V}\text{ar}\left(u_t^U\right) \to \infty$, the correction term goes to zero and there is no information about the payoff in prices.

Summing up, if payoffs have an unlearnable component, the identification and estimation of price informativeness should be modified. Regarding the identification, Proposition 4 shows how it is necessary to correct the expression from our baseline model. On the estimation side, both regressions require an instrument correlated with the learnable component of the innovation but not with the unlearnable one. Instead of looking for stock-specific instruments, we present our empirical results using Proposition 1, and use the results of this section to explore the potential bias via simulations. In the Appendix, after simulating this model, we show that the estimates of price informativeness obtained using the procedure from Proposition 1 are downward biased. As we describe there, the bias disappears when $\mathbb{V}\text{ar}\left(u_t^U\right) \to 0$, as expected.

4.4 Learning about Future Prices

In our analysis so far, we have looked at the information about the next period's payoff that is contained in the price. However, there are other measures that may be on interest to investors when making decisions. For example, an external observer may want to learn about next period's price from the current price. In this subsection, we adapt our analysis to price informativeness about the next period's price.

We assume that the price process at date t is given by

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t.$$

Iterating this expression forward and using the process for the change in payoff we can express the current price difference in terms of the future price difference, as follows

$$\Delta p_t = \overline{\phi}^p + \phi_0 \Delta x_t + \phi_1^p \Delta p_{t+1} + e_t^p, \tag{20}$$

where $\overline{\phi}^p \equiv \overline{\phi} - \frac{\phi_1}{\phi_0+\rho\phi_1}\left(\overline{\phi} + \phi_1 \mu_x\right)$, $\phi_1^p \equiv \frac{\phi_1}{\phi_0+\rho\phi_1}$, and $e_t^p \equiv \phi_n \Delta n_t - \frac{\phi_1}{\phi_0+\rho\phi_1}\left(\phi_1 u_{t+1} + \phi_n \Delta n_{t+1}\right)$. Hence, the unbiased signal contained in the price about the next period's price is

$$\tilde{\pi}_t = \frac{\Delta p_t - \left(\overline{\phi}^p + \phi_0 \Delta x_t + \mathbb{E}_t\left[e_t^p\right]\right)}{\phi_1^p} = \Delta p_{t+1} + \frac{1}{\phi_1^p} e_t^p.$$

Then, absolute and relative price informativeness are respectively given by

$$\tau_{\hat{\pi}} = \mathbb{V}ar\left[\hat{\pi}_t \middle| x_t, \Delta p_{t+1}\right]^{-1} = \mathbb{V}ar\left[\frac{1}{\phi_1^p} e_t^p \middle| x_t, \Delta p_{t+1}\right]^{-1} \quad \text{and} \quad \tau_{\hat{\pi}}^R \equiv \frac{\tau_{\hat{\pi}}}{\tau_{\hat{\pi}} + \tau_{\Delta p}}.$$

From Eq. (20) and the definition of e_t^p it follows that we cannot recover price informativeness from regressions of prices on future prices, since the error term of that regression would not be orthogonal to the future price and, hence, the estimated coefficients would be biased estimates of the structural parameters. However, using the structural mapping between the equilibrium price process and the coefficients in Regressions R1 and R2 we can recover price informativeness about the next period's price as the proposition below shows.

Proposition 5. (Identifying price informativeness about future price) *Relative price informativeness about the future price can be recovered from Regressions R1 and R2 as follows*

$$\tau_{\hat{\pi}}^{R} = \frac{1}{1 + \left(1 - R_{\Delta x}^{2} + \left(\frac{\zeta_0}{\beta_1}\right)^{2}\left(1 - R_{\Delta x, \Delta x'}^{2}\right)\right)}. \tag{21}$$

While we focus our empirical implementation on the behavior of price informativeness about future payoffs, there is scope to study further the behavior of price informativeness about future prices.

5 Empirical Implementation: Stock-Specific Price Informativeness

In this section, we make use of our identification results to construct and analyze measures of stock-specific relative price informativeness. We exclusively focus on relative price informativeness because it allows us to come up with meaningful and easily interpretable comparisons across stocks and over time. We recover a panel of stock-specific measures of price informativeness by running rolling time-series regressions of the form implied by Proposition 1 at the stock level using quarterly data.

We find that the distribution of informativeness across stocks is right-skewed, with time-series averages of the median and mean levels of price informativeness across all stocks and years respectively given by 1.84% and 4.31%. Our approach allows us to uncover both cross-sectional and time-series patterns regarding the behavior of price informativeness. In the cross-section, we find that that stocks that i) are larger, ii) turn over more quickly, and iii) have a higher institutional ownership share have higher price informativeness. In the time series, we find that the median and mean price informativeness have steadily increased over time since the mid-1980s. The standard deviation of price informativeness has also increased over this period. In the Appendix, we include additional results that show the robustness of our cross-sectional and time-series findings.

5.1 Data Description and Empirical Specification

We initially provide a brief description of the data and the sample selection procedure. The Appendix and the companion R notebooks include a more detailed description.[6] We obtain information on stock prices and accounting measures from the CRSP/Compustat dataset, as distributed by WRDS. Our sample selection procedure follows the conventional approach described in Bali, Engle

[6] See https://github.com/edavila/identifying_price_informativeness.

and Murray (2016). From the the Center for Research in Security Prices (CRSP), we obtain data on stock prices, market capitalization, turnover, S&P500 status, and industry (SIC) classification for all common US-based stocks listed on the NYSE, NASDAQ, and AMEX. From Compustat, we obtain accounting data that includes earnings and book values, at both quarterly and annual frequencies. In this section, we report the results of the analysis using quarterly data, which runs from 1961 until 2017. In the Appendix, we report the results of the analysis using annual data, which runs from 1950 until 2017. To match the timing of our model and to ensure that the accounting data were public on the trading date, we merge the Compustat data with CRSP data three-months ahead, although our findings are robust to using alternative windows. We use the personal consumption expenditure index (PCEPI), obtained from FRED, to deflate all nominal variables.

We implement Proposition 1 by running time-series regressions for each individual stock — indexed by j here — over rolling windows of 30 quarters (7.5 years).[7] We denote by p_t^j the log price of stock j, adjusted for splits. We use earnings — as measured by EBIT – as the relevant measure of payoffs, since stock-level measures of dividends are problematic. As discussed in Section 2, our model can be flexibly interpreted to use earnings as payoff measure. Since earnings can be negative, we compute Δx_t^j directly as a growth rate, as explained in the Appendix – we obtain comparable results when we compute Δx_t^j as the log-difference of the logistic transformation of standardized earnings. Formally, in a given rolling window, we run time-series regressions of the form

$$\Delta p_t^j = \overline{\beta}^j + \beta_0^j \Delta x_t^j + \beta_1^j \Delta x_{t+1}^j + d_t^{j,q} + \varepsilon_t^j \quad \Rightarrow R^{2,j}_{\Delta x, \Delta x'} \tag{22}$$

$$\Delta p_t^j = \overline{\zeta}^j + \zeta_0^j \Delta x_t^j + d_t^{j,q} + \hat{\varepsilon}_t^j \quad \Rightarrow R^{2,j}_{\Delta x}\ , \tag{23}$$

where Δp_t^j is a measure of capital gains, Δx_t^j and its one period ahead counterpart Δx_{t+1}^j are measures of earnings growth, and $d_t^{j,q}$ denote stock specific quarterly dummies. The introduction of $d_t^{j,q}$ accounts for seasonality patterns, and can be interpreted along the lines of Section 4.2. We estimate the regressions coefficients and errors using OLS. We respectively denote the R-squareds of the regressions (22) and (23) by $R^{2,j}_{\Delta x, \Delta x'}$ and the $R^{2,j}_{\Delta x}$. Hence, Regression R1 maps to Equation (22), while Regression R2 maps to Equation (23), but for the addition of the quarterly dummies, interpreted as public signals.

Consistently with Proposition 1, we recover relative price informativeness for stock j in a given period from Equations (22) and (23) as follows

$$\tau_\pi^{R,j} = \frac{R^{2,j}_{\Delta x, \Delta x'} - R^{2,j}_{\Delta x}}{1 - R^{2,j}_{\Delta x}}. \tag{24}$$

After restricting our results to stocks with contiguous observations and whose maximum leverage across observations is lower than 0.95, we end up with a panel of quarterly price informativeness measures for $2,959$ unique stocks.

[7] By working with the model in log-differences, we sidestep concerns associated with failures of stationarity — see e.g., Campbell (2017).

Table 1: Price informativeness: year-by-year summary statistics

t	$Median$	$Mean$	SD	$Skew$	$Kurt$	$P5$	$P25$	$P75$	$P95$	n
1980	0.0177	0.0384	0.0507	2.3098	6.5890	0.0006	0.0048	0.0532	0.1426	287
1981	0.0188	0.0383	0.0510	2.7435	11.3186	0.0005	0.0057	0.0526	0.1495	385
1982	0.0180	0.0385	0.0524	2.5854	9.0303	0.0004	0.0044	0.0533	0.1474	638
1983	0.0162	0.0358	0.0506	2.8678	11.9352	0.0005	0.0043	0.0485	0.1359	916
1984	0.0167	0.0354	0.0492	2.5405	7.9869	0.0005	0.0047	0.0443	0.1413	1303
1985	0.0160	0.0337	0.0465	3.0343	13.1555	0.0006	0.0049	0.0457	0.1230	1403
1986	0.0160	0.0341	0.0488	3.2878	16.4866	0.0004	0.0048	0.0431	0.1277	1363
1987	0.0168	0.0344	0.0465	2.7371	10.5861	0.0006	0.0047	0.0456	0.1250	1408
1988	0.0153	0.0342	0.0500	3.1442	14.0628	0.0003	0.0044	0.0430	0.1284	1350
1989	0.0159	0.0332	0.0465	3.0760	14.4146	0.0004	0.0044	0.0426	0.1219	1299
1990	0.0152	0.0329	0.0459	2.9421	12.9735	0.0006	0.0047	0.0428	0.1235	1281
1991	0.0141	0.0331	0.0487	3.0087	12.7699	0.0004	0.0035	0.0416	0.1288	1818
1992	0.0148	0.0330	0.0484	3.1930	15.0297	0.0003	0.0040	0.0406	0.1236	1997
1993	0.0151	0.0338	0.0490	2.7873	10.4996	0.0003	0.0037	0.0416	0.1338	2128
1994	0.0151	0.0345	0.0501	2.8120	10.6376	0.0003	0.0041	0.0434	0.1396	2233
1995	0.0150	0.0341	0.0491	2.7934	10.9260	0.0004	0.0038	0.0429	0.1361	2281
1996	0.0152	0.0347	0.0510	3.2885	18.0173	0.0003	0.0040	0.0429	0.1357	2379
1997	0.0171	0.0356	0.0485	2.8332	12.9698	0.0004	0.0046	0.0465	0.1295	2477
1998	0.0169	0.0364	0.0509	2.9784	13.7575	0.0004	0.0045	0.0476	0.1365	2370
1999	0.0183	0.0390	0.0539	2.9050	12.8642	0.0005	0.0047	0.0527	0.1483	2210
2000	0.0202	0.0415	0.0565	2.5904	9.0605	0.0005	0.0051	0.0532	0.1612	2106
2001	0.0199	0.0419	0.0570	2.6184	9.7162	0.0006	0.0049	0.0572	0.1634	2057
2002	0.0182	0.0403	0.0553	2.5182	8.3604	0.0005	0.0051	0.0524	0.1577	2098
2003	0.0180	0.0385	0.0521	2.5156	8.8717	0.0004	0.0044	0.0533	0.1455	2260
2004	0.0169	0.0376	0.0521	2.6805	10.2624	0.0004	0.0040	0.0514	0.1427	2401
2005	0.0171	0.0377	0.0533	2.9547	13.0413	0.0004	0.0043	0.0505	0.1434	2448
2006	0.0177	0.0377	0.0523	2.7295	10.9992	0.0004	0.0042	0.0493	0.1426	2526
2007	0.0192	0.0405	0.0567	2.7475	10.6076	0.0004	0.0046	0.0517	0.1597	2520
2008	0.0209	0.0459	0.0673	3.2676	16.1572	0.0006	0.0061	0.0563	0.1817	2486
2009	0.0209	0.0515	0.0782	3.0029	11.9775	0.0005	0.0055	0.0647	0.2087	2499
2010	0.0193	0.0482	0.0738	3.1643	14.2904	0.0005	0.0046	0.0591	0.2033	2513
2011	0.0200	0.0472	0.0714	3.1912	14.8721	0.0004	0.0052	0.0578	0.1913	2447
2012	0.0201	0.0484	0.0726	3.1499	14.7723	0.0004	0.0050	0.0616	0.2011	2373
2013	0.0206	0.0484	0.0734	3.1878	14.9204	0.0005	0.0049	0.0594	0.1964	2309
2014	0.0209	0.0483	0.0721	3.0329	12.6298	0.0004	0.0052	0.0596	0.1927	2319
2015	0.0200	0.0476	0.0699	2.8858	11.0189	0.0005	0.0058	0.0596	0.1931	2319
2016	0.0197	0.0470	0.0684	2.8007	10.7733	0.0004	0.0052	0.0598	0.1876	2263
2017	0.0188	0.0467	0.0683	2.7674	10.4160	0.0003	0.0047	0.0592	0.1861	2190

Note: Table 1 reports year-by-year summary statistics on the panel of price informativeness measures recovered. It provides information on the median, mean, standard deviation, skewness, excess kurtosis, 5th, 25th, 75th, and 95th percentiles of each yearly distribution, as well as the number of stocks in each year. Since our panel of price informativeness is quarterly, we average the measures of quarterly price informativeness at the yearly level before computing the summary statistics. We start reporting summary statistics in 1980, since that is the first year with more than 250 stocks. Informativeness in year t is computed over a rolling window of 30 quarters (7.5 years) prior.

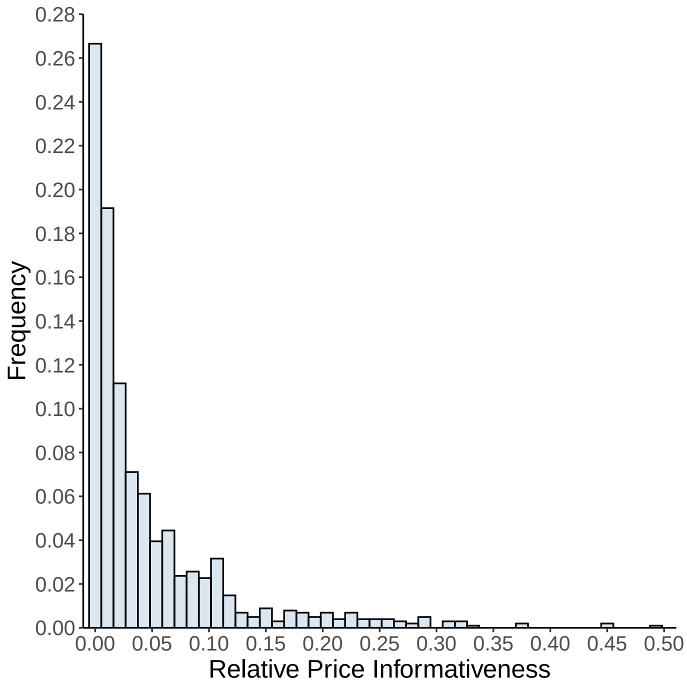

Figure 2: Price informativeness: relative frequency histogram

Note: Figure 2 shows a relative frequency histogram of price informativeness for a representative time period, the last quarter of 2015. Note that informativeness is computed over a rolling window of 30 quarters (7.5 years) prior. The histogram features $2,208$ stocks.

Table 1 reports year-by-year summary statistics of the distribution of stock-specific price informativeness, starting in 1980. Throughout the paper, note that informativeness in year t is computed over a rolling window of 30 quarters (7.5 years) prior. We illustrate our results graphically in Figure 2, which presents a relative frequency histogram of price informativeness for a representative specific time period (last quarter of 2015).

The distribution of informativeness that we recover is right-skewed every year, with a mean that is roughly 0.02 larger than the median of the distribution of informativeness. Because the distribution of informativeness is skewed, the median is often perceived to be a better measure of central tendency. The 95% percentile of the distribution stays between 0.012 and 0.02, which means that an external observer would rarely put more than a 2% weight on the price when updating beliefs to form a posterior over future payoffs. Since we have not included additional controls in the regressions besides quarter fixed-effects, our results should be interpreted as the price informativeness for an external observer who exclusively observes prices and past payoffs — see the Appendix for the results including additional controls, which correspond to a notion of informativeness in which the external observer has a larger information set, as explained in Section 4.2.

5.2 Price Informativeness in the Cross-Section

By computing stock-specific measures of price informativeness, we are able to establish a new set of cross-sectional patterns relating price informativeness to stock characteristics. We focus on five stock characteristics that have been widely used to explain patterns in the cross-section of stock returns — see, e.g., Bali, Engle and Murray (2016). These are i) size, measured as the natural log of stocks market capitalization, ii) value, measured as the ratio between a stock's book value and its market capitalization, iii) turnover, measured as the ratio of trading volume to shares outstanding, iv) idiosyncratic volatility, measured as the standard deviation — over a 30 month period — of the difference between the returns of a stock and the market return, and v) institutional ownership, measured as the proportion of shares held by institutional investors.

In Table 2, we report the estimates of panel regressions of relative price informativeness (in twentiles) on each of the five explanatory variables, using year-fixed effects. The coefficients that we report can be interpreted as a weighted average of the slopes of running year-by-year regressions of price informativeness of a given explanatory variable (size, value, turnover, return volatility, institutional ownership). Figures OA-2 through OA-6 provide an alternative graphical illustration of our results. These figures show that the cross-sectional relations identified in Table 2 are stable over time.[8]

Our cross-sectional analysis yields several robust patterns. First, we find a strong positive cross-sectional relation between a stock's size (market capitalization) and price informativeness, that is, large stocks have higher price informativeness. Second, we find a negative and weak cross-sectional relation between a stock's book-to-market ratio and price informativeness, that is, value stocks have lower price informativeness.[9] Third, we find a strong positive cross-sectional relation between a stock's turnover and price informativeness, that is, stocks that trade frequently have higher price informativeness. Fourth, we find a positive and weak cross-sectional relation between a stock's idiosyncratic return volatility and price informativeness, that is, stocks whose returns are more volatile have have higher price informativeness. Finally, we find a strong positive cross-sectional relation between a stock's institutional ownership share and price informativeness, that is, stocks owned mostly by institution investors have higher price informativeness.

There are no clear theoretical predictions on the sign of these cross-sectional relations. For instance, on the one hand, one could conceive that sophisticated investors acquire more private information about stocks with higher market capitalization because they can benefit from the information at a larger scale. This would imply a positive relation between price informativeness and size. However, on the other hand, it may be that larger firms attract the attention of more unsophisticated traders, which would make the prices of those stocks noisy and uninformative. Similarly, in the case of turnover, one could argue that high turnover reflects a large amount of non-payoff relevant trades and expect low price informativeness, or postulate that high turnover reflects a large volume of informed trading which would be associated with high price informativeness.

[8] In the Appendix, we show that these cross-sectional patterns remain valid using pooled measures of price informativeness over time.

[9] This is consistent with the findings in Farboodi et al. (2019).

Table 2: Cross-sectional results (1)

	Estimate	Std. Error	t-stat
Size	0.00222	0.000215	10.35
Value	-0.00051	0.000591	-0.86
Turnover	0.00031	0.000026	11.84
Idiosyncratic Volatility	0.01637	0.008324	1.97
Institutional Ownership	0.01523	0.001695	8.98

Note: Table 2 reports the estimates ($\widehat{a_1^c}$) of panel regressions of price informativeness on cross-sectional characteristics (in twentiles) with year-fixed effects (ξ_t): $\tau_{\pi,t}^{R,b} = a_0^c + a_1^c c_t^b + \xi_t + \epsilon_{b,t}$, where $\tau_\pi^{R,b,t}$ denotes the average price informative per bin (twentile) in a given period, c_t^b denotes the value of the given characteristic per bin (twentile) in a given period, ξ_t denotes a year-fixed effect, a_0^c and a_1^c are parameters, and $\epsilon_{b,t}$ is an error term. Figures OA-2 through OA-6 provide the graphical counterpart of the results in this table. Size is measured as the natural log of stocks market capitalization, value is measured as the ratio between a stock's book value and its market capitalization, turnover is measured as the ratio of trading volume to shares outstanding, idiosyncratic volatility is measured as the standard deviation — over a 30 month period — of the difference between the returns of a stock and the market return, and institutional ownership is measured as the proportion of stocks held by institutional investors.

Similar arguments can be given for the other characteristics.

That said, our results in Sections 2 and 3 allow us to provide a structural interpretation of our results. For instance, simply by relying on Equation (5) and (4), we can conclude that stocks with high informativeness, have a high ϕ_1 (sensitivity of the asset price to the future payoff) and $\tau_{\Delta n}$ (precision of the non-payoff relevant component, i.e., noise), relative to ϕ_n (sensitivity of the asset price to its non-payoff relevant component) and τ_u (precision of innovation to the payoff). Therefore, our empirical findings suggest that, for most stocks, either the ratio of demand sensitivities to information relative to noise $\frac{\phi_1}{\phi_n}$ is low or that the noise is too volatile relative to the volatility of the payoff, or both. As explained above, one would need additional information to separately identify noise from information in prices.

Figures 3 and 4 illustrate additional cross-sectional patterns of the behavior of informativeness by exchange, S&P 500 status, and sector. Instead of focusing on mean or median comparisons, we find it more informative to graphically compare the distributions of informativeness by characteristic after extracting time-fixed effects. While there are small differences, the relations seem less strong than those identified in Table 2. First, we compare across exchanges and find that stocks listed in the NYSE have higher median informativeness than those in the NASDAQ, which appear to be more informative than those listed in the AMEX. Second, we study whether price informativeness vary among stocks that belong to the S&P500 and those that do not. Consistently with our findings on size, we find that stocks outside of the S&P have a lower price informativeness on average. Finally, we study the behavior or price informativeness across sectors. We find that the median price informativeness is highest in the wholesale/retail and the finance/insurance sectors, and lowest in the service sector.

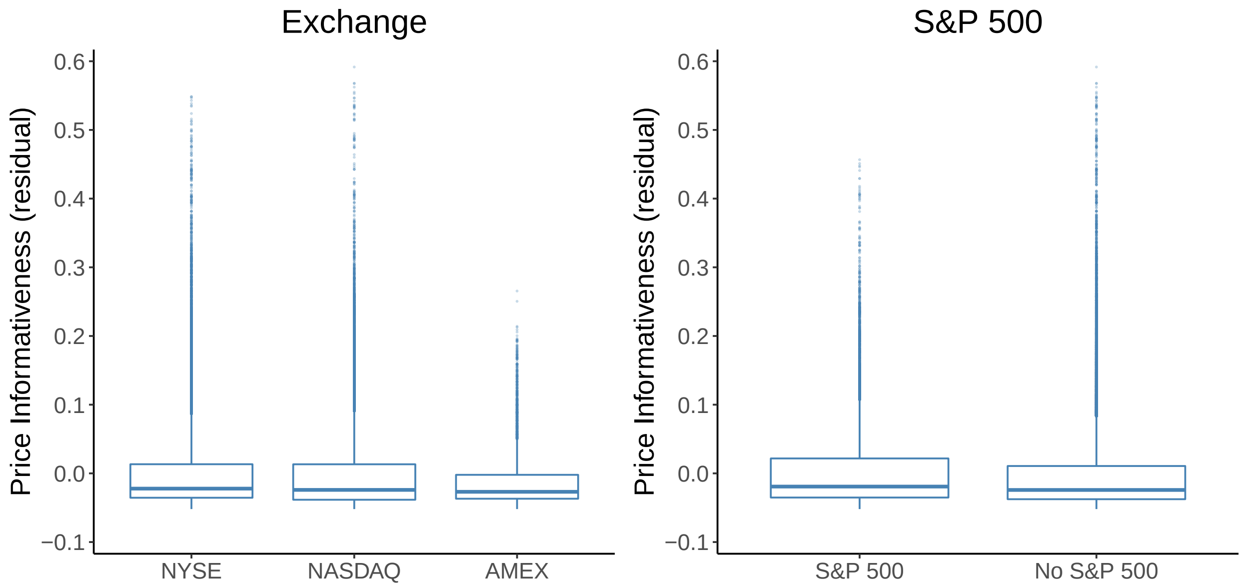

Figure 3: Cross-sectional results (2)

Note: The left panel in Figure 3 shows a box plot by exchange of the residuals of a regression of relative price informativeness on year-fixed effects. The left panel in Figure 3 shows a box plot by S&P 500 status of the residuals of a regression of relative price informativeness on year-fixed effects. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to ±1.5 times the interquartile range.

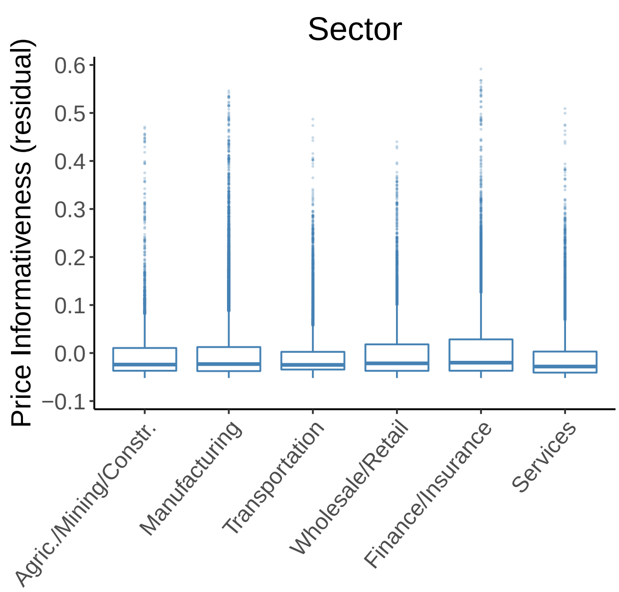

Figure 4: Cross-sectional results (3)

Note: Figure 4 shows a box plot by one-digit SIC industry code of the residuals of a regression of relative price informativeness on year-fixed effects. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to 1.5 times the interquartile range. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to ±1.5 times the interquartile range.

5.3 Price Informativeness Over Time

We also study how the distribution of stock-specific price informativeness evolves over time.[10] Table 1 includes a large amount of information about the time evolution of the distribution of informativeness. To better illustrate the results, we show the behavior of the median, the mean, and the standard deviation of the cross-section of informativeness between 1980 and 2017 graphically in Figure 5.

We find that both the median and the mean of the distribution of informativeness feature increasing trends starting in the mid-1980s. The median moves from roughly 1% to 2% between 1986 and 2017, while the mean moves from roughly 3% in 1986 until roughly 4.5% by 2017. The large and dispersed estimates of mean and median informativeness before 1985 are due to smaller sample size in that period. For that reason, we've decided to emphasize the steady increasing trend that starts in the mid-1980s. As in the cross-sectional case, there are no clear theoretical predictions on the behavior of these moments over time. The positive long-run trend that we identify is consistent with an increase in the precision of private signals or with a reduction of noise trading in financial markets among most stocks.

We also find that the standard deviation of informativeness has a positive long-run trend in our sample. In this case, there is a large spike in the cross-sectional standard deviation of informativeness around the global financial crisis of 2008 – other measures of dispersion have a similar behavior. Through the lens of our framework, this fact is consistent with significant changes in the amount of private information and noise among different stocks around that period.

Finally, Table OA-1 in the Appendix reports the cross-sectional correlation between price informativeness measured in year t and price informativeness measured in prior years, following the methodology of chapter 4 of Bali, Engle and Murray (2016). This table shows that our informativeness measures are persistent over time, especially in recent years. As one might expect, the strength of the correlation decays over time, with one-year cross correlations above 0.7, while 5 year cross-correlations can be as low as 0.1.

6 Conclusion

We have shown that the outcomes of regressions of price changes on changes in fundamentals can be combined to recover exact measures of price informativeness within a large class of linear or linearized models. Empirically, we compute a panel of stock-specific measures of price informativeness and find that the median and mean levels of price informativeness fluctuates around levels of 2% and 4%, respectively. These values can be interpreted as the weight that an external observer puts on the price signalwhen forming a posterior belief about future payoffs. Cross-sectionally, we find that price informativeness is robustly higher for stocks with higher market capitalization, that

[10] To keep the paper focused, we exclusively study there the behavior of the panel of stock-specific price informativeness measures. There is scope to apply our methodology to aggregate data in order to generate a time-series of aggregate price informativeness.

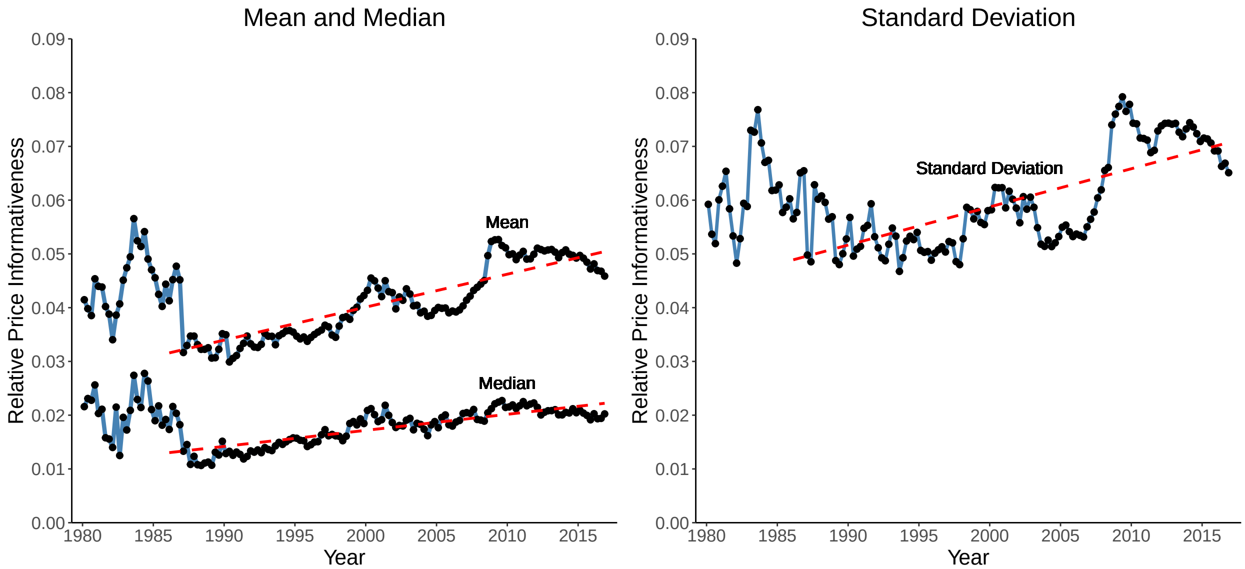

Figure 5: Price informativeness over time

Note: The left panel in Figure 5 shows the time series evolution of the mean and median relative price informativeness. The right panel in Figure 5 shows the time series evolution of the standard deviation of price informativeness. The red dashed lines show linear trends starting in 1986. In both panels, the dots correspond to the average within a quarter of the price informativeness measures computed using quarterly data.

trade more frequently, and that have a higher institutional share. Over time, we find that the mean and median price informativeness have steadily increased since the mid 1980's.

Our methodology opens the door to answering a broad set of questions related to the role of information aggregation in asset markets. Empirically, there is scope to further explore the relation between price informativeness measures and other characteristics in the cross-section or over time. It also seems worthwhile to document and explain the behavior of price informativeness in different contexts, perhaps internationally or in different markets. Theoretically, our results can be used to discipline theories of information and learning in financial markets. Finally, we hope that our approach encourages further research on how to formally identify price informativeness in richer models, for instance, those with feedback effects or with significant non-linearities.

References

Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski. 2014. "Dynamic Dispersed Information and the Credit Spread Puzzle." *NBER Working Paper.*

Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski. 2015. "A Theory of Asset Prices based on Heterogeneous Information." *NBER Working Paper.*

Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski. 2017. "Imperfect Financial Markets and Shareholder Incentives in Partial and General Equilibrium." *NBER Working Paper.*

Bai, Jennie, Thomas Philippon, and Alexi Savov. 2015. "Have financial markets become more informative?" *Journal of Financial Economics, Forthcoming.*

Bali, Turan G., Robert F. Engle, and Scott Murray. 2016. *Empirical Asset Pricing: The Cross Section of Stock Returns.* John Wiley & Sons.

Blackwell, David. 1953. "Equivalent comparisons of experiments." *The Annals of Mathematical Statistics*, 265–272.

Bond, Philip, Alex Edmans, and Itay Goldstein. 2012. "The Real Effects of Financial Markets." *Annual Review of Financial Economics*, 4(1): 339–360.

Campbell, John Y. 2017. *Financial decisions and markets: a course in asset pricing.* Princeton University Press.

Campbell, John Y., and Robert J. Shiller. 1988. "The dividend-price ratio and expectations of future dividends and discount factors." *The Review of Financial Studies*, 1(3): 195–228.

Chen, Qi, Itay Goldstein, and Wei Jiang. 2006. "Price informativeness and investment sensitivity to stock price." *The Review of Financial Studies*, 20(3): 619–650.

Cochrane, John. 2005. *Asset Pricing: (Revised).* . Revised ed., Princeton University Press.

Davila, Eduardo, and Cecilia Parlatore. 2017. "Trading Costs and Informational Efficiency." *Working Paper.*

De Long, J Bradford, Andrei Shleifer, Lawrence H Summers, and Robert J Waldmann. 1990. "Noise Trader Risk in Financial Markets." *Journal of Political Economy*, 98(4): 703–738.

Diamond, Douglas W, and Robert E Verrecchia. 1981. "Information aggregation in a noisy rational expectations economy." *Journal of Financial Economics*, 9(3): 221–235.

Durnev, Art, Randall Morck, and Bernard Yeung. 2004. "Value-enhancing capital budgeting and firm-specific stock return variation." *The Journal of Finance*, 59(1): 65–105.

Easley, D., M. O'Hara, and J. Paperman. 1998. "Financial analysts and information-based trade." *Journal of Financial Markets*, 1: 175–201.

Farboodi, Maryam, Adrien Matray, Laura Veldkamp, and Venky Venkateswaran. 2019. "Where Has All the Data Gone?" *Working Paper.*

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

Hayek, FA. 1945. "The Use of Knowledge in Society." *The American Economic Review*, 35(4): 519–530.

Hellwig, M.F. 1980. "On the aggregation of information in competitive markets." *Journal of Economic Theory*, 22(3): 477–498.

Hou, Kewei, Lin Peng, and Wei Xiong. 2013. "Is R2 a Measure of Market Inefficiency?" *working paper.*

Kacperczyk, Marcin, Savitar Sundaresan, and Tianyu Wang. forthcoming. "Do Foreign Institutional Investors Improve Market Efficiency?" *Review of Financial Studies.*

Morck, Randall, Bernard Yeung, and Wayne Yu. 2000. "Tehe information content of stock markets: Why do emerging markets have synchronous stock price movements?" *Journal of Financial Economics*, 58: 215–260.

Roll, Richard. 1988. "R2." *The Journal of Finance*, 43(3): 541–566.

Stambaugh, R.F. 1999. "Predictive regressions." *Journal of Financial Economics*, 54(3): 375–421.

Veldkamp, Laura. 2011. *Information Choice in Macroeconomics and Finance.* Princeton University Press.

Vives, Xavier. 2008. *Information and Learning in Markets: the Impact of Market Microstructure.* Princeton University Press.

Wurgler, Jeffrey. 2000. "Financial markets and the allocation of capital." *Journal of Financial Economics*, 58: 187–214.

Appendix

A Proofs and Derivations

A.1 Section 2: Proofs and derivations

Proof of Proposition 1. (Identifying price informativeness)

For completeness, we reproduce here Regressions R1 and R2:

$$\Delta p_t = \overline{\beta} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t+1} + e_t \tag{R1}$$

$$\Delta p_t = \overline{\zeta} + \zeta_0 \Delta x_t + \varepsilon_t^{\zeta}. \tag{R2}$$

Note that the R-squareds of both regressions can be expressed as follows

$$R^2_{\Delta x, \Delta x'} = 1 - \frac{\mathbb{V}\mathrm{ar}\left(e_t\right)}{\mathbb{V}\mathrm{ar}\left(\Delta p_t\right)} \quad \text{and} \quad R^2_{\Delta x} = \frac{\mathbb{V}\mathrm{ar}\left(\zeta_0 \Delta x_t\right)}{\mathbb{V}\mathrm{ar}\left(\Delta p_t\right)}.$$

After substituting Equation (1) in Regression R1, the following relation holds

$$\Delta p_t = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n} + \left(\phi_0 + \rho \phi_1\right) \Delta x_t + \phi_1 u_t + \phi_n \varepsilon_t^n. \tag{25}$$

By comparing Regression R2 with the structural Equation (25), it follows that $\overline{\zeta} = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n}$, $\zeta_0 = \phi_0 + \rho \phi_1$, and $\varepsilon_t^{\zeta} = \phi_1 u_t + \phi_n \varepsilon_t^{\Delta n}$.

From Equation (25), the following variance decomposition must hold

$$\begin{aligned} \mathbb{V}\mathrm{ar}\left(\Delta p_t\right) &= \mathbb{V}\mathrm{ar}\left(\zeta_0 \Delta x_t\right) + \mathbb{V}\mathrm{ar}\left(\phi_1 u_t + \phi_n \varepsilon_t^{\Delta n}\right) \\ &= \mathbb{V}\mathrm{ar}\left(\zeta_0 \Delta x_t\right) + \left(\phi_1\right)^2 \mathbb{V}\mathrm{ar}\left(u_t\right) + \mathbb{V}\mathrm{ar}\left(e_t\right) \end{aligned}$$

which can be rearranged to express $\frac{\tau_\pi}{\tau_u}$ as follows

$$1 = \underbrace{\frac{\mathbb{V}\mathrm{ar}\left(\zeta_0 \Delta x_t\right)}{\mathbb{V}\mathrm{ar}\left(\Delta p_t\right)}}_{R^2_{\Delta x}} + \underbrace{\frac{\mathbb{V}\mathrm{ar}\left(e_t\right)}{\mathbb{V}\mathrm{ar}\left(\Delta p_t\right)}}_{1 - R^2_{\Delta x, \Delta x'}} \left(\underbrace{\frac{\left(\phi_1\right)^2}{\mathbb{V}\mathrm{ar}\left(e_t\right)} \mathbb{V}\mathrm{ar}\left(u_t\right)}_{\frac{\tau_\pi}{\tau_u}} + 1 \right) \Rightarrow \frac{\tau_\pi}{\tau_u} = \frac{R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x, \Delta x'}}$$

Therefore, relative price informativeness can be written as

$$\tau_\pi^R = \frac{\tau_\pi}{\tau_\pi + \tau_u} = \frac{1}{1 + \frac{1}{\frac{\tau_\pi}{\tau_u}}} = \frac{R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x}}.$$

□

Absolute Price Informativeness

Proposition 6. (Identifying absolute price informativeness) *Let $\overline{\beta}$, β_0, and β_1 denote the coefficients of the following regression of log-price differences on realized and future log-payoff differences,*

$$\Delta p_t = \overline{\beta} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t+1} + e_t, \tag{R1}$$

where $\Delta p_t = p_t - p_{t-1}$ denotes the date t change in log-price, $\Delta x_t = x_t - x_{t-1}$ and $\Delta x_{t+1} = x_{t+1} - x_t$ respectively denote the date t and $t+1$ log-payoff differences.

Proof. By comparing Regression R1 with the structural Equation (2), it follows that $\overline{\beta} = \overline{\phi} + \phi_n \mu_{\Delta n}$, $\beta_0 = \phi_0$, $\beta_1 = \phi_1$, and $e_t = \phi_n \varepsilon_t^{\Delta n}$. Consequently, $\sigma_e^2 = \mathbb{V}\mathrm{ar}\left[e_t\right] = \left(\phi_n\right)^2 \mathbb{V}\mathrm{ar}\left[\varepsilon_t^{\Delta n}\right] = \left(\phi_n\right)^2 \tau_{\Delta n}^{-1}$. Therefore, we can recover absolute price informativeness as follows

$$\tau_\pi = \frac{\left(\beta_1\right)^2}{\sigma_e^2} = \frac{\left(\phi_1\right)^2}{\left(\phi_n\right)^2 \tau_{\Delta n}^{-1}} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_{\Delta n}.$$

Given the assumptions on u_t and Δn_t, it is straightforward to show that the OLS estimates of Regressions R1 and R2 are consistent, which implies that price informativeness can be consistently estimated as $\widehat{\tau_\pi} = \frac{\left(\widehat{\beta_1}\right)^2}{\widehat{\sigma_e^2}}$. Formally, $\mathrm{plim}\left(\widehat{\tau_\pi}\right) = \mathrm{plim}\left(\frac{\left(\widehat{\beta_1}\right)^2}{\widehat{\sigma_e^2}}\right) = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_{\Delta n} = \tau_\pi$. $\square$

Kalman gain for a Bayesian external observer

The priors of a Bayesian external observer over u_t and Δn_t are given by

$$u_t \sim N\left(\mu_u, \tau_u^{-1}\right) \quad \text{and} \quad \Delta n_t \sim N\left(\mu_{\Delta n}, \tau_{\Delta n}^{-1}\right).$$

The external observer only learns from the price. The unbiased signal contained in the price about the innovation to the payoff is

$$\pi_t = u_t + \frac{\phi_n}{\phi_1}\left(\Delta n_t - \mu_{\Delta n}\right),$$

where

$$\pi_t | \, u_t \sim N\left(u_t, \tau_\pi^{-1}\right).$$

A standard application of Bayesian updating (see Vives (2008) or Veldkamp (2011)) immediately implies that the posterior distribution of an external observer who makes use of the price as a signal about the innovation to the fundamental is given by

$$\begin{aligned} \hat{\mu}_t &\equiv \mathbb{E}\left[u_t | \, \pi_t\right] = (1 - K)\,\mu_t + K\pi_t \\ \hat{\tau}_u &\equiv \mathbb{V}ar\left[u_t | \, \pi_t\right] = \left(\tau_u + \tau_\pi\right)^{-1}, \end{aligned}$$

where $K = \frac{\tau_\pi}{\tau_u + \tau_\pi}$ is the Kalman gain and it represents the weight that a Bayesian observer that only learns from the price puts on the information contained in the price. This system is equivalent to Equation (7).

To see that the reduction in uncertainty about the innovation to the fundamental for the external external observer is given by the Kalman gain note that

$$1-\frac{\mathbb{V}ar\left[u_t|\,\pi_t\right]}{\mathbb{V}ar\left[u_t\right]}=1-\frac{\left(\tau_u+\tau_\pi\right)^{-1}}{\tau_u^{-1}}=\frac{\tau_\pi}{\tau_u+\tau_\pi}=K.$$

A.2 Section 3: Proofs and derivations

Portfolio Demand Approximation

The optimality condition of an investor that maximizes Equation (10) subject to the wealth accumulation constraint in Equation (11) is given by

$$\mathbb{E}\left[U_i'\left(w_{t+1}^i\right)\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\middle|\mathcal{I}_t^i\right]=0. \tag{26}$$

We approximate an investor's first order condition in three steps.

First, we take a first order Taylor expansion of an investor's future marginal utility $U'\left(w_{t+1}^i\right)$ around the current date t wealth level w_t^i. Formally, we approximate $U'\left(w_{t+1}^i\right)$ as follows

$$U'\left(w_{t+1}^i\right)\approx U'\left(w_t^i\right)+U''\left(w_t^i\right)\Delta w_{t+1}^i,$$

which allows us to express Equation (26) as

$$U'\left(w_t^i\right)\mathbb{E}_i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]+U''\left(w_t^i\right)w_t^i\mathbb{E}_i\left[\left(\left(R^f-1\right)+\theta_t^i\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\right)\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\right]\approx 0.$$

Second, we impose that terms that involve the product of two or more net interest rates are negligible. In continuous time, these terms would be of order $(dt)^2$. Formally, it follows that

$$\left(R^f-1\right)\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]\approx 0 \quad \text{and} \quad \left(\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]\right)^2\approx 0,$$

which allows us to express Equation (26) as

$$U'\left(w_t^i\right)\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]+U''\left(w_t^i\right)w_t^i\theta_t^i\mathbb{V}ar_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}\right]\approx 0.$$

Therefore, we can express an investor's risky portfolio share θ_t^i as

$$\theta_t^i\approx\frac{1}{\gamma^i}\frac{\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]}{\mathbb{V}ar_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}\right]}, \tag{27}$$

where $\gamma^i\equiv-\frac{w^iU''(w^i)}{U'(w^i)}$ denotes the coefficient of relative risk aversion.

Third, as in Campbell and Shiller (1988), we take a log-linear approximation of returns around a predetermined dividend-price ratio. Formally, note that

$$
\begin{aligned}
\frac{X_{t+1}+P_{t+1}}{P_t} &= e^{\ln\left(\frac{\left(1+\frac{P_{t+1}}{X_{t+1}}\right)\frac{X_{t+1}}{X_t}}{\frac{P_t}{X_t}}\right)} \\
\ln\left(\frac{X_{t+1}+P_{t+1}}{P_t}\right) &= \ln\left(1+\frac{P_{t+1}}{X_{t+1}}\right)+x_{t+1}-x_t-(p_t-x_t) \\
&= \ln\left(1+e^{p_{t+1}-x_{t+1}}\right)+\Delta x_{t+1}-(p_t-x_t),
\end{aligned}
$$

where $y_t = \ln Y_t$ for all variables Y. Following Campbell and Shiller (1988), we approximate the first term around a point $PX = e^{p-x}$, to find that

$$
\begin{aligned}
\ln\left(1+e^{p_{t+1}-x_{t+1}}\right) &\approx \ln(1+PX)+\frac{PX}{PX+1}\left(p_{t+1}-x_{t+1}-(p-x)\right). \\
&= k_0+k_1\left(p_{t+1}-x_{t+1}\right),
\end{aligned}
$$

where $k_1 \equiv \frac{PX}{PX+1}$ and $k_0 \equiv \ln(1+PX)-k_1(p-x)$.

Therefore, starting from Equation (27), we can express an investor's risky portfolio share θ_t^i as

$$
\theta_t^i \approx \frac{1}{\gamma^i}\frac{k_0+k_1\mathbb{E}_t^i\left[p_{t+1}-x_{t+1}\right]+\mathbb{E}_t^i\left[\Delta x_{t+1}\right]-(p_t-x_t)-r^f}{\mathbb{V}ar_t^i\left[k_1\left(p_{t+1}-x_{t+1}\right)+\Delta x_{t+1}\right]}, \tag{28}
$$

where we define $r^f \equiv \ln R^f$ and we used that $e^y \approx 1+y$.

Forming expectations

In order to characterize the equilibrium it is necessary to characterize investors' expectations. We conjecture and subsequently verify that $k_1\mathbb{E}_t^i\left[p_{t+1}-x_{t+1}\right]+\mathbb{E}_t^i\left[\Delta x_{t+1}\right]$ is linear in s_t^i, $\overline{n}_t^i$, and x_t and that $\mathbb{V}ar_t^i\left[k_1\left(p_{t+1}-x_{t+1}\right)+\Delta x_{t+1}\right]$ is a constant. Under this conjecture, θ_t^i is a linear function of s_t^i, x_t, $\overline{n}_t^i$ and it is given by

$$
\theta_t^i \approx \alpha_x^i x_t+\alpha_s^i s_t^i+\alpha_n^i\overline{n}_t^i-\alpha_p^i p_t+\psi^i.
$$

Using this expression and the market clearing condition $\int \theta_t^i w_t^i di = Q$ implies

$$
p_t = \frac{\overline{\alpha_x}}{\overline{\alpha_p}}x_t+\frac{\overline{\alpha_s}}{\overline{\alpha_p}}u_t+\frac{\overline{\alpha_n}}{\overline{\alpha_p}}n_t+\frac{\overline{\psi}}{\overline{\alpha_p}},
$$

where $\overline{\alpha_h} \equiv \int \alpha_h^i w_t^i di$ for $h = x, s, n, p$ and $\overline{\psi} \equiv \int \psi^i w_t^i di - Q$. As in Vives (2008), we make use the Strong Law of Large Numbers, since the sequence of independent random variables $\left\{\alpha_h^i w_t^i \tau_h^{-1}\right\}$ has uniformly bounded variance and finite means. This expression can also be written as

$$
p_t = \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}}-\frac{\overline{\alpha_s}}{\overline{\alpha_p}}\right)x_t+\frac{\overline{\alpha_s}}{\overline{\alpha_p}}x_{t+1}+\frac{\overline{\alpha_n}}{\overline{\alpha_p}}n_t+\frac{\overline{\psi}}{\overline{\alpha_p}}-\frac{\overline{\alpha_s}}{\overline{\alpha_p}}\mu_{\Delta x}. \tag{29}
$$

Investors in the model learn from the price. The information contained in the price for an investor in the model is

$$\hat{\pi}_t = \frac{\overline{\alpha_p}}{\overline{\alpha_s}} \left(p_t - \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \mu_{\Delta n} + \frac{\overline{\psi}}{\overline{\alpha_p}} \right) \right)$$

which has a precision

$$\tau_{\hat{\pi}} \equiv \mathbb{V}ar \left[\hat{\pi}_t \middle| u_t, \{x_s\}_{s \leq t}, p_{t-1} \right]^{-1} = \left(\frac{\overline{\alpha_s}}{\overline{\alpha_n}} \right)^2 \tau_{\Delta n}.$$

Note that given the information set of the investor, $\mathbb{V}ar \left[n_t \middle| u_t, \{x_s\}_{s \leq t}, p_{t-1} \right] = \mathbb{V}ar \left[\Delta n_t \middle| u_t, \{x_s\}_{s \leq t}, p_{t-1} \right]$. Then,

$$\mathbb{E}_t^i [u_t] = \mathbb{E} \left[u_t | \mathcal{I}_t^i \right] = \frac{\tau_s s_t^i + \tau_u \overline{n}_t^i + \tau_{\hat{\pi}} \hat{\pi}_t}{\tau_s + \tau_u + \tau_{\hat{\pi}}} = \frac{\tau_s s_t^i + \tau_u \overline{n}_t^i + \tau_{\hat{\pi}} \frac{\overline{\alpha_p}}{\overline{\alpha_s}} \left(p_t - \frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t - \frac{\overline{\alpha_n}}{\overline{\alpha_s}} \mu_{\Delta n} - \frac{\overline{\psi}}{\overline{\alpha_p}} \right)}{\tau_s + \tau_u + \tau_{\hat{\pi}}}$$

and

$$\mathbb{V}ar \left[u_t \middle| \mathcal{I}_t^i \right] = (\tau_s + \tau_u + \tau_{\hat{\pi}})^{-1},$$

where $\mathcal{I}_t^i = \left\{ s_t^i, \overline{n}_t^i, \{p_s\}_{s \leq t}, \{x_s\}_{s \leq t} \right\}$. Note that these two expressions imply that our conjectures above are satisfied. To see this, note that

$$k_1 \mathbb{E}_t^i [p_{t+1} - x_{t+1}] + \mathbb{E}_t^i [\Delta x_{t+1}] = k_1 \mathbb{E}_t^i \left[\frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_{t+1} + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_{t+1} + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} n_{t+1} + \frac{\overline{\psi}}{\overline{\alpha_p}} - x_{t+1} \right] + \mu_{\Delta x} + \mathbb{E}_t^i [u_t] \tag{30}$$

$$\begin{aligned}
&= k_1 \left(\mathbb{E}_t^i \left[\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 \right) x_{t+1} + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_{t+1} \right] + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \mu_{\Delta n} + \frac{\overline{\psi}}{\overline{\alpha_p}} \right) + \mu_{\Delta x} + \mathbb{E}_t^i [u_t] \\
&= k_1 \left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 \right) \left(\mu_{\Delta x} + \mathbb{E}_t^i [u_t] \right) + \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 \right) x_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \mu_{\Delta n} + \frac{\overline{\psi}}{\overline{\alpha_p}} \right) + \mu_{\Delta x} + \mathbb{E}_t^i [u_t] \\
&= k_1 \left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \left(\mu_{\Delta x} + \mathbb{E}_t^i [u_t] \right) + + \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 \right) x_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \mu_{\Delta n} + \frac{\overline{\psi}}{\overline{\alpha_p}} \right)
\end{aligned}$$

where we used that $\mathbb{E}_t^i [u_{t+1}] = 0$ and that $\mathbb{E}_t^i [\varepsilon_{t+1}^n] = 0$. Moreover,

$$\mathbb{V}ar_t^i [k_1 (p_{t+1} - x_{t+1}) + \Delta x_{t+1}] = \mathbb{V}ar_t^i \left[k_1 \left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \Delta x_{t+1} + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_{t+1} + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} n_{t+1} \right) \right] \tag{31}$$

$$\begin{aligned}
&= k_1^2 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right)^2 \mathbb{V}ar_t^i [u_t] + k_1^2 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} \right)^2 \mathbb{V}ar_t^i [u_{t+1}] + k_1^2 \left(\frac{\overline{\alpha_n}}{\overline{\alpha_p}} \right)^2 \mathbb{V}ar_t^i [\varepsilon_t^n] \\
&= k_1^2 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right)^2 (\tau_s + \tau_u + \tau_{\hat{\pi}})^{-1} + k_1^2 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} \right)^2 \tau_u^{-1} + k_1^2 \left(\frac{\overline{\alpha_n}}{\overline{\alpha_p}} \right)^2 \tau_{\Delta n}^{-1}.
\end{aligned}$$

Using these expressions in the first order condition and matching coefficients gives

$$\alpha_x^i = \frac{1}{\kappa_i} k_1 \left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \left(1 - \frac{\tau_{\hat{\pi}} \frac{\overline{\alpha_x}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \right) \right) \tag{32}$$

$$\alpha_s^i = \frac{1}{\kappa_i} k_1 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \frac{\tau_s}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \tag{33}$$

$$\alpha_n^i = \frac{1}{\kappa_i} k_1 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \frac{\tau_u}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \tag{34}$$

$$\alpha_p^i = \frac{1}{\kappa_i} \left(1 - k_1 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \frac{\tau_{\hat{\pi}} \frac{\overline{\alpha_p}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \right) \tag{35}$$

$$\psi^i = \frac{1}{\kappa_i} \left(k_0 + k_1 \left(- \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \left(\frac{\tau_{\pi} \frac{\overline{\alpha_p}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} - \mu_{\Delta x} \right) + 1 \right) \left(\frac{\overline{\alpha_n}}{\overline{\alpha_p}} \mu_{\Delta n} + \frac{\overline{\psi}}{\overline{\alpha_p}} \right) - r^f \right) \tag{36}$$

where $\kappa_i \equiv \gamma^i \mathbb{V}ar_t^i \left[k_1 \left(p_{t+1} - x_{t+1} \right) + \Delta x_{t+1} \right]$.

Proof of Lemma 1

Iterating forward Eq. (29) and taking differences

$$\Delta p_t = \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}} \right) \Delta x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} \Delta x_{t+1} + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \Delta n_t.$$

This maps to the price process in the general framework by setting $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, and $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$.

□

Proof of Lemma 2

When investors are the identical, the noise in their signal does not disappear from the price and the price in (29) becomes

$$p_t = \left(\frac{\alpha_x}{\alpha_p} - \frac{\alpha_s}{\alpha_p} \right) x_t + \frac{\alpha_s}{\alpha_p} \left(x_{t+1} + \varepsilon_{st} \right) + \frac{\alpha_n}{\alpha_p} n_t + \frac{\psi - Q}{\alpha_p} - \frac{\alpha_s}{\alpha_p} \mu_{\Delta x},$$

where the demand coefficients are given by the system in Equations (32) - (36). Iterating backwards this price and taking differences we have

$$\Delta p_t = \left(\frac{\alpha_x}{\alpha_p} - \frac{\alpha_s}{\alpha_p} \right) \Delta x_t + \frac{\alpha_s}{\alpha_p} \Delta x_{t+1} + \frac{\alpha_n}{\alpha_p} \left(\Delta n_t + \frac{\alpha_s}{\alpha_n} \Delta \varepsilon_{st} \right).$$

Setting $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, and $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$ and where $\Delta \hat{n}_t \equiv \Delta \overline{n}_t + \frac{\alpha_s}{\alpha_n} \Delta \varepsilon_{st}$ maps to the process in the general framework.

Note that from Equations (32) - (36) one can see that there is time-varying risk aversion $\frac{\alpha_{ht}}{\alpha_{pt}} = \frac{\alpha_{ht-1}}{\alpha_{pt-1}} = \frac{\alpha_h}{\alpha_p}$ for all t and $h = x, s, n$, and that $\frac{\psi_t - Q}{\alpha_{pt}} - \frac{\psi_{t-1} - Q}{\alpha_{pt-1}} = \Delta\gamma_t \frac{\mathbb{V}ar_t^i[k_1(p_{t+1} - x_{t+1}) + \Delta x_{t+1}]}{\left(1 - k_1\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)\frac{\tau_{\hat{\pi}}\frac{\overline{\alpha_p}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}}\right)} Q = \frac{\Delta\gamma_t}{\gamma_t \alpha_{pt}} Q$. In this case the price process is

$$\Delta p_t = \left(\frac{\alpha_x}{\alpha_p} - \frac{\alpha_s}{\alpha_p}\right)\Delta x_t + \frac{\alpha_s}{\alpha_p}\Delta x_{t+1} + \frac{\alpha_n}{\alpha_p}\left(\Delta n_t + \frac{\alpha_s}{\alpha_n}\Delta\varepsilon_{st} + \frac{\Delta\gamma_t}{\gamma_t \alpha_{nt}} Q\right)$$

and setting $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, and $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$ and where $\Delta\hat{n}_t \equiv \Delta\overline{n}_t + \frac{\alpha_s}{\alpha_n}\Delta\varepsilon_{st} + \frac{\Delta\gamma_t}{\gamma_t \alpha_{nt}} Q$ maps to the process in the general framework. In this case, the noise in the price can come from time-varying risk aversion.

Proof of Lemma 3

The case in which there are informed and uninformed investors and noise traders is a special case of the model in Section 3.1 with three types of agents. Then, the demand for informed investors is

$$\theta_t^I \approx \alpha_x^I x_t + \alpha_s^I u_t - \alpha_p^I p_t + \psi^I,$$

the demand for uninformed investors is

$$\theta_t^U \approx \alpha_x^U x_t + \alpha_n^U \overline{n}_t^U - \alpha_p^U p_t + \psi^U,$$

and the demand of noise traders is given by δ. Market clearing and the SLLN imply that the equilibrium price is

$$p_t \approx \frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \overline{n}_t + \frac{\overline{\psi}}{\overline{\alpha_p}} + \delta_t,$$

where $\overline{\alpha_y} \equiv \int_{I \cup U} \alpha_y^i w_t^i di$ denotes the wealth-weighted cross-sectional average of α_y^i over the set of informed and uninformed investors with $\alpha_s^i = 0$ for all uninformed investors, $\alpha_n^i = 0$ for all informed investors, and $\overline{\psi} \equiv \int_{I \cup U} \psi^i w_t^i di - Q$.

Taking first differences for this price process we have

$$\Delta p_t \approx \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta\tilde{n}_t,$$

where the coefficients $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, and $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$ are equilibrium outcomes and $\Delta\tilde{n}_t \equiv \Delta\overline{n}_t + \frac{\overline{\alpha_p}}{\overline{\alpha_n}}\Delta\delta_t$, which proves our claim.

$\square$

Proof of Lemma 4

From Assumption 2 and market clearing it follows that

$$Q = \int \alpha_s^i s_t^i w_t^i di + \int \alpha_x^i w_t^i di x_t + \int \alpha_n^i \overline{n}_t^i w_t^i di - \int \alpha_p^i w_t^i di p_t + \int \psi^i w_t^i di.$$

Using Assumption 1 we have

$$Q = \int \alpha_s^i \left(u_t + \varepsilon_{st}^i\right) w_t^i di + \overline{\alpha_x} x_t + \int \alpha_n^i \left(n_t + \varepsilon_{nt}^i\right) w_t^i di - \overline{\alpha_p} p_t + \int \psi^i w_t^i di.$$

Using the Strong Law of Large Numbers, since the sequence of independent random variables $\left\{\alpha_h^i w_t^i \tau_h^{-1}\right\}$ has uniformly bounded variance and finite means for $h = s, n$, we have the following equilibrium price

$$p_t \approx \frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \overline{n}_t + \frac{\overline{\psi}}{\overline{\alpha_p}}.$$

Taking differences gives

$$\Delta p_t \approx \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t,$$

where $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, and $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$. This proves our result.

$\square$

A.3 Section 4: Proofs and derivations

A.3.1 Proof of Proposition 2 (Identifying price informativeness when payoff and noise are correlated)

When

$$\Delta n_t = \mu_{\Delta n} + \omega u_t + \varepsilon_t^n,$$

the price process

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t$$

can be written as

$$\Delta p_t = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n} + \left(\phi_0 + \rho \phi_1\right) \Delta x_t + \left(\phi_1 + \phi_n \omega\right) u_t + \phi_n \varepsilon_t^n. \tag{37}$$

Hence, the unbiased signal about the innovation u_t contained in the price change is

$$\pi_t = \frac{1}{\left(\phi_1 + \phi_n \omega\right)} \left(\Delta p_t - \left(\overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n} + \left(\phi_0 + \rho \phi_1\right) \Delta x_t\right)\right) = u_t + \frac{\phi_n}{\left(\phi_1 + \phi_n \omega\right)} \varepsilon_t^n,$$

and absolute price informativeness is given by

$$\tau_\pi \equiv \mathbb{V}ar\left[\pi_t \middle| u_t, \Delta x_t\right]^{-1} = \left(\frac{\phi_1 + \phi_n \omega}{\phi_n}\right)^2 \tau_{\Delta n}.$$

The price process can also be written as a function of Δx_t and Δx_{t+1} as follows

$$\Delta p_t = \overline{\phi} - \phi_n \omega \mu_{\Delta x} + \phi_n \mu_{\Delta n} + (\phi_0 - \phi_n \omega)\,\Delta x_t + (\phi_1 + \phi_n \omega)\,\Delta x_{t+1} + \phi_n \varepsilon_t^n.$$

Comparing this equation with Regression R1 it is easy to see that $\overline{\beta} = \overline{\phi} - \phi_n \omega \mu_{\Delta x} + \phi_n \mu_{\Delta n}$, $\beta_0 = \phi_0 - \phi_n \omega$, $\beta_1 = \phi_1 + \omega \phi_n$, and $e_t = \phi_n \varepsilon_t^n$ with $e_t \perp \Delta x_t, \Delta x_{t+1}$. Moreover, comparing Equation (37) with Regression R2, it follows that $\overline{\zeta} = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n}$, $\zeta_0 = \phi_0 + \rho \phi_1$, and $e_t^{\zeta} = (\phi_1 + \phi_n \omega)\, u_t + \phi_n \varepsilon_t^n$ with $e_t^{\zeta} \perp \Delta x_t$. Then, we can recover relative price informativeness in the same way as in the baseline model, as follows

$$\tau_\pi^R = \frac{R^2_{\Delta x, \Delta x'} - R^2_{\Delta x}}{1 - R^2_{\Delta x}},$$

where the steps to recover absolute price informativeness are exactly the same as those in the baseline model.

□

A.3.2 Proof of Proposition 3 (Identifying price informativeness with public signals)

When investors have access to public signals, we assume that the equilibrium price process is given

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_\chi \cdot \Delta \chi_t + \phi_n \Delta n_t. \tag{38}$$

a) If the public signals are not available to the external observer, then the model with public information can be cast in terms of having noise correlated with the payoff by replacing the noise term in Eq. (2) by

$$\phi_n' \Delta n_t' \equiv \phi_\chi \cdot \Delta \chi_t + \phi_n \Delta n_t = \phi_n \mu_{\Delta n} + \phi_\chi \cdot \omega \cdot u_t + \phi_\chi \Delta \overline{\varepsilon}_t^{\chi} + \phi_n \varepsilon_t^n.$$

In this case, Proposition 2 holds.

b) If the public signals are available to the external observer, one can extend the identification results as follows. Using that $\Delta x_{t+1} = \mu_{\Delta x} + \rho \Delta x_t + u_t$ and that $\Delta n_t = \mu_{\Delta n} + \varepsilon_t^n$, we can write the price process as

$$\Delta p_t = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n} + (\phi_0 + \phi_1 \rho)\,\Delta x_t + \phi_1 u_t + \phi_\chi \cdot \Delta \chi_t + \phi_n \varepsilon_t^n \tag{39}$$

Hence, the unbiased signal contained in the price for an external observer who has access to the public signals is

$$\pi_t' \equiv \frac{1}{\phi_1}\left(\Delta p_t - \left(\overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n} + (\phi_0 + \phi_1 \rho)\,\Delta x_t + \phi_\chi \cdot \Delta \chi_t\right)\right) = u_t + \frac{\phi_n}{\phi_1}\varepsilon_t^n,$$

and price informativeness is

$$\tau_{\pi'} \equiv \mathbb{V}ar\left[\left.\pi'_t\right| u_t, \Delta x_t, \Delta\chi_t\right]^{-1} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_{\Delta n}.$$

Comparing the expression for the equilibrium price in Eq. (38) with Regression R1–PS below

$$\Delta p_t = \overline{\beta}' + \beta'_0 \Delta x_t + \beta'_1 \Delta x_{t+1} + \beta'_2 \cdot \Delta\chi_t + e'_t, \tag{R1–PS}$$

we have that $\beta'_1 = \phi_1$ and $\mathbb{V}ar\left[e'_t\right] = \phi_n^2 \tau_{\Delta n}^{-1}$. Therefore, absolute price informativeness can be recovered as $\tau_{\pi'} = \frac{(\beta'_1)^2}{\mathbb{V}ar[e'_t]}$.

Moreover, comparing Regression R2–PS below

$$\Delta p_t = \overline{\zeta}' + \zeta'_0 \Delta x_t + \zeta'_1 \cdot \Delta\chi_t + e_t^{\zeta'}, \tag{R2–PS}$$

with the expression for the price process in Equation (39) we have that $e_t^{\zeta'} \equiv \phi_1 u_t + \phi_n \varepsilon_t^n$. Taking the variance on both sides of Regression R2–PS gives

$$1 = \frac{\mathbb{V}ar_t\left[\zeta'_0 \Delta x_t + \zeta'_1 \cdot \Delta\chi_t\right]}{\mathbb{V}ar\left[p_t\right]} + \frac{\mathbb{V}ar\left[e'_t\right]}{\mathbb{V}ar\left[p_t\right]}\left(\underbrace{\frac{\mathbb{V}ar_t\left[\phi_1 u_t\right]}{\mathbb{V}ar\left[e'_t\right]}}_{=\frac{\tau_{\pi'}}{\tau_u}} + 1\right).$$

Noting that

$$\frac{\mathbb{V}ar\left[e'_t\right]}{\mathbb{V}ar\left[p_t\right]} = 1 - R^2_{\Delta x, \Delta x', \Delta\chi} \quad \text{and} \quad \frac{\mathbb{V}ar_t\left[\zeta'_0 \Delta x_t + \zeta'_1 \cdot \Delta\chi_t\right]}{\mathbb{V}ar\left[p_t\right]} = R^2_{\Delta x, \Delta\chi}$$

implies

$$\tau^R_{\pi'} \equiv \frac{\tau_{\pi'}}{\tau_u + \tau_{\pi'}} = \frac{R^2_{\Delta x, \Delta x', \Delta\chi} - R^2_{\Delta x, \Delta\chi}}{1 - R^2_{\Delta x, \Delta\chi}},$$

where $R^2_{\Delta x, \Delta x', \Delta\chi}$ and $R^2_{\Delta x, \Delta\chi}$ are the R-squareds of Regression R1–PS and Regression R2–PS, respectively.

□

A.3.3 Proof of Proposition 4 *(*Identifying price informativeness about learnable component of payoffs)

When there is an unlearnable component of payoffs, the portfolio demand of an investor i can be approximated by

$$\theta_t^i \approx \alpha_x^i x_t + \alpha_s^i s_t^i + \alpha_n^i \overline{n}_t^i - \alpha_p^i p_t + \psi^i,$$

the coefficients α_x^i, α_s^i, α_n^i, and α_p^i are positive scalars that represent the individual demand sensitivities to the contemporary payoff, the private signal, the private trading needs, and the asset price, respectively, and ψ^i can be positive or negative and incorporates the risk premium.

Using the market clearing condition with this approximation we have that the approximated equilibrium (log) price is

$$p_t = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_t^L + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} n_t + \frac{\overline{\psi}}{\overline{\alpha_p}},$$

where $\overline{\alpha_h} \equiv \int \alpha_h^i w_t^i di$ for $h = x, s, n, p$ and $\overline{\psi} \equiv \int \psi^i w_t^i di - Q$. Using that the payoff process is

$$\Delta x_{t+1} = \mu_{\Delta x} + u_t^U + u_t^L$$

the price can be written as

$$p_t = \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}} \right) x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} \left(x_{t+1} - u_t^U \right) + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} n_t + \frac{\overline{\psi} - \overline{\alpha_s} \mu_{\Delta x}}{\overline{\alpha_p}}$$

and taking first differences this becomes

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t - \phi_1 \Delta u_t^U,$$

where $\overline{\phi} = 0$, $\phi_0 = \frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$, $\phi_1 = \frac{\overline{\alpha_s}}{\overline{\alpha_p}}$ and $\phi_n = \frac{\overline{\alpha_n}}{\overline{\alpha_p}}$. This price process can also be written as

$$\Delta p_t = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_0 \Delta x_t + \phi_1 u_t^L + \phi_1 u_{t-1}^U + \phi_n \Delta n_t.$$

Then, the unbiased signal about the change in the learnable component of the innovation contained in the price is

$$\pi_t^L \equiv \frac{1}{\phi_1} \left(\Delta p_t - \left(\overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n} + \phi_0 \Delta x_t \right) \right) = u_t^L + u_{t-1}^U + \frac{\phi_n}{\phi_1} \varepsilon_t^n.$$

and absolute and relative price informativeness are respectively given by

$$\tau_\pi^L \equiv \mathbb{V}ar \left[u_{t-1}^U + \frac{\phi_n}{\phi_1} \varepsilon_t^n \middle| u_t^L, \Delta x_t \right]^{-1} = \left(\left(\tau_u^L + \tau_u^U \right)^{-1} + \left(\frac{\phi_n}{\phi_1} \right)^2 \tau_{\Delta n}^{-1} \right)^{-1} \tag{40}$$

and

$$\tau_\pi^{LR} = \left(\frac{\tau_u^L}{\tau_u^L + \tau_u^U} + \left(\frac{\phi_n}{\phi_1} \right)^2 \left(\frac{\tau_{\Delta n}}{\tau_u^L} \right)^{-1} \right)^{-1}.$$

When there is an unlearnable component of the payoff, the error term in regression R1 is given by $e_t^{\Delta x, \Delta x'} = \phi_n \varepsilon_t^n - \phi_1 \Delta u_t^U$ and the error term in Regression R2 is given by $e_t^{\Delta x} = \phi_n \varepsilon_t^n + \phi_1 u_t^L + \phi_1 u_{t-1}^U$. Then, we have that

$$\begin{aligned} \mathbb{V}\mathrm{ar} \left(e_t^{\Delta x, \Delta x'} \right) &= \mathbb{V}\mathrm{ar} \left(\phi_1 u_t^U \right) + \mathbb{V}\mathrm{ar} \left(\phi_1 u_{t-1}^U \right) + \mathbb{V}\mathrm{ar} \left(\phi_n \varepsilon_t^n \right) \\ \mathbb{V}\mathrm{ar} \left(e_t^{\Delta x} \right) &= \mathbb{V}\mathrm{ar} \left(\phi_1 u_t^L \right) + \mathbb{V}\mathrm{ar} \left(\phi_1 u_{t-1}^U \right) + \mathbb{V}\mathrm{ar} \left(\phi_n \varepsilon_t^n \right), \end{aligned}$$

and

$$\mathbb{V}\text{ar}\left(e_t^{\Delta x}\middle|\,\Delta x_t\right)=\mathbb{V}\text{ar}\left(\phi_1 u_t^L\right)+\mathbb{V}\text{ar}\left(\phi_1 u_{t-1}^U\middle|\,\Delta x_t\right)+\mathbb{V}\text{ar}\left(\phi_n \varepsilon_t^n\right).$$

First note that

$$\mathbb{V}\text{ar}\left(e_t^{\Delta x}\right)-\mathbb{V}\text{ar}\left(e_t^{\Delta x,\Delta x'}\right)=\mathbb{V}\text{ar}\left(\phi_1 u_t^L\right)-\mathbb{V}\text{ar}\left(\phi_1 u_t^U\right).$$

Then,

$$\begin{aligned}\frac{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\right)-\mathbb{V}\text{ar}\left(e_t^{\Delta x,\Delta x'}\right)}{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\middle|\,\Delta x_t\right)}&=\frac{\mathbb{V}\text{ar}\left(\phi_1 u_t^L\right)-\mathbb{V}\text{ar}\left(\phi_1 u_t^U\right)}{\mathbb{V}\text{ar}\left(\phi_1 u_t^L\right)+\mathbb{V}\text{ar}\left(\phi_1 u_{t-1}^U+\phi_n\varepsilon_t^n\middle|\,\Delta x_t\right)}\\&=\frac{1}{1+\frac{\mathbb{V}\text{ar}\left(u_{t-1}^U+\frac{\phi_n}{\phi_1}\varepsilon_t^n\middle|\Delta x_t\right)}{\mathbb{V}\text{ar}\left(u_t^L\right)}}\left(1-\frac{\mathbb{V}\text{ar}\left(u_t^U\right)}{\mathbb{V}\text{ar}\left(u_t^L\right)}\right)\\&=\frac{\tau_\pi^L}{\tau_\pi^L+\tau_u^L}\left(1-\frac{\mathbb{V}\text{ar}\left(u_t^U\right)}{\mathbb{V}\text{ar}\left(u_t^L\right)}\right).\end{aligned}$$

Rewriting this expression we have

$$\tau_\pi^{LR}=\frac{\tau_\pi^L}{\tau_\pi^L+\tau_u^L}=\frac{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\right)-\mathbb{V}\text{ar}\left(e_t^{\Delta x,\Delta x'}\right)}{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\middle|\,\Delta x_t\right)}\frac{1}{1-\frac{\mathbb{V}\text{ar}\left(u_t^U\right)}{\mathbb{V}\text{ar}\left(u_t^L\right)}}.$$

Finally, using that

$$\mathbb{V}\text{ar}\left(u_t\right)=\mathbb{V}\text{ar}\left(u_t^L\right)+\mathbb{V}\text{ar}\left(u_t^U\right)$$

and that

$$\frac{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\right)-\mathbb{V}\text{ar}\left(e_t^{\Delta x,\Delta x'}\right)}{\phi_1^2\mathbb{V}\text{ar}\left(u_t\right)}=\frac{\mathbb{V}\text{ar}\left(u_t^L\right)}{\mathbb{V}\text{ar}\left(u_t\right)}-\frac{\mathbb{V}\text{ar}\left(u_t^U\right)}{\mathbb{V}\text{ar}\left(u_t\right)}$$

we have

$$\frac{\mathbb{V}\text{ar}\left(u_t^U\right)}{\mathbb{V}\text{ar}\left(u_t^L\right)}=\frac{1-\frac{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\right)-\mathbb{V}\text{ar}\left(e_t^{\Delta x,\Delta x'}\right)}{\phi_1^2\mathbb{V}\text{ar}\left(u_t\right)}}{1+\frac{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\right)-\mathbb{V}\text{ar}\left(e_t^{\Delta x,\Delta x'}\right)}{\phi_1^2\mathbb{V}\text{ar}\left(u_t\right)}},$$

where $\mathbb{V}\text{ar}\left(u_t\right)$ is variance of the residuals in a regression of Δx_{t+1} on Δx_t.

Finally, note that

$$\frac{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\right)-\mathbb{V}\text{ar}\left(e_t^{\Delta x,\Delta x'}\right)}{1-\frac{\mathbb{V}\text{ar}\left(u_t^U\right)}{\mathbb{V}\text{ar}\left(u_t^L\right)}}=\frac{\phi_1^2\left(\mathbb{V}\text{ar}\left(u_t^L\right)-\mathbb{V}\text{ar}\left(u_t^U\right)\right)}{\mathbb{V}\text{ar}\left(u_t^L\right)-\mathbb{V}\text{ar}\left(u_t^U\right)}\mathbb{V}\text{ar}\left(u_t^L\right)=\phi_1^2\mathbb{V}\text{ar}\left(u_t^L\right).$$

Then, if $\mathbb{V}\text{ar}\left(u_t^L\right)=\mathbb{V}\text{ar}\left(u_t^U\right)$ we have that $\mathbb{V}\text{ar}\left(u_t\right)=2\mathbb{V}\text{ar}\left(u_t^L\right)$ and relative price informativeness can be recovered as

$$\tau_{\pi}^{LR} = \frac{\frac{\phi_1^2}{2}\mathbb{V}\text{ar}\left(u_t\right)}{\mathbb{V}\text{ar}\left(e_t^{\Delta x}\middle|\,\Delta x_t\right)},$$

which proves our results.

$\square$

A.3.4 Proof of Proposition 5 (Identifying price informativeness about future prices)

The equilibrium price process is given by

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta x_{t+1} + \phi_n \Delta n_t.$$

Iterating this expression forward we have

$$\begin{aligned}\Delta p_{t+1} &= \overline{\phi} + \phi_0 \Delta x_{t+1} + \phi_1 \Delta x_{t+2} + \phi_n \Delta n_{t+1} \\ &= \overline{\phi} + \phi_1 \Delta \mu_x + \left(\phi_0 + \rho\phi_1\right) \Delta x_{t+1} + \phi_1 u_{t+1} + \phi_n \Delta n_{t+1},\end{aligned} \tag{41}$$

where we substituted for the process for Δx_{t+2}. This last equation implies that

$$\Delta x_{t+1} = \frac{\Delta p_{t+1} - \left(\overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_1 u_{t+1} + \phi_n \Delta n_{t+1}\right)}{\phi_0 + \rho\phi_1}.$$

Using this expression in (41) we can express the current price difference in terms of the future price difference, as follows

$$\Delta p_t = \overline{\phi} - \frac{\phi_1}{\phi_0 + \rho\phi_1}\left(\overline{\phi} + \phi_1 \mu_{\Delta x}\right) + \phi_0 \Delta x_t + \frac{\phi_1}{\phi_0 + \rho\phi_1}\Delta p_{t+1} - \frac{\phi_1}{\phi_0 + \rho\phi_1}\left(\phi_1 u_{t+1} + \phi_n \Delta n_{t+1}\right) + \phi_n \Delta n_t. \tag{42}$$

The unbiased signal contained in the price about the next period's price is

$$\tilde{\pi}_t = \frac{\Delta p_t - \left(\overline{\phi} + \phi_n \mu_{\Delta n} - \frac{\phi_1}{\phi_0 + \rho\phi_1}\left(\overline{\phi} + \phi_n \mu_{\Delta n} + \phi_1 \mu_{\Delta x}\right) + \phi_0 \Delta x_t\right)}{\frac{\phi_1}{\phi + \rho\phi_1}} = \Delta p_{t+1} - \left(\phi_1 u_{t+1} + \phi_n \varepsilon_{t+1}^n\right) + \frac{\phi_n}{\frac{\phi_1}{\phi + \rho\phi_1}}\varepsilon_t^n$$

Absolute price informativeness is given by

$$\tau_{\tilde{\pi}} = \mathbb{V}ar\left[\tilde{\pi}_t\middle|\,x_t, \Delta p_{t+1}\right]^{-1} = \left(\mathbb{V}\text{ar}\left(\phi_1 u_{t+1} + \phi_n \varepsilon_{t+1}^n\right) + \left(\frac{\phi_0 + \rho\phi_1}{\phi_1}\right)^2 \mathbb{V}\text{ar}\left(\phi_n \varepsilon_{t+1}^n\right)\right)^{-1} \tag{43}$$

and relative price informativeness is

$$\tau_{\tilde{\pi}}^R = \frac{\tau_{\hat{\pi}}}{\tau_{\hat{\pi}} + \tau_{\Delta p}}.$$

From the structural mapping of the coefficients in Regressions R1–PS and R2–PS it follows that $e_t = \phi_n \varepsilon_t^n$, $e_t^\zeta = \phi_1 u_{t+1} + \phi_n \varepsilon_t^n$. Moreover, since innovations to the noise are i.i.d. it follows that

$$\mathbb{V}\mathrm{ar}\left(\phi_1 u_{t+1} + \phi_n \varepsilon_{t+1}^n\right) = \mathbb{V}\mathrm{ar}\left(\phi_1 u_{t+1} + \phi_n \varepsilon_t^n\right).$$

Moreover, $\beta_1 = \phi_1$ and $\zeta_0 = \phi_0 + \rho\phi_1$. Then,

$$\tau_{\hat{\pi}} = \left(\mathbb{V}ar\left[e_t^\zeta\right] + \left(\frac{\zeta_0}{\beta_1}\right)^2 \mathbb{V}ar\left[e_t\right]\right)^{-1}. \tag{44}$$

Using Equation (44), we have that

$$\frac{\tau_{\Delta p}}{\tau_{\hat{\pi}}} = \frac{\mathbb{V}ar\left[e_t^\zeta\right] + \left(\frac{\zeta_0}{\beta_1}\right)^2 \mathbb{V}ar\left[e_t\right]}{\mathbb{V}ar\left[\Delta p\right]}.$$

Then, since $\frac{\mathbb{V}ar\left[e_t^\zeta\right]}{\mathbb{V}ar[\Delta p]} = 1 - R^2_{\Delta x}$ and $\frac{\mathbb{V}ar[e_t]}{\mathbb{V}ar[\Delta p]} = 1 - R^2_{\Delta x, \Delta x'}$, it follows that

$$\tau_{\hat{\pi}}^R = \frac{1}{1 + \left(1 - R^2_{\Delta x} + \left(\frac{\zeta_0}{\beta_1}\right)^2 \left(1 - R^2_{\Delta x, \Delta x'}\right)\right)}.$$

□

A.3.5 Simulation: Learnable and unlearnable payoff

In this section, we simulate the model with an unlearnable payoff component studied in Section 4.3. This simulation is helpful to understand the sign and magnitude of the potential biased caused by estimating the model via OLS and using an incorrect identification procedure.

Figure (6) shows the distribution of two different estimates of relative price informativeness in an environment in which the innovation to the asset payoff has a learnable and an unlearnable component, as described in Section 4.3. The red line shows the mean estimate of price informativeness estimated using Proposition 1 (OLS estimation of Regressions R1 and R2). The blue line shows the mean estimate of price informativeness estimated using Proposition 4 (IV estimation of Regressions R1 and R2, using u_t^L as instrument for Regression R2, and u_t^L and u_{t-1}^L as instruments for Regression R1. The dashed areas report the 5% and 95% of the distribution of estimates for each estimator. In this simulation, the estimates of price informativeness using Proposition 1 are roughly 20% lower than the true value of $\tau_\pi^L = 0.127$.

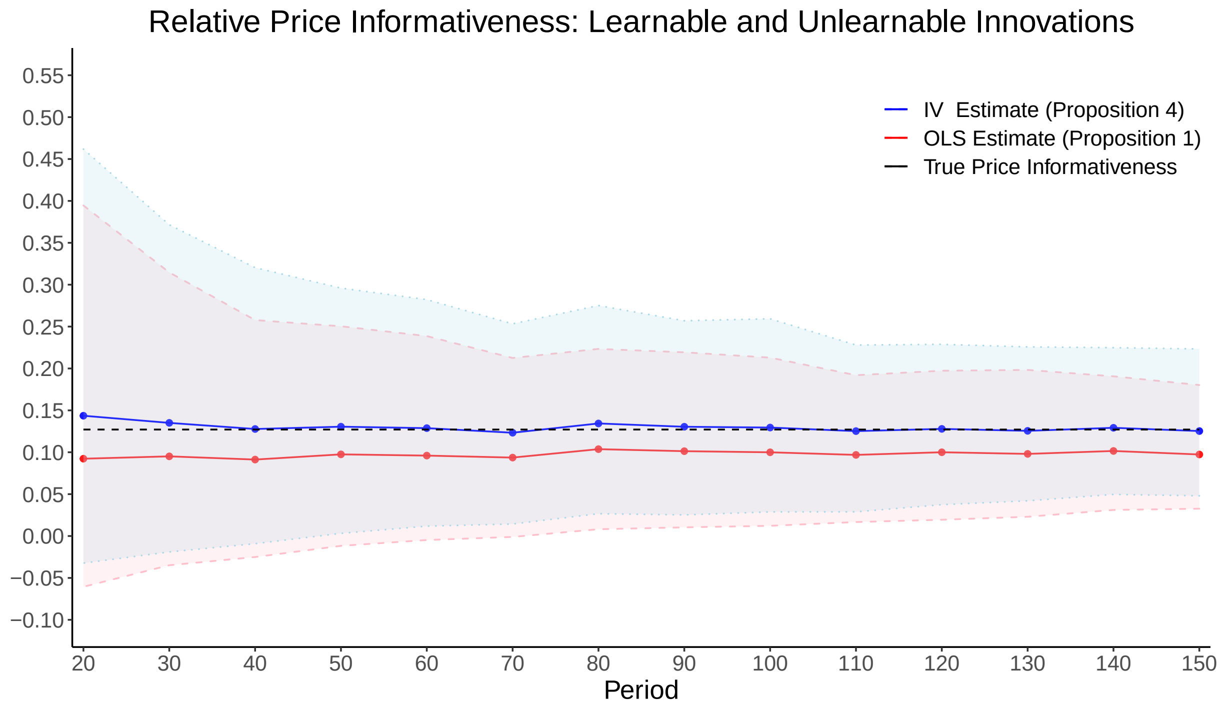

Figure 6: Price informativeness: estimates with learnable and unlearnable payoff innovations

Note: Figure 6 shows the distribution of two different estimates of relative price informativeness in an environment in which the innovation to the asset payoff has a learnable and an unlearnable component, as described in Section 4.3. We simulate the model of Section 4.3 with parameters: $\mu_{\Delta x} = \mu_{\Delta n} = 0$, $\tau_u^L = 2$, $\tau_u^U = 8$, $\tau_n = 0.3$, and $\overline{\phi} = 0$, $\phi_0 = 1$, $\phi_1 = 1$, $\phi_n = 1$. For every number of periods between 20 to 150, in intervals of 10, we simulate the model $N = 500$ times, and report the mean estimate and the 5% and 95% estimates of price informativeness. The dashed black line shows the true value of relative price informativeness. The red line shows the mean estimate of price informativeness estimated using Proposition 1 (OLS estimation of Regressions R1 and R2). The blue line shows the mean estimate of price informativeness estimated using Proposition 4 (IV estimation of Regressions R1 and R2, using u_t^L as instrument for Regression R2, and u_t^L and u_{t-1}^L as instruments for Regression R1. The dashed areas report the 5% and 95% of the distribution of estimates for each estimator.

Online Appendix

A Detailed Data Description

This Appendix describes in more detail the data used for the empirical implementation of the results in Section 5. See https://github.com/edavila/identifying_price_informativeness for additional details and replicating files.

We obtain stock market price data from the Center for Research in Security Prices (CRSP) for the time period between January 1, 1950 and December 31, 2019. First, we import monthly price data from the Monthly Stock File (msf) for ordinary common shares (shrcd = 10 or 11). Second, we import delisting prices and other delisting information from the monthly delisting file (msedelist). Third, we import market-returns from the monthly stock indicators file (msi). Lastly, we import the start and/or end date(s) of when a stock has been part of the S&P 500 from dsp500list. We restrict our sample to securities listed on the NYSE, AMEX or the Nasdaq (exchcd = 1, 2, or 3). We compute market capitalization by multiplying the stock price by the number of shares outstanding. For companies with multiple securities, we sum the market cap for all the company's securities and keep only the permno with the highest market capitalization. We define turnover as the ratio of volume to shares outstanding.

From FRED, we download monthly time series for Personal Consumption Expenditure (PCEPI), 1-Year and 10-Year Treasury Rates (GS1, GS10), Unemployment Rate (UNRATE), Personal Consumption Expenditures (PCE) and Personal Income (PI).

We import firm performance data from both the COMPUSTAT Fundamentals Annual Data & the Fundamentals Quarterly Data in the standard, consolidated, industrial format for domestic firms (INDFMT='INDL' and DATAFMT='STD' and CONSOL='C' and POPSRC='D') for observations between January 1, 1950 and December 31, 2019 . For future linking with CRSP, we also import GVKEYs and permnos from the CRSP/COMPUSTAT Merged (CCM) database, keeping the following linktypes: "LU", "LC", or "LS", and for which the the issue marker is primary (linkprim = "P" or "C"). For both the annual and quarterly data, we only keep observations where the observation date is between the beginning and end of the period for which the CCM link is valid.

We form book value (book) for annual[quarterly] data as shareholders equity (seq[q]) + deferred assets plus investment tax credit (txditc[q]) - pstk[q] (preferred stock). To deal with missing values, we replace seqq with common equity plus preferred equity (ceq[q] + pstk[q]) if an observation of the former but not the latter is unavailable, and if both of are unavailable, we replace with total assets minus total liability (at[q] - lt[q]). If an observation for txditc[q] and/or pstk[q] are missing, we replace it with 0. We respectively use oiadpq and ebit as our payoff measures in the quarterly and annual datasets. We merge for both the annual and quarterly datasets, where in the shifted specifications, we shift CRSP respectively one quarter and one month back. After merging the COMPUSTAT and CRSP datasets using the timing describing in the text, we use PCEPI to

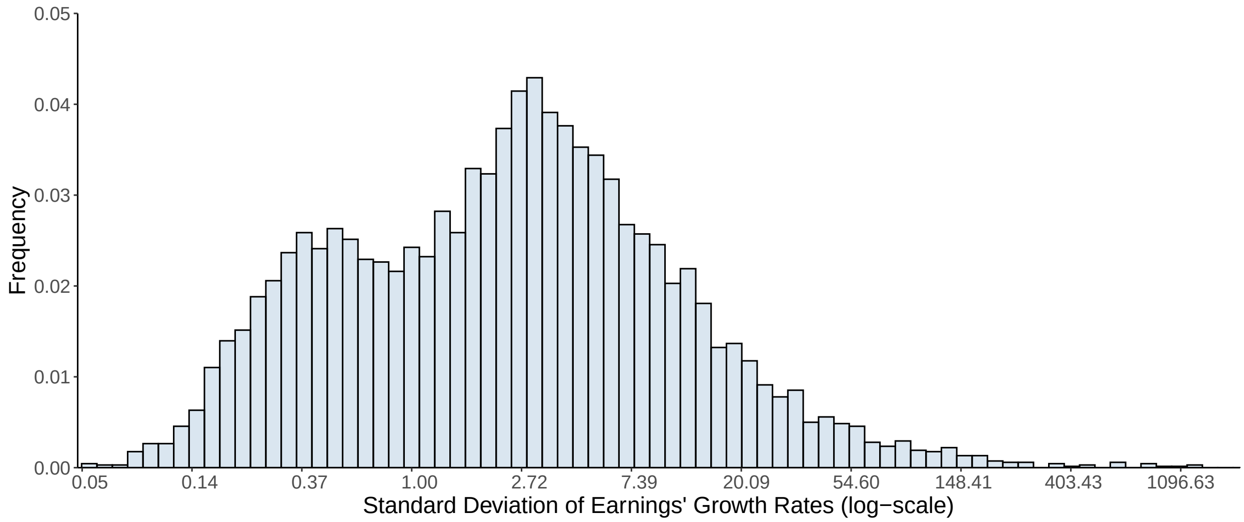

Figure OA-1: Cross-sectional standard deviation of earnings' growth rates

Note: Figure OA-1 shows a relative frequency histogram of the time series standard deviation of earnings growth rates across stocks. This histogram includes features 6,803 stocks. The median and mean of the average growth rate of earnings across the stocks represented in this figure are, respectively, 0.11 and 0.56.

deflate all nominal variables. We also discard stocks with non-finite prices and whose payoff is always 0 or NA. We winsorize payoff and price values at the 2.5th and 97.5th percentile to reduce the impact of outliers. We compute growth rates of payoffs as follows. When the lagged payoff is positive, the growth rate is defined as ${}^{\text{payoff}}/_{\text{payoff}_{t-1}} - 1$. When the lagged payoff is negative, the growth rate is defined as ${}^{\text{payoff}_t}/_{|\text{payoff}_{t-1}|} + 1$.

Figure OA-1 illustrates the distribution of stock-specific standard deviation of quarterly earnings' growth rates in our sample of stocks with more than 40 observations. As one would expect, the volatility of earnings across stocks varies widely in the cross-section.

B Empirical Implementation: Additional Results

In this section, we report additional empirical results. First, we include a series of figures that give more insight into the cross-sectional results presented in Table 2. Second, we present the results of the model after included a set of controls (public signals). Third, we present cross-sectional relations the cross-sectional relation that we identify remains valid Finally, we present the results using annual observations, instead of quarterly.

B.1 Cross-sectional relation: graphical illustration

Figures OA-2 through OA-6 are the counterpart of the cross-sectional results presented in Table 2. Each figure shows scatter plots of cross-sectional regressions of relative price informativeness (in twentiles) on each of the five variables considered: size, value, turnover, return volatility, and institutional ownership, for each of the years between 1981 and 2016. This figures made clear that

Table OA-1: Persistence of price informativeness

t	$\rho_{t,t-1}(\tau_\pi^R)$	$\rho_{t,t-2}(\tau_\pi^R)$	$\rho_{t,t-3}(\tau_\pi^R)$	$\rho_{t,t-4}(\tau_\pi^R)$	$\rho_{t,t-5}(\tau_\pi^R)$
1980	0.7162	0.5848	0.6299	0.6018	0.4854
1981	0.7696	0.5089	0.3970	0.4810	0.4214
1982	0.6976	0.4993	0.3399	0.2513	0.2822
1983	0.8285	0.5573	0.3901	0.2248	0.1394
1984	0.7581	0.5134	0.3121	0.2140	0.1018
1985	0.7577	0.4746	0.3303	0.1458	0.1094
1986	0.7731	0.5329	0.3546	0.2753	0.1027
1987	0.7929	0.5735	0.3927	0.2323	0.1649
1988	0.8298	0.5230	0.3810	0.2613	0.1356
1989	0.9018	0.6937	0.4481	0.3234	0.2383
1990	0.7838	0.6388	0.4819	0.3031	0.2054
1991	0.8398	0.5827	0.4774	0.3697	0.2488
1992	0.8647	0.6744	0.4716	0.3876	0.2865
1993	0.8722	0.7161	0.5791	0.4141	0.3456
1994	0.8811	0.7202	0.5955	0.4805	0.3422
1995	0.8925	0.7570	0.6221	0.5167	0.4430
1996	0.8778	0.7371	0.6380	0.5251	0.4553
1997	0.8381	0.7030	0.5770	0.4938	0.4117
1998	0.8043	0.5836	0.5015	0.3743	0.2972
1999	0.8404	0.6195	0.4403	0.3736	0.2664
2000	0.8244	0.6071	0.4514	0.3339	0.2697
2001	0.7947	0.5857	0.4339	0.3180	0.2351
2002	0.7786	0.5606	0.4154	0.3047	0.2192
2003	0.8686	0.6461	0.4659	0.3431	0.2602
2004	0.8811	0.7221	0.5450	0.3989	0.2963
2005	0.8991	0.7370	0.6050	0.4689	0.3459
2006	0.8980	0.7921	0.6563	0.5412	0.4078
2007	0.8608	0.7368	0.6504	0.5276	0.4331
2008	0.7145	0.5745	0.4834	0.4176	0.3533
2009	0.7714	0.4768	0.3793	0.3285	0.2860
2010	0.8953	0.6313	0.3740	0.3063	0.2762
2011	0.8967	0.7690	0.5111	0.2746	0.1945
2012	0.9152	0.7970	0.6963	0.4358	0.1963
2013	0.9214	0.8173	0.7285	0.6360	0.3573
2014	0.9346	0.8278	0.7182	0.6353	0.5505
2015	0.9119	0.8455	0.7582	0.6530	0.5889
2016	0.9296	0.8309	0.7734	0.7102	0.6136
2017	0.9395	0.8508	0.7664	0.7154	0.6610

Note: Table OA-1 reports the cross-sectional correlation between price informativeness measured in year t and price informativeness measures in year $t-k$, where $k = \{1, 2, 3, 4, 5\}$. Since our panel of price informativeness is quarterly, we average the measures of quarterly price informativeness at the yearly level before computing the correlations. We start reporting the correlations in 1980, since that is the first year with more than 250 stocks. Informativeness in year t is computed over a rolling window of 30 quarters (7.5 years) prior.

Table OA-2: Cross-sectional results, public signals

	Estimate	Std. Error	t-stat
Size	0.0023	0.000261	8.80
Value	-0.0011	0.000715	-1.58
Turnover	0.0003	0.000029	10.31
Idiosyncratic Volatility	0.0193	0.010162	1.90
Institutional Ownership	0.0171	0.002100	8.12

Note: Table OA-2 reports the estimates ($\widehat{a_1^c}$) of panel regressions of price informativeness on cross-sectional characteristics (in twentiles) with year-fixed effects (ξ_t): $\tau_{\pi,t}^{R,b} = a_0^c + a_1^c c_t^b + \xi_t + \epsilon_{b,t.}$, where $\tau_{\pi}^{R,b,t}$ denotes the average price informative per bin (twentile) in a given period, c_t^b denotes the value of the given characteristic per bin (twentile) in a given period, ξ_t denotes a year-fixed effect, a_0^c and a_1^c are parameters, and $\epsilon_{b,t}$ is an error term. Figures OA-2 through OA-6 provide the graphical counterpart of the results in this table. Size is measured as the natural log of stocks market capitalization, value is measured as the ratio between a stock's book value and its market capitalization, turnover is measured as the ratio of trading volume to shares outstanding, idiosyncratic volatility is measured as the standard deviation — over a 30 month period — of the difference between the returns of a stock and the market return, and institutional ownership is measured as the proportion of stocks held by institutional investors.

the positive relations between price informativeness and size, turnover, and institutional ownership are robust across time.

B.2 Public signals/additional controls

While in the body of the paper we report measures of price informativeness that do not include additional controls – public signals, in the language of Section 4.2 – in this section we report the results once we include several controls. Formally, the results reported here are the outcome of running the following two regressions

$$\Delta p_t^j = \overline{\beta}^j + \beta_0^j \Delta x_t^j + \beta_1^j \Delta x_{t+1}^j + \beta^c \cdot \Delta w_t^{j,q} + d_t^{j,q} + \varepsilon_t^j \quad \Rightarrow R^{2,j}_{\Delta x, \Delta x'} \tag{45}$$

$$\Delta p_t^j = \overline{\zeta}^j + \zeta_0^j \Delta x_t^j + \beta^c \cdot \Delta w_t^{j,q} + d_t^{j,q} + \hat{\varepsilon}_t^j \quad \Rightarrow R^{2,j}_{\Delta x}\ , \tag{46}$$

where $w_t^{j,q}$ denotes a given set of controls/public signal. The results reported here use the following aggregate variables are controls for all stocks: i) changes in the one- and ten-year treasury rates, ii) changes in unemployment rates, iii) change in log consumption, and iv) changes in log income. We obtain similar results using stock-specific controls too.

Figure OA-7, which is the counterpart of Figure 2, shows a relative frequency histogram of price informativeness for a representative time period, the last quarter of 2015. Figure OA-8, which is the counterpart of Figure 5, shows the time series evolution of the mean, median, and standard deviation of relative price informativeness. Table OA-2 and Figures OA-9 and OA-10, which are the counterparts of Table 2 and Figures 3 and 4, show the cross-sectional properties of the distribution of price informativeness across stocks.

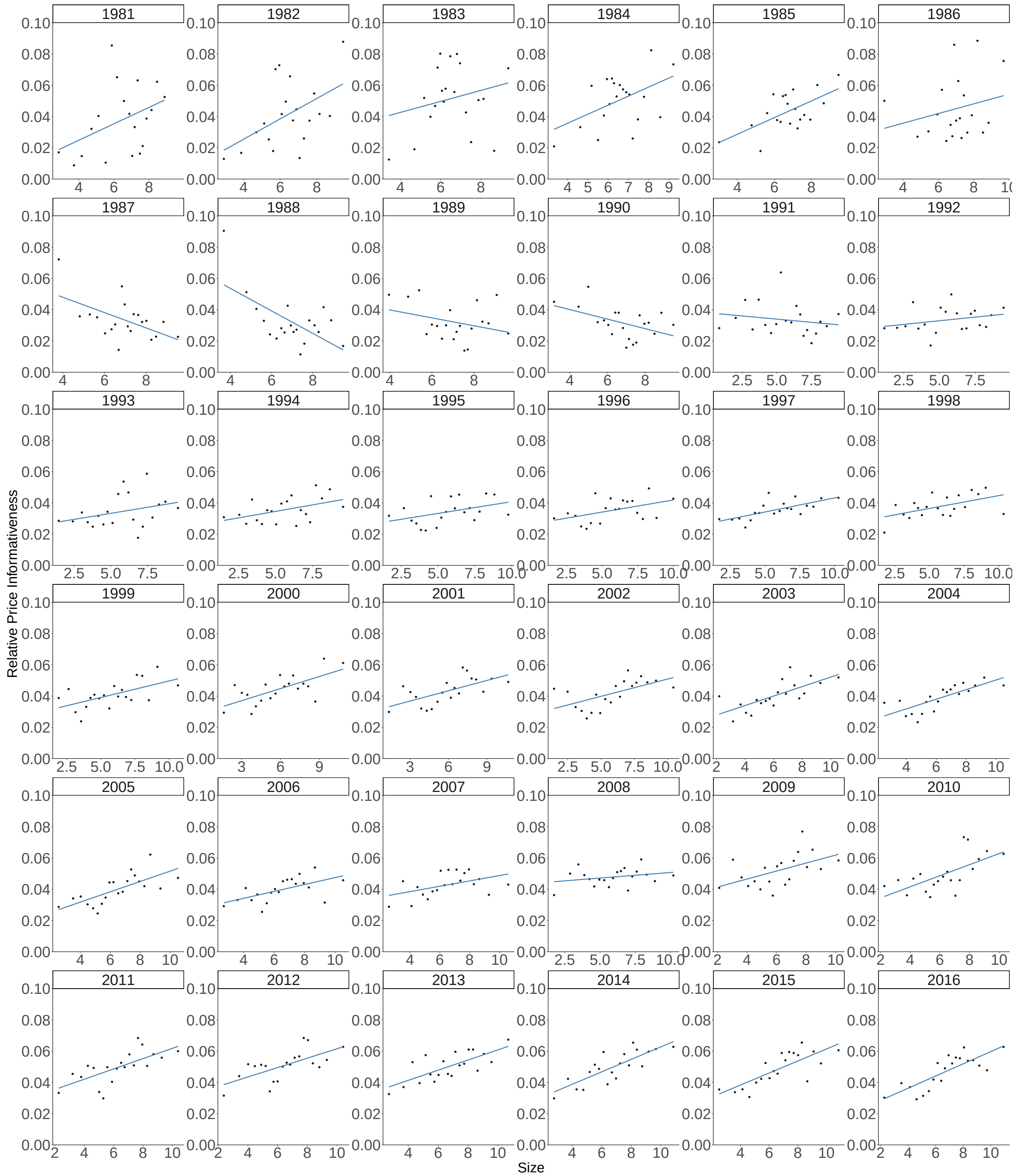

Figure OA-2: Price informativeness and size

Note: Figure OA-2 shows year-by-year cross-sectional regressions of relative price informativeness (in twentiles) on size, defined as the log of market capitalization – see e.g. Bali, Engle and Murray (2016). The estimate reported in in Table 2 can be interpreted as a weighted averaged of the year-by-year slope coefficient illustrated here.

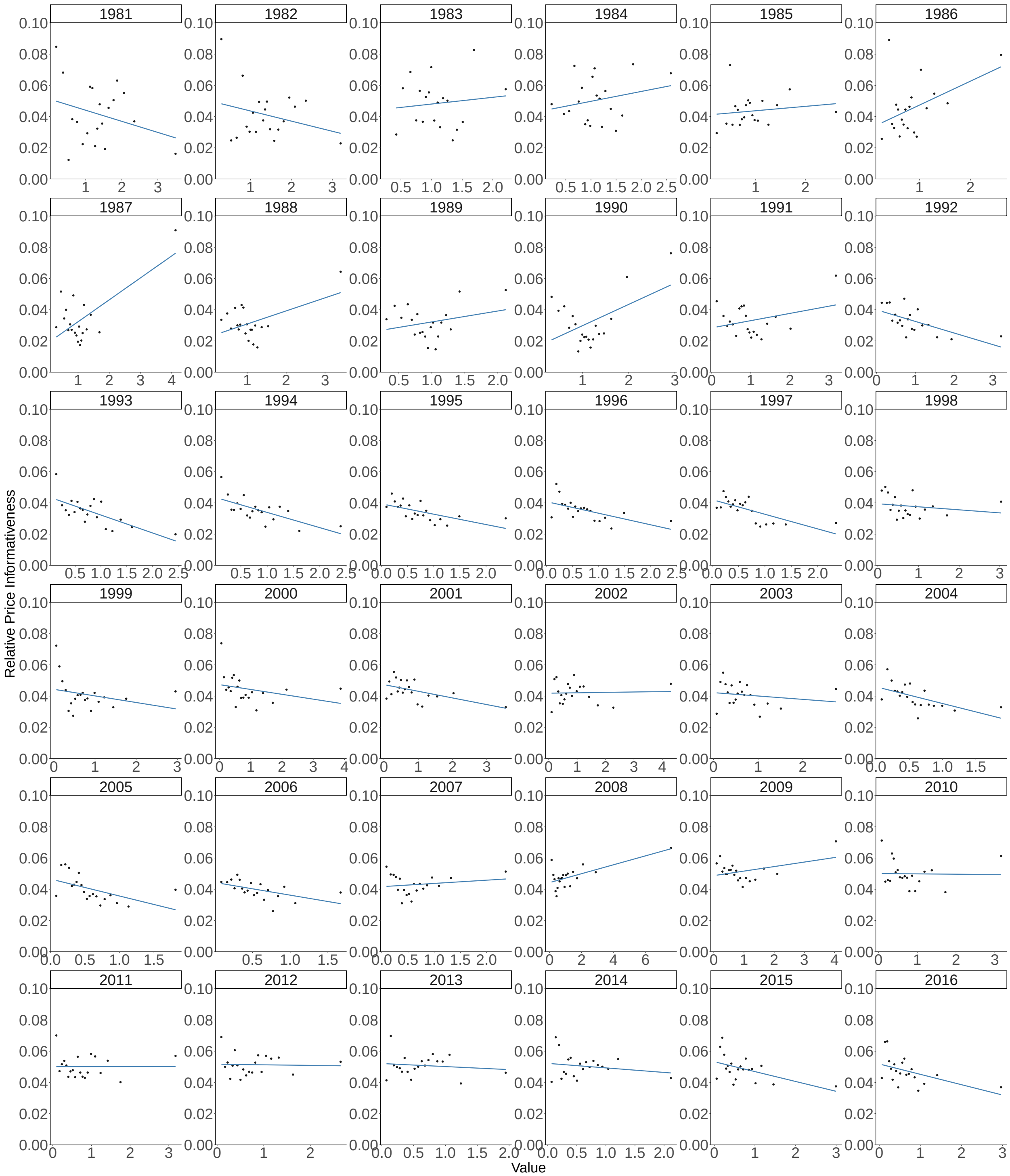

Figure OA-3: Price informativeness and value

Note: Figure OA-3 shows year-by-year cross-sectional regressions of relative price informativeness (in twentiles) on value, defined as the ratio between a stock's book value and its market capitalization. The estimate reported in in Table 2 can be interpreted as a weighted averaged of the year-by-year slope coefficient illustrated here.

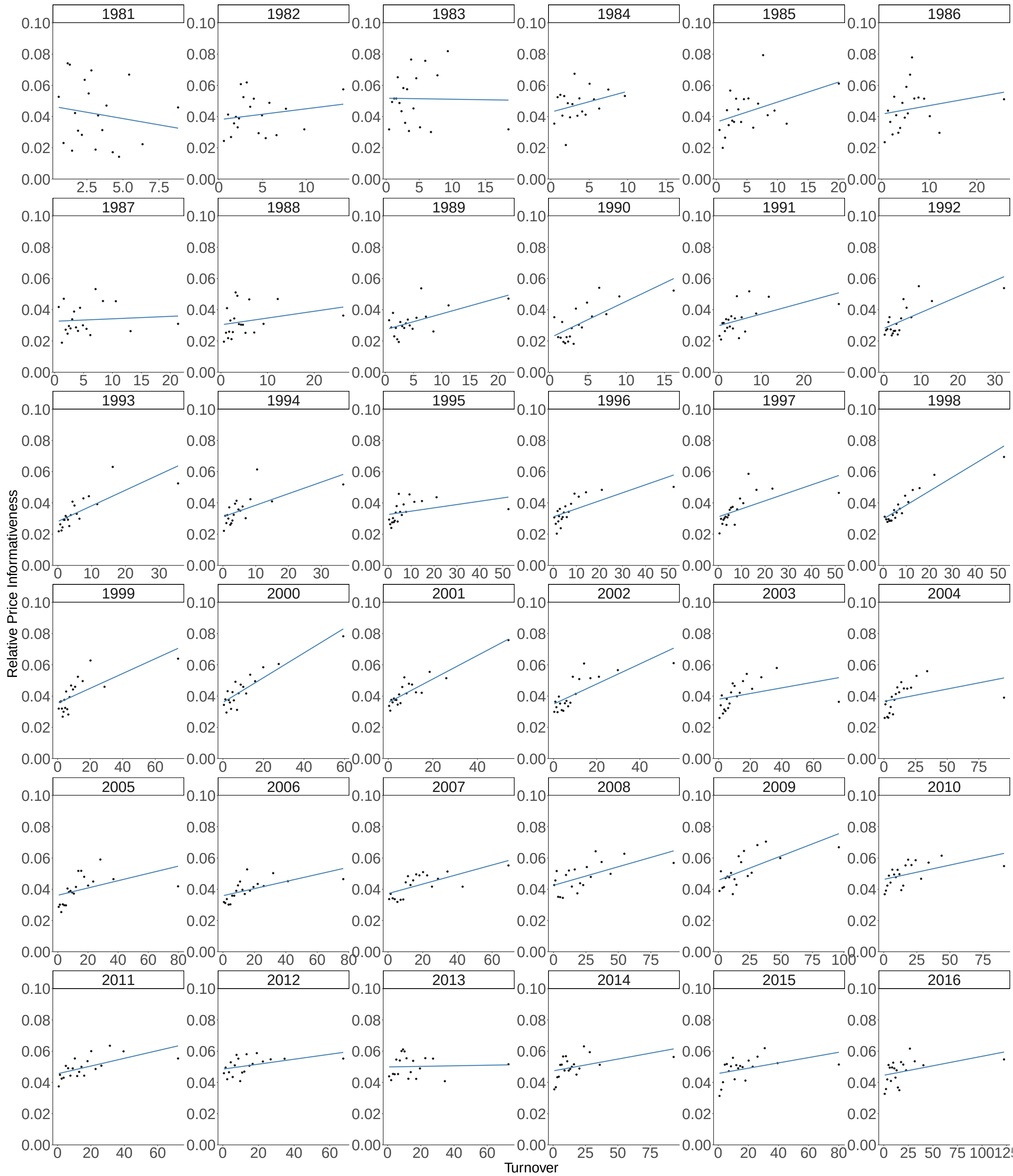

Figure OA-4: Price informativeness and turnover

Note: Figure OA-4 shows year-by-year cross-sectional regressions of relative price informativeness (in twentiles) on turnover, defined as the ratio of trading volume to shares outstanding. The estimate reported in Table 2 can be interpreted as a weighted averaged of the year-by-year slope coefficient illustrated here.

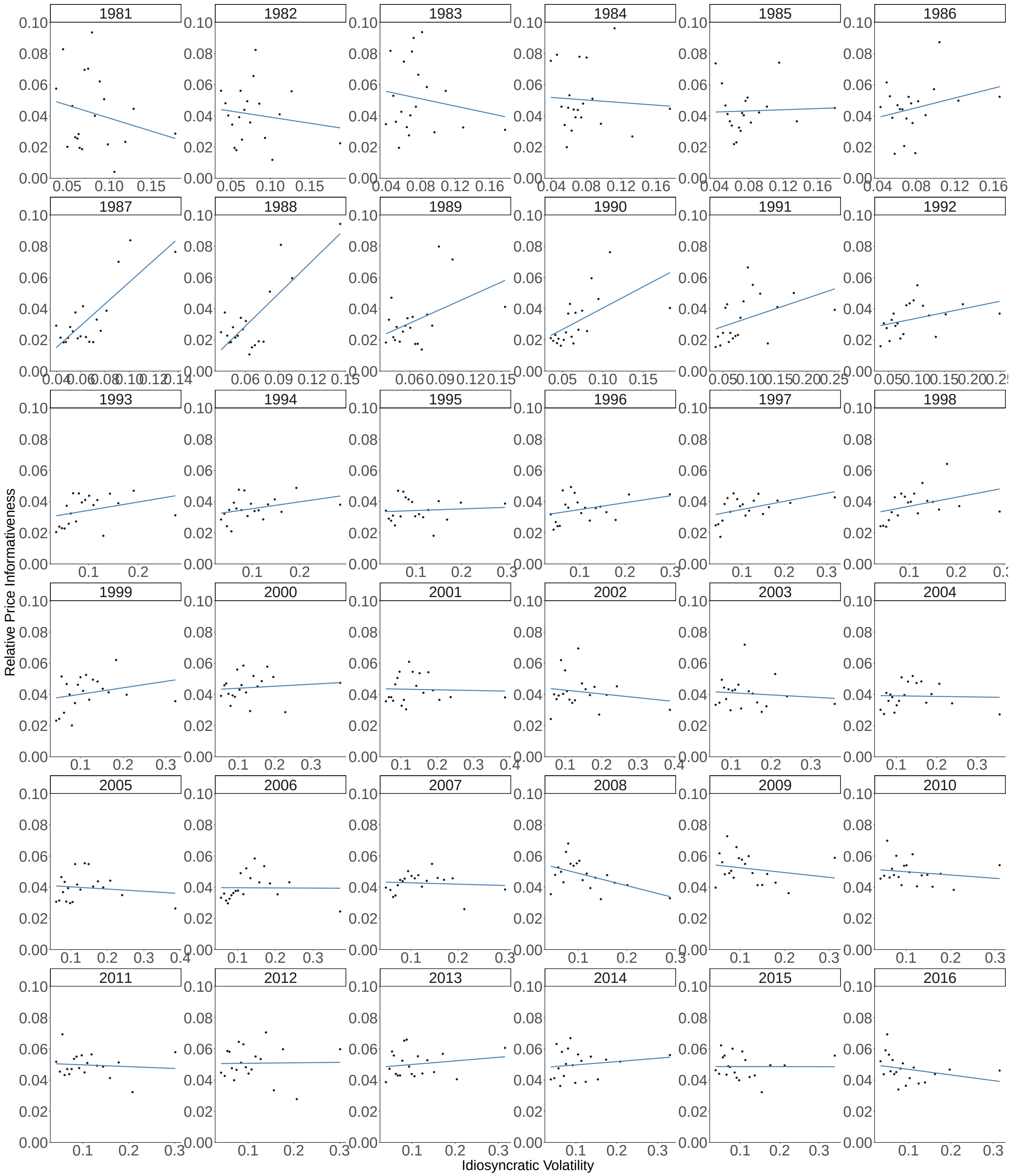

Figure OA-5: Price informativeness and idiosyncratic return volatility

Note: Figure OA-5 shows year-by-year cross-sectional regressions of relative price informativeness (in twentiles) on idiosyncratic volatility, define as the standard deviation over a 30 month period of the difference between the returns of a stock and the market return. The estimate reported in Table 2 can be interpreted as a weighted averaged of the year-by-year slope coefficient illustrated here.

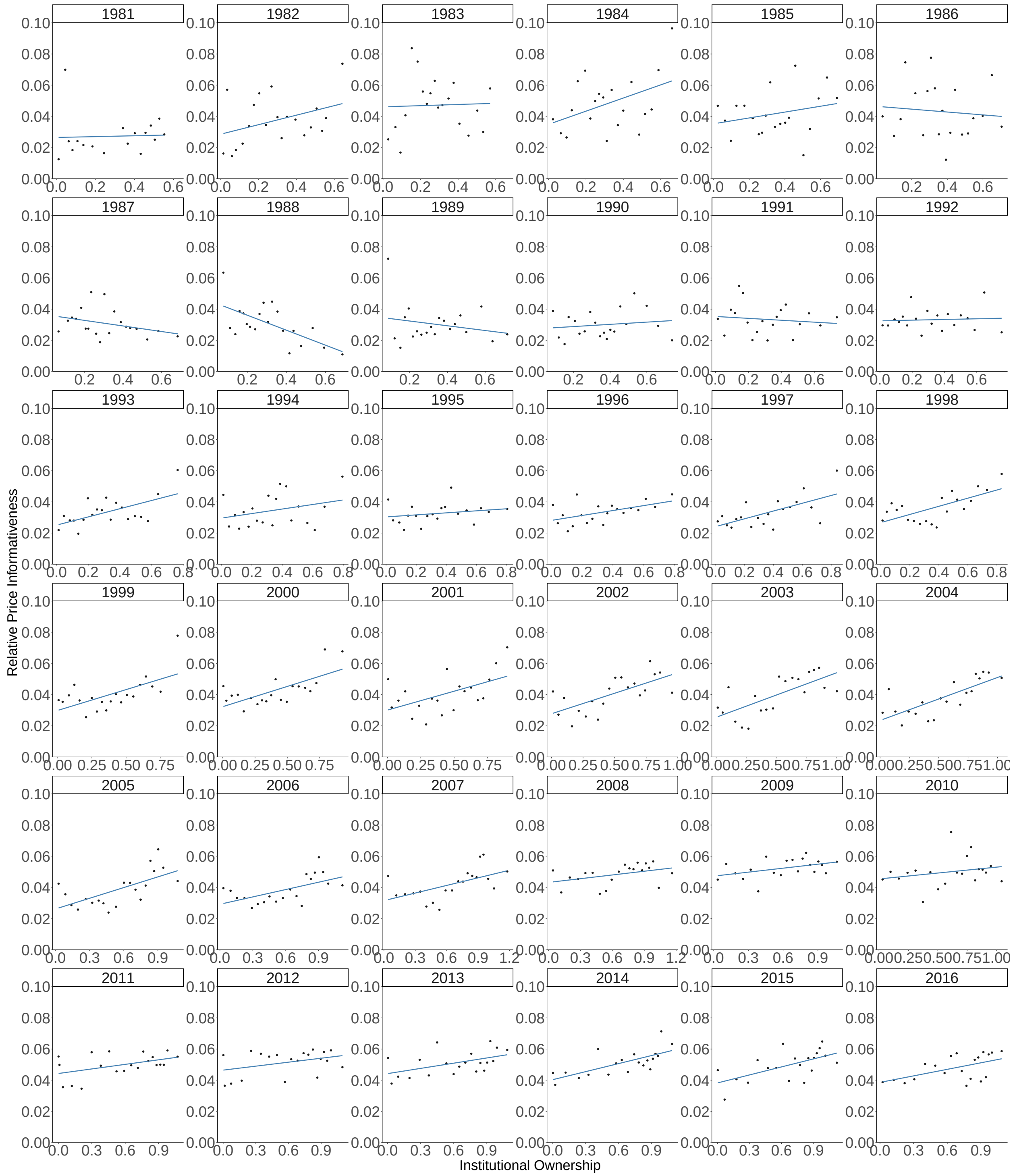

Figure OA-6: Price informativeness and institutional ownership

Note: Figure OA-6 shows year-by-year cross-sectional regressions of relative price informativeness (in twentiles) on institutional ownership, defined as the proportion of stocks held by institutional investors. The estimate reported in Table 2 can be interpreted as a weighted averaged of the year-by-year slope coefficient illustrated here.

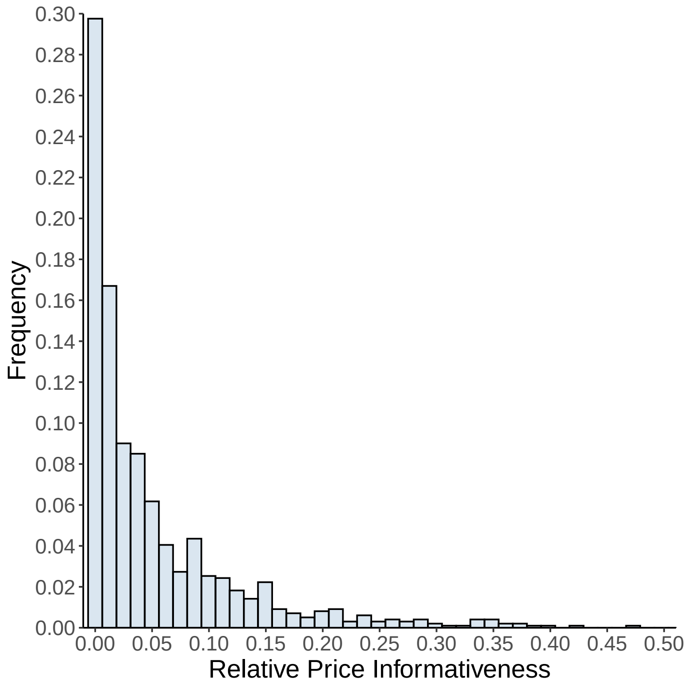

Figure OA-7: Price informativeness: relative frequency histogram, public signals

Note: Figure OA-7 shows a relative frequency histogram of price informativeness for a representative time period, the last quarter of 2015. Note that informativeness is computed over a rolling window of 30 quarters (7.5 years) prior.

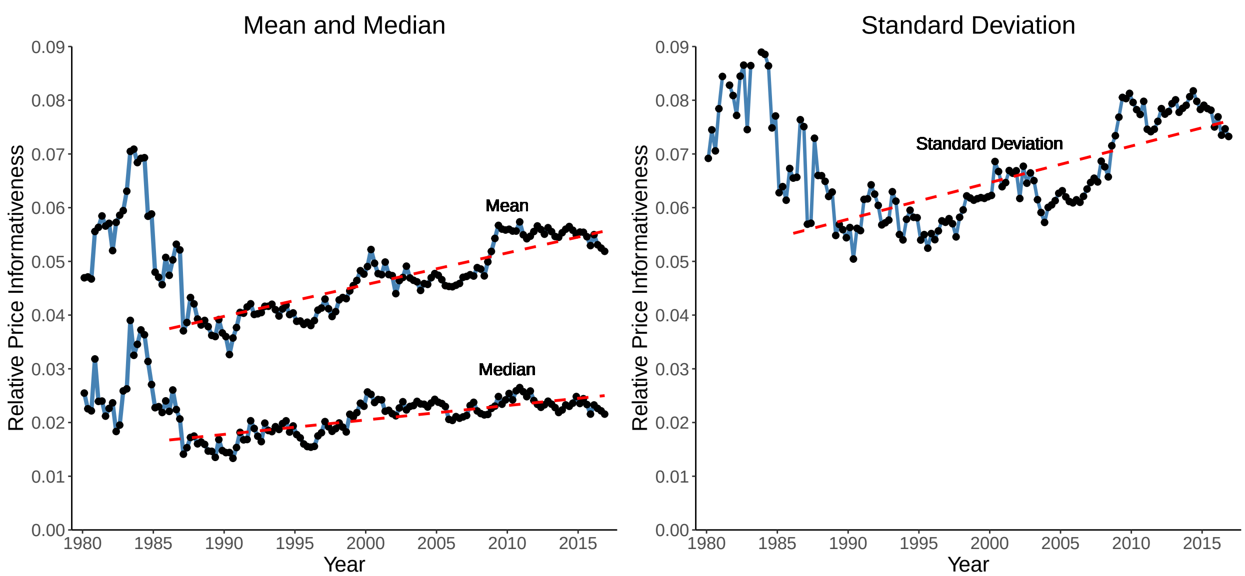

Figure OA-8: Price informativeness over time, public signals

Note: The left panel in Figure OA-8 shows the time series evolution of the mean and median relative price informativeness. The right panel in Figure OA-8 shows the time series evolution of the standard deviation of price informativeness. The red dashed lines show linear trends starting in 1986.

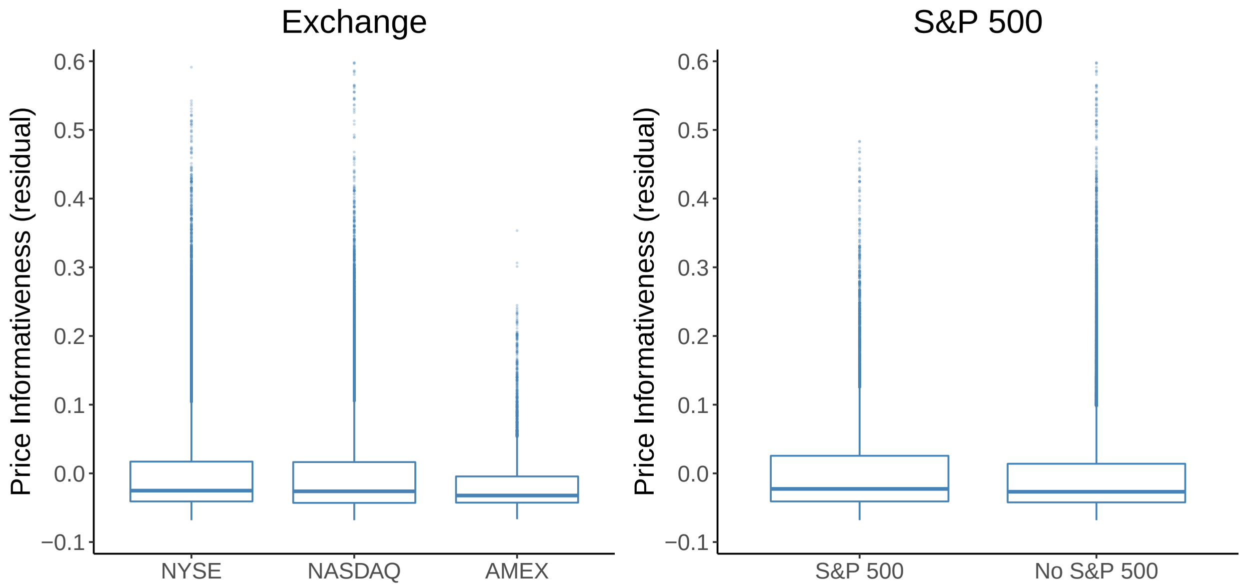

Figure OA-9: Cross-sectional results, public signals

Note: The left panel in Figure OA-9 shows a box plot by exchange of the residuals of a regression of relative price informativeness on year-fixed effects. The left panel in Figure OA-9 shows a box plot by S&P 500 status of the residuals of a regression of relative price informativeness on year-fixed effects. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to ±1.5 times the interquartile range.

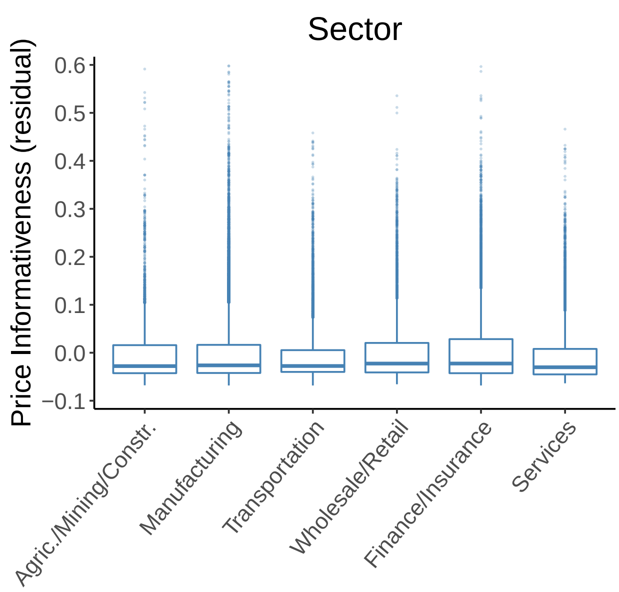

Figure OA-10: Cross-sectional results, public signals

Note: Figure OA-10 shows a box plot by one-digit SIC industry code of the residuals of a regression of relative price informativeness on year-fixed effects. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to 1.5 times the interquartile range. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to ±1.5 times the interquartile range.

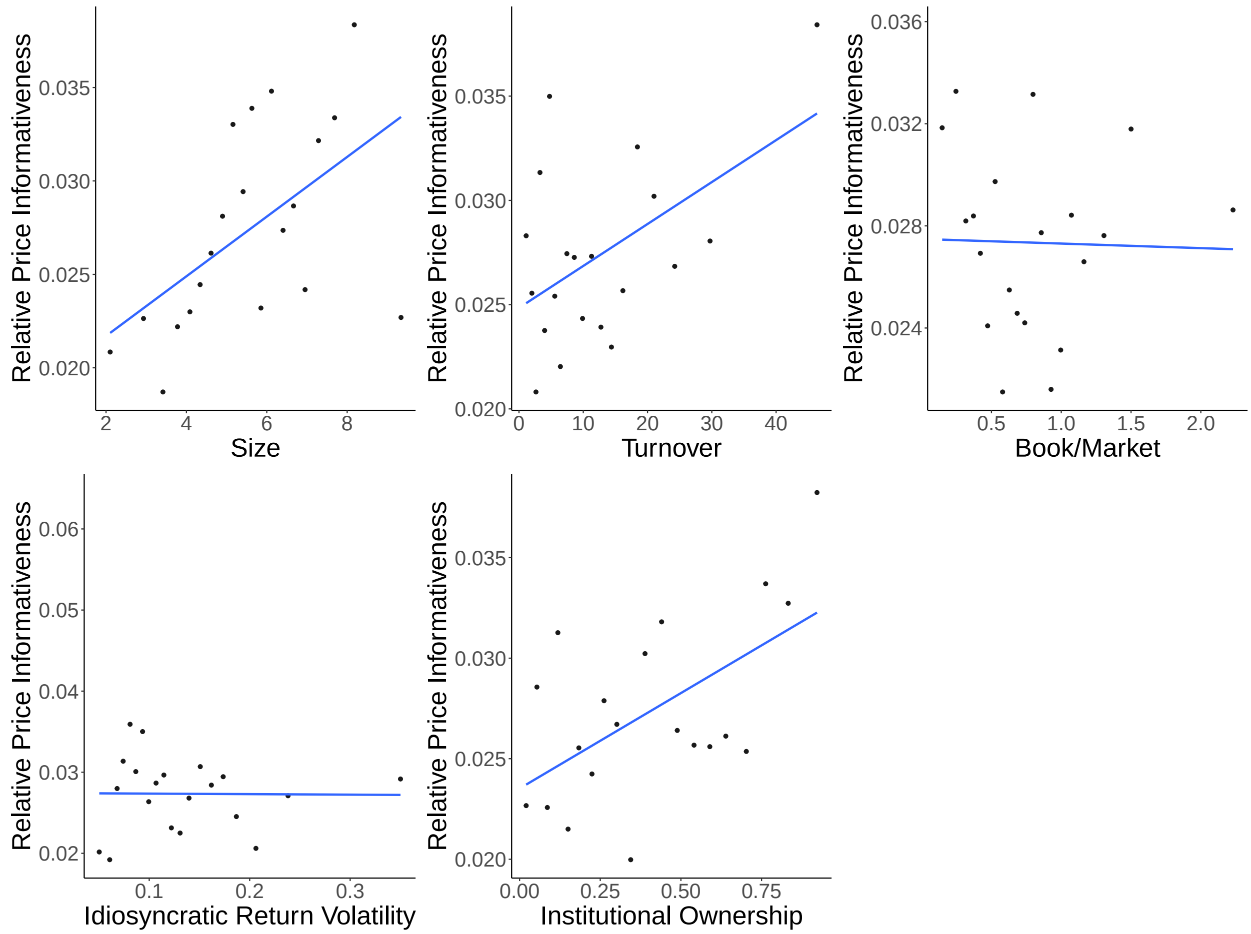

Figure OA-11: Price informativeness: cross sectional results, largest time series

Note: Figure OA-11 shows cross-sectional regressions of relative price informativeness (in twentiles) on size, value, turnover, return volatility, and institutional ownership. The estimate reported in Table 2 can be interpreted as a weighted averaged of the year-by-year slope coefficient illustrated here.

B.3 Price informativeness in the cross-section: full sample results

In the body of the paper, we compute price informativeness using rolling windows of 40 quarters. Alternatively, we can compute a single measure of price informativeness for each stock using the largest possible of time series data available for each. In this case, we obtain a single price informativeness measure per stock. The upside of this approach is that it uses all the available information for a given stock. The main drawback of this approach is that the recovered informativeness measures rely on observations from different time periods. Figure OA-11 shows the cross-sectional results established in Table 2 remain valid in this case.

B.4 Annual observations

In this section, we report results when using annual data, instead of quarterly data. These results show that both the cross-sectional and the time-series findings identified using quarterly observa-

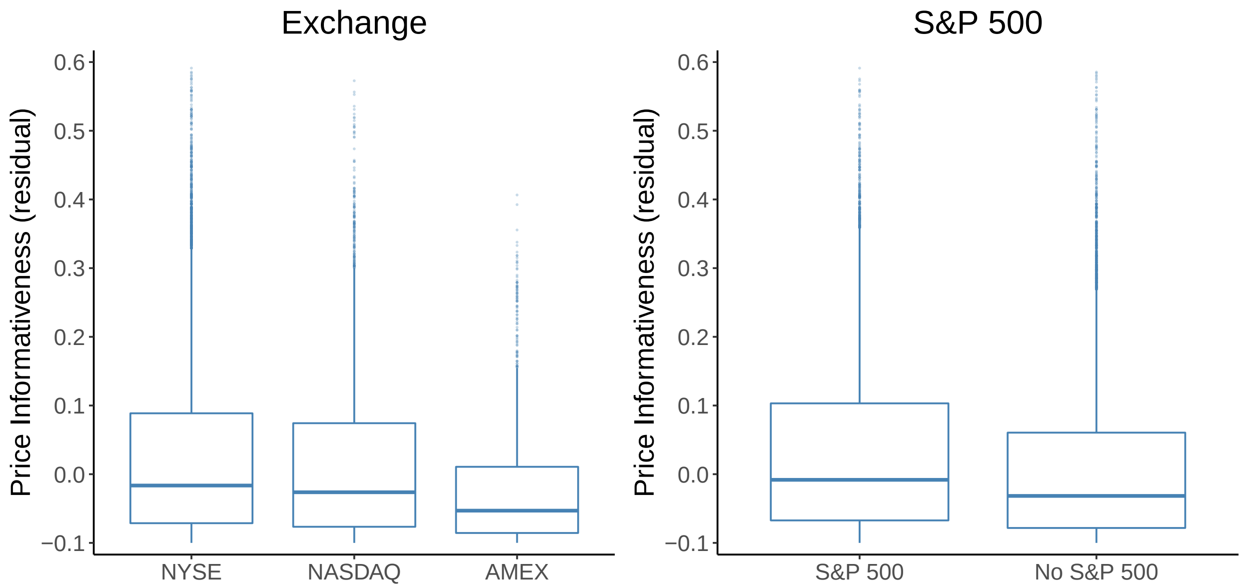

Figure OA-12: Cross-sectional results, annual data

Note: The left panel in Figure OA-12 shows a box plot by exchange of the residuals of a regression of relative price informativeness on year-fixed effects. The left panel in Figure OA-12 shows a box plot by S&P 500 status of the residuals of a regression of relative price informativeness on year-fixed effects. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to ±1.5 times the interquartile range.

tions remain true when using annual data.

Table OA-3: Price informativeness: year-by-year summary statistics, annual data

t	$Median$	$Mean$	SD	$Skew$	$Kurt$	$P5$	$P25$	$P75$	$P95$	n
1980	0.0605	0.1005	0.1098	1.3136	1.0969	0.0005	0.0122	0.1561	0.3468	325
1981	0.0805	0.1183	0.1163	1.0880	0.5231	0.0007	0.0206	0.1897	0.3450	330
1982	0.0862	0.1205	0.1224	1.3648	1.6106	0.0010	0.0204	0.1810	0.3660	498
1983	0.0745	0.1127	0.1205	1.6009	2.6929	0.0007	0.0198	0.1618	0.3550	548
1984	0.0700	0.1065	0.1155	1.5872	2.4329	0.0009	0.0188	0.1479	0.3690	549
1985	0.0615	0.1000	0.1068	1.4595	1.7157	0.0005	0.0181	0.1471	0.3355	559
1986	0.0517	0.0896	0.0996	1.4745	1.7431	0.0004	0.0130	0.1341	0.2973	645
1987	0.0503	0.0884	0.0994	1.5413	2.0074	0.0003	0.0152	0.1315	0.3056	649
1988	0.0475	0.0852	0.1008	1.7193	3.0337	0.0005	0.0130	0.1223	0.3153	665
1989	0.0471	0.0845	0.0996	1.7150	3.0903	0.0005	0.0108	0.1222	0.2977	703
1990	0.0485	0.0879	0.1024	1.7527	3.6143	0.0007	0.0119	0.1260	0.3040	728
1991	0.0484	0.0869	0.1014	1.6711	2.7390	0.0006	0.0114	0.1272	0.3125	758
1992	0.0490	0.0876	0.1024	1.6822	2.6015	0.0004	0.0118	0.1257	0.3263	929
1993	0.0504	0.0871	0.1011	1.6797	2.8568	0.0005	0.0121	0.1267	0.3046	1084
1994	0.0483	0.0875	0.1026	1.6647	2.4477	0.0006	0.0118	0.1226	0.3116	1116
1995	0.0505	0.0912	0.1051	1.5758	2.0894	0.0007	0.0120	0.1346	0.3320	1106
1996	0.0491	0.0916	0.1070	1.6750	2.8497	0.0007	0.0116	0.1352	0.3321	1058
1997	0.0514	0.0912	0.1057	1.7379	3.2153	0.0010	0.0134	0.1340	0.3100	1034
1998	0.0491	0.0904	0.1064	1.7439	3.2574	0.0005	0.0115	0.1355	0.3195	985
1999	0.0462	0.0911	0.1123	1.8568	3.7130	0.0006	0.0103	0.1294	0.3393	946
2000	0.0525	0.1008	0.1216	1.7250	2.8762	0.0004	0.0126	0.1454	0.3629	915
2001	0.0522	0.0968	0.1160	1.8303	3.7779	0.0004	0.0121	0.1378	0.3491	916
2002	0.0502	0.0900	0.1084	1.7981	3.3087	0.0004	0.0100	0.1287	0.3285	926
2003	0.0522	0.0945	0.1098	1.6024	2.4414	0.0004	0.0110	0.1349	0.3258	966
2004	0.0577	0.0960	0.1109	1.5923	2.5432	0.0004	0.0110	0.1394	0.3273	997
2005	0.0533	0.0919	0.1075	1.5791	2.3983	0.0004	0.0105	0.1343	0.3234	1004
2006	0.0532	0.0922	0.1068	1.5606	2.2134	0.0005	0.0110	0.1376	0.3256	1017
2007	0.0549	0.1001	0.1171	1.6268	2.5102	0.0004	0.0121	0.1501	0.3564	1039
2008	0.0602	0.1164	0.1363	1.5380	1.9849	0.0005	0.0134	0.1752	0.3960	1043
2009	0.0658	0.1262	0.1491	1.6213	2.4922	0.0004	0.0165	0.1832	0.4391	1017
2010	0.0640	0.1247	0.1480	1.6283	2.5649	0.0006	0.0161	0.1901	0.4374	1009
2011	0.0645	0.1279	0.1506	1.5369	2.1180	0.0005	0.0149	0.1983	0.4403	1028
2012	0.0630	0.1258	0.1508	1.5881	2.3045	0.0006	0.0140	0.1914	0.4402	1062
2013	0.0635	0.1236	0.1479	1.5607	2.1880	0.0003	0.0142	0.1941	0.4405	1191
2014	0.0669	0.1281	0.1494	1.4830	1.8592	0.0006	0.0147	0.1985	0.4397	1223
2015	0.0696	0.1286	0.1472	1.4709	1.8840	0.0006	0.0143	0.1984	0.4426	1239
2016	0.0697	0.1286	0.1491	1.4858	1.8313	0.0008	0.0159	0.1918	0.4407	1270

Note: Table OA-3 reports year-by-year summary statistics on the panel of price informativeness measures recovered. It provides information on the median, mean, standard deviation, skewness, excess kurtosis, 5th, 25th, 75th, and 95th percentiles of each yearly distribution, as well as the number of stocks in each year. Since our panel of price informativeness is quarterly, we average the measures of quarterly price informativeness at the yearly level before computing the summary statistics. We start reporting summary statistics in 1980, since that is the first year with more than 250 stocks. Informativeness in year t is computed over a rolling window of 30 quarters (7.5 years) prior.

Table OA-4: Cross-sectional results, annual data

	Estimate	Std. Error	t-stat
Size	0.00873	0.000445	19.63
Value	-0.01197	0.001613	-7.42
Turnover	0.00096	0.000083	11.54
Idiosyncratic Volatility	-0.22194	0.019394	-11.44
Institutional Ownership	0.07366	0.004254	17.32

Note: Table OA-4 reports the estimates ($\widehat{a_1^c}$) of panel regressions of price informativeness on cross-sectional characteristics (in twentiles) with year-fixed effects (ξ_t): $\tau^{R,b}_{\pi,t} = a^c_0 + a^c_1 c^b_t + \xi_t + \epsilon_{b,t}.$, where $\tau^{R,b,t}_{\pi}$ denotes the average price informative per bin (twentile) in a given period, c^b_t denotes the value of the given characteristic per bin (twentile) in a given period, ξ_t denotes a year-fixed effect, a^c_0 and a^c_1 are parameters, and $\epsilon_{b,t}$ is an error term. Figures OA-2 through OA-6 provide the graphical counterpart of the results in this table. Size is measured as the natural log of stocks market capitalization, value is measured as the ratio between a stock's book value and its market capitalization, turnover is measured as the ratio of trading volume to shares outstanding, idiosyncratic volatility is measured as the standard deviation — over a 30 month period — of the difference between the returns of a stock and the market return, and institutional ownership is measured as the proportion of stocks held by institutional investors.

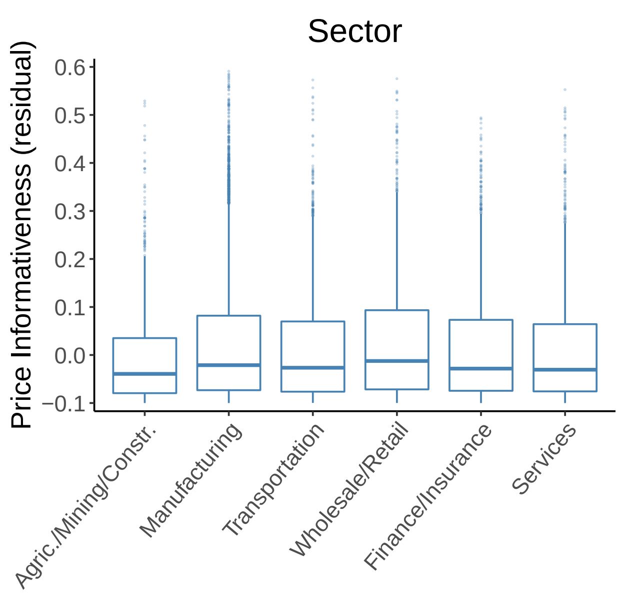

Figure OA-13: Cross-sectional results, annual data

Note: Figure OA-13 shows a box plot by one-digit SIC industry code of the residuals of a regression of relative price informativeness on year-fixed effects. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to 1.5 times the interquartile range. The solid middle line represents the median. The top and bottom of the box represent the 75% and 25% percentiles. The whiskers extend up to ±1.5 times the interquartile range.

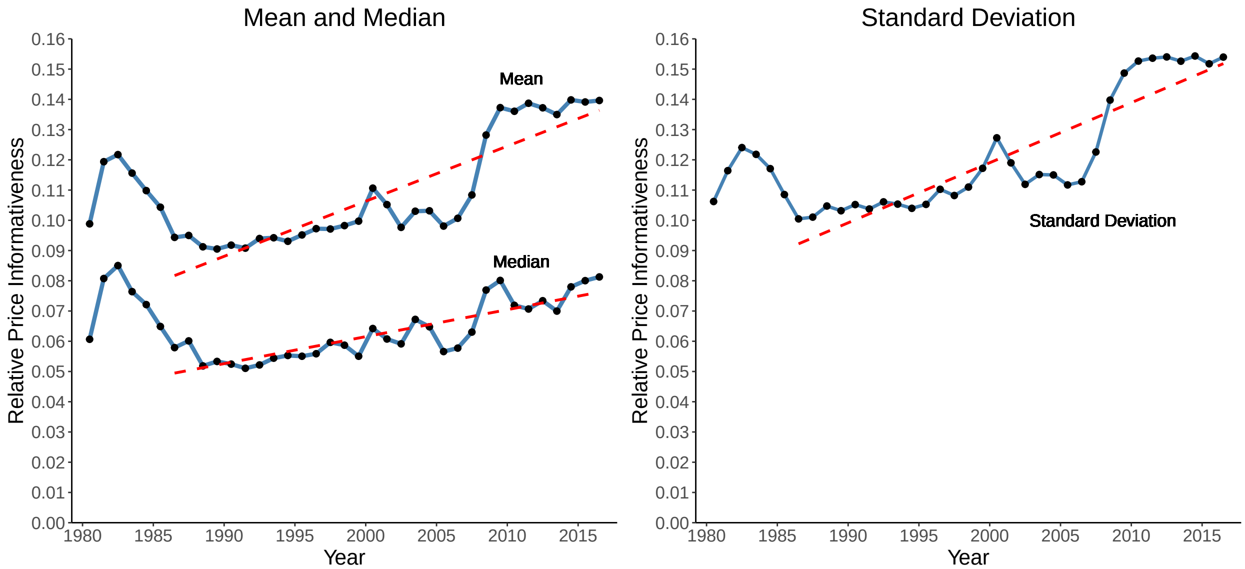

Figure OA-14: Price informativeness over time, annual data

Note: The left panel in Figure OA-14 shows the time series evolution of the mean and median relative price informativeness. The right panel in Figure OA-14 shows the time series evolution of the standard deviation of price informativeness. The red dashed lines show linear trends starting in 1986. In both panels, the dots correspond to the average within a year of the price informativeness measures computed using annual data.

C Alternative Modeling Frameworks

Our identification methodology extends to any linear or log-linear setup. In this section, we illustrate how to extend our results in the context of three different specifications. First, we extend our approximate results to the case in which the payoff follows a stationary AR(1) process. Second, we develop our identification results using an exact linear formulation for the price process under difference stationary and stationary specifications for the payoff. We also provide the respective CARA-Normal models to microfound these exact linear formulations.

C.1 Log-Linear model in levels

General framework and identification

We consider a discrete time environment with dates $t = 0, 1, 2, \ldots, \infty$, in which investors trade a risky asset in fixed supply at a (log) price p_t at each date t. We assume that the (log) payoff of the risky asset at date $t+1$, x_{t+1}, follows a stationary AR(1) process

$$x_{t+1} = \mu_x + \rho x_t + u_t, \tag{47}$$

where μ_x is a scalar, $|\rho| < 1$, and where the innovations to the payoff, u_t, have mean zero, a finite variance denoted by $\mathbb{V}ar\left[u_t\right] = \sigma_u^2 = \tau_u^{-1}$, and are identically and independently distributed over time.[11] We assume that the equilibrium price is given by

$$p_t = \overline{\phi} + \phi_0 x_t + \phi_1 x_{t+1} + \phi_n n_t, \tag{48}$$

where $\overline{\phi}$, ϕ_0, ϕ_1, and ϕ_n are parameters and where n_t represents the aggregate component of investors' trading motives that are orthogonal to the asset payoff, given by $n_t = \mu_n + \varepsilon_t^n$, where $\mathbb{E}\left[\varepsilon_t^n\right]$ and $\mathbb{V}ar\left[\varepsilon_t^n\right] = \sigma_n^2 = \tau_n^{-1}$. For simplicity, we assume that u_t and n_t are independent.

In this environment, the unbiased signal of the innovation to future payoffs u_t contained in the price, which we denote by $\tilde{\pi}_t$, is given by

$$\tilde{\pi}_t \equiv \frac{p_t - \left(\overline{\phi} + \phi_1 \mu_x + \phi_n \mu_n + (\phi_0 + \rho\phi_1)\, x_t\right)}{\phi_1} = u_t + \frac{\phi_n}{\phi_1}\left(n_t - \mu_n\right)$$

and absolute and relative price informativeness are respectively given by

$$\tau_{\tilde{\pi}} \equiv \left(\mathbb{V}\mathrm{ar}\left[\tilde{\pi}_t \middle| x_{t+1}, x_t\right]\right)^{-1} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_n \quad \text{and} \quad \tau_{\tilde{\pi}}^R \equiv \frac{\tau_{\tilde{\pi}}}{\tau_{\tilde{\pi}} + \tau_u}.$$

Proposition 7. (Identifying price informativeness log-linear)

a) Absolute price informativeness. Let $\overline{\beta}$, β_0, and β_1 denote the coefficients of the following regression of log-prices on realized and future log-payoffs,

$$p_t = \overline{\beta} + \beta_0 x_t + \beta_1 x_{t+1} + e_t, \tag{R1-LL}$$

where p_t denotes the date t log-price, x_t and x_{t+1} respectively denote the dates t and $t+1$ log-payoff, and where $\sigma_e^2 = \mathbb{V}ar[e_t]$ denotes the variance of the error. Then, absolute price informativeness, τ_π, can be recovered by

$$\tau_{\tilde{\pi}} = \frac{\beta_1^2}{\sigma_e^2}. \tag{49}$$

The OLS estimation of Regression R1-LL yields consistent estimates of β_1 and σ_e^2.

b) Relative Price Informativeness. Let $R_{x,x'}^2$ denote the R-squared of Regression R1-LL. Let R_x^2, ζ, and ζ_0 respectively denote the R-squared and the coefficients of the following regression of log-price on log-payoff,

$$p_t = \overline{\zeta} + \zeta_0 x_t + e_t^{\zeta}. \tag{R2-LL}$$

Then, relative price informativeness, $\tau_{\tilde{\pi}}^R$, can be recovered by

$$\tau_{\tilde{\pi}}^R = \frac{R_{x,x'}^2 - R_x^2}{1 - R_x^2}. \tag{50}$$

The OLS estimation of Regressions R1-LL and R2-LL yields consistent estimates of $R_{x,x'}^2$ and R_x^2.

Proof. a) By comparing Regression R1-LL with the structural Equation (48), it follows that $\overline{\beta} = \overline{\phi} +$

[11] As in the body of the paper, we index the innovation to the date $t+1$ payoff u_t by t — instead of $t+1$ — because investors may be able to learn about it at date t.

$\phi_n\mu_n$, $\beta_0 = \phi_0$, $\beta_1 = \phi_1$, and $e_t = \phi_n\varepsilon_t^n$. Consequently, $\sigma_e^2 = \mathbb{V}\mathrm{ar}\left[e_t\right] = (\phi_n)^2\,\mathbb{V}\mathrm{ar}\left[\varepsilon_t^n\right] = (\phi_n)^2\,\tau_n^{-1}$. Therefore, we can recover absolute price informativeness as follows

$$\tau_{\tilde{\pi}} = \frac{(\beta_1)^2}{\sigma_e^2} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_n.$$

Given Equations (47) and (48), as well as the assumptions on u_t and n_t, it is straightforward to show that the OLS estimates of Regressions R1-LL and R2-LL are consistent, which implies that price informativeness can be consistently estimated as $\widehat{\tau_{\tilde{\pi}}} = \frac{\left(\widehat{\beta_1}\right)^2}{\widehat{\sigma_e^2}}$. Formally, $\operatorname{plim}\left(\widehat{\tau_{\tilde{\pi}}}\right) = \operatorname{plim}\left(\frac{\left(\widehat{\beta_1}\right)^2}{\widehat{\sigma_e^2}}\right) = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_n = \tau_\pi$.

b) Note that the R-squareds of Regressions R1-LL and R2-LL can be expressed as follows

$$R^2_{x,x'} = 1 - \frac{\mathbb{V}\mathrm{ar}\left(e_t\right)}{\mathbb{V}\mathrm{ar}\left(p_t\right)} \quad \text{and} \quad R^2_{\Delta x} = \frac{\mathbb{V}\mathrm{ar}\left(\zeta_0 x_t\right)}{\mathbb{V}\mathrm{ar}\left(p_t\right)}.$$

After substituting Equation (47) in Regression R1-LL, the following relation holds

$$p_t = \overline{\phi} + \phi_1\mu_x + \phi_n\mu_n + \left(\phi_0 + \rho\phi_1\right) x_t + \phi_1 u_t + \phi_n\varepsilon_t^n. \tag{51}$$

By comparing Regression R2-LL with the structural Equation (51), it follows that $\overline{\zeta} = \overline{\phi} + \phi_1\mu_x + \phi_n\mu_n$, $\zeta_0 = \phi_0 + \rho\phi_1$, and $\varepsilon_t^\zeta = \phi_1 u_t + \phi_n\varepsilon_t^n$.

From Equation (51), the following variance decomposition must hold

$$\begin{aligned}\mathbb{V}\mathrm{ar}\left(p_t\right) &= \mathbb{V}\mathrm{ar}\left(\zeta_0 x_t\right) + \mathbb{V}\mathrm{ar}\left(\phi_1 u_t + \phi_n\varepsilon_t^n\right)\\ &= \mathbb{V}\mathrm{ar}\left(\zeta_0 x_t\right) + (\phi_1)^2\,\mathbb{V}\mathrm{ar}\left(u_t\right) + \mathbb{V}\mathrm{ar}\left(e_t\right)\end{aligned}$$

which can be rearranged to express $\frac{\tau_{\tilde{\pi}}}{\tau_u}$ as follows

$$1 = \underbrace{\frac{\mathbb{V}\mathrm{ar}\left(\zeta_0 x_t\right)}{\mathbb{V}\mathrm{ar}\left(p_t\right)}}_{R_x^2} + \underbrace{\frac{\mathbb{V}\mathrm{ar}\left(e_t\right)}{\mathbb{V}\mathrm{ar}\left(p_t\right)}}_{1-R^2_{x,x'}}\left(\underbrace{\frac{(\phi_1)^2}{\mathbb{V}\mathrm{ar}\left(e_t\right)}\mathbb{V}\mathrm{ar}\left(u_t\right)}_{\frac{\tau_{\tilde{\pi}}}{\tau_u}} + 1\right) \Rightarrow \frac{\tau_{\tilde{\pi}}}{\tau_u} = \frac{R^2_{x,x'} - R^2_x}{1 - R^2_{x,x'}}$$

Therefore, relative price informativeness can be written as

$$\tau^R_{\tilde{\pi}} = \frac{\tau_{\tilde{\pi}}}{\tau_{\tilde{\pi}} + \tau_u} = \frac{1}{1 + \frac{1}{\frac{\tau_{\tilde{\pi}}}{\tau_u}}} = \frac{R^2_{x,x'} - R^2_x}{1 - R^2_x}.$$

□

Microfoundation

Time is discrete, with dates denoted by $t = 0, 1, 2, \ldots, \infty$. The economy is populated by a continuum of investors, indexed by $i \in I$, who live for two dates. An investor born at date t has well-behaved expected utility preferences over terminal wealth w_{t+1}^i, with flow utility given by $U_i\left(w_{t+1}^i\right)$, where $U_i'(\cdot) > 0$ and $U_i''(\cdot) < 0$.

There are two long-term assets in the economy: a risk-free asset in perfectly elastic supply, with gross return $R^f > 1$, and a risky asset in fixed supply Q, whose date t (log) payoff is $x_t = \ln(X_t)$ and which trades at a (log) price $p_t = \ln(P_t)$. The process followed by x_t is given by

$$x_{t+1} = \mu_x + \rho x_t + u_t, \tag{52}$$

where $\Delta x_{t+1} = x_t - x_{t-1}$, μ_x is a scalar, $|\rho| < 1$, and $x_0 = \Delta x_0 = 0$. The realized payoff x_t is common knowledge to all investors before the price p_t is determined. The realized payoff at date $t+1$, x_{t+1}, is only only revealed to investors at date $t+1$.

We assume that investors receive private signals about the innovation to the risky asset payoff. Formally, each investor receives a signal about the payoff innovation u_t given by

$$s_t^i = u_t + \varepsilon_{st}^i \quad \text{with} \quad \varepsilon_{st}^i \sim N\left(0, \tau_s^{-1}\right),$$

where $\varepsilon_{st}^i \perp \varepsilon_{st}^j$ for all $i \neq j$, and $u_t \perp \varepsilon_{st}^i$ for all t and all i.

We also assume that investors also have private trading motives that arise from random heterogeneous priors that are random in the aggregate. Formally, each investor i born at date t has a prior over u_t given by

$$u_t \sim_i N\left(\overline{n}_t^i, \tau_u^{-1}\right),$$

where

$$\overline{n}_t^i = n_t + \varepsilon_{\overline{n}t}^i \quad \text{with} \quad \varepsilon_{\overline{n}t}^i \overset{\text{iid}}{\sim} N\left(0, \tau_{\overline{n}}^{-1}\right),$$

and

$$n_t = \mu_n + \varepsilon_t^n \quad \text{with} \quad \varepsilon_t^n \sim N\left(0, \tau_n^{-1}\right),$$

where μ_n is a scalar, and where $\varepsilon_t^n \perp \varepsilon_{\overline{n}t}^i$ for all t and all i. The variable n_t, which can be interpreted as the aggregate sentiment in the economy, is not observed and acts as a source of aggregate noise, preventing the asset price from being fully revealing.

Each investor i born at date t is endowed with wealth w_t^i, and optimally chooses a portfolio share in the risky asset, denoted by θ_t^i, to solve

$$\max_{\theta_t^i} \mathbb{E}_t^i\left[U_i\left(w_{t+1}^i\right)\right] \tag{53}$$

subject to a wealth accumulation constraint

$$w_{t+1}^i = \left(R^f + \theta_t^i\left(\frac{X_{t+1} + P_{t+1}}{P_t} - R^f\right)\right) w_t^i, \tag{54}$$

where the information set of an investor i in period t is given by $\mathcal{I}_t^i = \{s_t^i, \overline{n}_t^i, x_t, p_t\}$.

The optimality condition of an investor that maximizes Equation (53) subject to the wealth accumulation constraint in Equation (54) is given by

$$\mathbb{E}\left[U_i'\left(w_{t+1}^i\right)\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\middle|\mathcal{I}_t^i\right]=0. \tag{55}$$

We approximate an investor's first order condition in three steps.

First , we take a first order Taylor expansion of an investor's future marginal utility $U'\left(w_{t+1}^i\right)$ around the current date t wealth level w_t^i. Formally, we approximate $U'\left(w_{t+1}^i\right)$ as follows

$$U'\left(w_{t+1}^i\right)\approx U'\left(w_t^i\right)+U''\left(w_t^i\right)\Delta w_{t+1}^i,$$

which allows us to express Equation (55) as

$$U'\left(w_t^i\right)\mathbb{E}_i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]+U''\left(w_t^i\right)w_t^i\mathbb{E}_i\left[\left(\left(R^f-1\right)+\theta_t^i\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\right)\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\right]\approx 0.$$

Second, we impose that terms that involve the product of two or more net interest rates are negligible. In continuous time, these terms would be of order $(dt)^2$. Formally, it follows that

$$\left(R^f-1\right)\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]\approx 0 \quad \text{and} \quad \left(\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]\right)^2\approx 0,$$

which allows us to express Equation (55) as

$$U'\left(w_t^i\right)\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]+U''\left(w_t^i\right)w_t^i\theta_t^i\mathbb{V}ar_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}\right]\approx 0.$$

Therefore, we can express an investor's risky portfolio share θ_t^i as

$$\theta_t^i\approx\frac{1}{\gamma^i}\frac{\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]}{\mathbb{V}ar_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}\right]}, \tag{56}$$

where $\gamma^i\equiv-\frac{w^iU''\left(w^i\right)}{U'\left(w^i\right)}$ denotes the coefficient of relative risk aversion.

Third, as in Campbell and Shiller (1988), we take a log-linear approximation of returns around a predetermined dividend-price ratio. Formally, note that

$$\begin{aligned}\frac{X_{t+1}+P_{t+1}}{P_t}&=e^{\ln\left(\frac{\left(1+\frac{P_{t+1}}{X_{t+1}}\right)\frac{X_{t+1}}{X_t}}{\frac{P_t}{X_t}}\right)}\\ \ln\left[\frac{X_{t+1}+P_{t+1}}{P_t}\right]&=\ln\left(1+\frac{P_{t+1}}{X_{t+1}}\right)+x_{t+1}-x_t-\left(p_t-x_t\right)\\ &=\ln\left(1+e^{p_{t+1}-x_{t+1}}\right)+\Delta x_{t+1}-\left(p_t-x_t\right),\end{aligned}$$

where we define $r^f=\ln R^f$. Following Campbell and Shiller (1988), we approximate the first term around a point $PX = e^{p-x}$, to find that

$$\ln\left(1+e^{\ln P_{t+1}-\ln X_{t+1}}\right) \approx \ln(1+PX) + \frac{PX}{PX+1}\left(p_{t+1}-x_{t+1}-p-x\right).$$
$$= k_0 + k_1\left(p_{t+1}-x_{t+1}\right),$$

where $k_1 = \frac{PX}{PX+1}$ and $k_0 = \ln(1+PX) - k_1(p-x)$.

Therefore, starting from Equation (56), we have that the risky asset demand of an investor i can be approximated as

$$\theta_t^i \approx \frac{1}{\gamma^i}\frac{k_0 + k_1\mathbb{E}_t^i\left[p_{t+1}-x_{t+1}\right] + \mathbb{E}_t^i\left[\Delta x_{t+1}\right] - (p_t - x_t) - r^f}{\mathbb{V}ar^i\left[k_1\left(p_{t+1}-x_{t+1}\right)+\Delta x_{t+1}|\mathcal{I}_t^i\right]}, \tag{57}$$

where we define $r^f \equiv \ln R^f$ and we used that $e^y \approx 1+y$.

In order to characterize the equilibrium it is necessary to characterize investors' expectations. We conjecture and subsequently verify that $k_1\mathbb{E}_t^i\left[p_{t+1}-x_{t+1}\right] + \mathbb{E}_t^i\left[\Delta x_{t+1}\right]$ is linear in s_t^i, $\overline{n}_t^i$, and x_t and that $\mathbb{V}ar_t^i\left[k_1\left(p_{t+1}-x_{t+1}\right)+\Delta x_{t+1}\right]$ is a constant, which we denote by V. Under this conjecture, θ_t^i is a linear function of $s_t^i, x_t, \overline{n}_t^i$ and it is given by

$$\theta_t^i = \alpha_x^i x_t + \alpha_s^i s_t^i + \alpha_n^i \overline{n}_t^i - \alpha_p^i p_t + \psi^i.$$

Using this expression and the market clearing condition

$$\int \theta_t^i w_t^i di = Q$$

implies

$$p_t = \frac{\overline{\alpha_x}}{\overline{\alpha_p}}x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}}u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}}n_t + \frac{\overline{\psi}}{\overline{\alpha_p}}. \tag{58}$$

This expression can also be written as

$$p_t = \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}\rho\right)x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}}x_{t+1} + \frac{\overline{\alpha_n}}{\overline{\alpha_p}}n_t + \left(\frac{\overline{\psi}}{\overline{\alpha_p}} - \frac{\overline{\alpha_s}}{\overline{\alpha_p}}\mu_x\right)$$

Investors in the model learn from the price. The information contained in the price for an investor in the model is

$$\hat{\pi}_t = \frac{\overline{\alpha_p}}{\overline{\alpha_s}}\left(p_t - \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}}x_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}}\mu_n - \frac{\overline{\psi}}{\overline{\alpha_p}}\right)\right)$$

which has a precision

$$\tau_{\hat{\pi}} \equiv \mathbb{V}ar\left[|\hat{\pi}_t|\, u_t, x_t\right]^{-1} = \left(\frac{\overline{\alpha_s}}{\overline{\alpha_n}}\right)^2\tau_n.$$

Then,

$$\mathbb{E}_t^i\left[u_t\right] = \mathbb{E}\left[u_t|s_t^i, \overline{n}_t^i, p_t\right] = \frac{\tau_s s_t^i + \tau_u \overline{n}_t^i + \tau_{\hat{\pi}}\hat{\pi}_t}{\tau_s + \tau_u + \tau_{\hat{\pi}}} = \frac{\tau_s s_t^i + \tau_u \overline{n}_t^i + \tau_{\hat{\pi}}\frac{\overline{\alpha_p}}{\overline{\alpha_s}}\left(p_t - \frac{\overline{\alpha_x}}{\overline{\alpha_p}}x_t - \frac{\overline{\alpha_n}}{\overline{\alpha_s}}\mu_n - \frac{\overline{\psi}}{\overline{\alpha_p}}\right)}{\tau_s + \tau_u + \tau_{\hat{\pi}}}$$

and

$$\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right] = \left(\tau_s + \tau_u + \tau_{\hat{\pi}}\right)^{-1}.$$

Note that these two expressions imply that our conjectures above are satisfied. To see this note that

$$\begin{aligned}
k_1\mathbb{E}_t^i\left[p_{t+1} - x_{t+1}\right] + \mathbb{E}_t^i\left[\Delta x_{t+1}\right] &= k_1\mathbb{E}_t^i\left[\frac{\overline{\alpha_x}}{\overline{\alpha_p}}x_{t+1} + \frac{\overline{\alpha_s}}{\overline{\alpha_p}}u_{t+1} + \frac{\overline{\alpha_n}}{\overline{\alpha_p}}n_{t+1} + \frac{\overline{\psi}}{\overline{\alpha_p}} - x_{t+1}\right] + \mathbb{E}_t^i\left[\mu_x + (\rho-1)\,x_t + u_t\right] \\
&= k_1\left(\mathbb{E}_t^i\left[\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1\right)x_{t+1} + \frac{\overline{\alpha_s}}{\overline{\alpha_p}}u_{t+1}\right] + \frac{\overline{\alpha_n}}{\overline{\alpha_p}}\mu_n + \frac{\overline{\psi}}{\overline{\alpha_p}}\right) + (\rho-1)\,x_t + \mu_x + \mathbb{E}_t^i\left[u_t\right] \\
&= k_1\left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1\right)\left(\mu_x + \mathbb{E}_t^i\left[u_t\right]\right) + \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1\right)\rho x_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}}\mu_n + \frac{\overline{\psi}}{\overline{\alpha_p}}\right) + (\rho-1)\,x_t + \mu_x + \\
&= k_1\left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)\left(\mu_x + \mathbb{E}_t^i\left[u_t\right]\right) + \left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1\right)\rho + \frac{(\rho-1)}{k_1}\right)x_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}}\mu_n + \frac{\overline{\psi}}{\overline{\alpha_p}}\right)
\end{aligned}$$

where we used that $\mathbb{E}_t^i\left[u_{t+1}\right] = 0$ and that $\mathbb{E}_t^i\left[\varepsilon_{t+1}^n\right] = 0$. Moreover,

$$\begin{aligned}
\mathbb{V}ar^i\left[k_1\left(p_{t+1} - x_{t+1}\right) + \Delta x_{t+1}|\mathcal{I}_t^i\right] &= \mathbb{V}ar^i\left[k_1\left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1\right)x_{t+1} + \frac{\overline{\alpha_s}}{\overline{\alpha_p}}u_{t+1} + \frac{\overline{\alpha_n}}{\overline{\alpha_p}}n_{t+1}\right) + u_t|\mathcal{I}_t^i\right] \\
&= k_1^2\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)^2\mathbb{V}ar^i\left[u_t|\mathcal{I}_t^i\right] + k_1^2\left(\frac{\overline{\alpha_s}}{\overline{\alpha_p}}\right)^2\mathbb{V}ar^i\left[u_{t+1}|\mathcal{I}_t^i\right] + k_1^2\left(\frac{\overline{\alpha_n}}{\overline{\alpha_p}}\right)^2\mathbb{V}ar \\
&= k_1^2\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)^2\left(\tau_s + \tau_u + \tau_{\hat{\pi}}\right)^{-1} + k_1^2\left(\frac{\overline{\alpha_s}}{\overline{\alpha_p}}\right)^2\tau_u^{-1} + k_1^2\left(\frac{\overline{\alpha_n}}{\overline{\alpha_p}}\right)^2\tau_n^{-1}.
\end{aligned}$$

Using these expression in the first order condition and matching coefficients gives

$$\begin{aligned}
\alpha_x^i &= \frac{1}{\kappa_i}k_1\left(-\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)\frac{\tau_\pi\frac{\overline{\alpha_x}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} + \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)\rho\right) \\
\alpha_s^i &= \frac{1}{\kappa_i}k_1\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)\frac{\tau_s}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \\
\alpha_n^i &= \frac{1}{\kappa_i}k_1\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)\frac{\tau_u}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \\
\alpha_p^i &= \frac{1}{\kappa_i}\left(k_1\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)\frac{\tau_\pi\frac{\overline{\alpha_p}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} - 1\right) \\
\psi^i &= \frac{1}{\kappa_i}\left(k_0 + k_1\left(-\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)\left(\frac{\tau_\pi\frac{\overline{\alpha_p}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} - \mu_x\right) + 1\right)\left(\frac{\overline{\alpha_n}}{\overline{\alpha_p}}\mu_n + \frac{\overline{\psi}}{\overline{\alpha_p}}\right) - r^f\right)
\end{aligned}$$

where $\kappa_i \equiv \gamma^i\mathbb{V}ar^i\left[k_1\left(p_{t+1} - x_{t+1}\right) + \Delta x_{t+1}|\mathcal{I}_t^i\right]$.

In this equilibrium, our guess in Eq. (57) is verified and the equilibrium price is linear and can be expressed as in Eq. (48).

C.2 An Exact CARA-Normal Formulation

In an earlier version of this paper, we developed our identification results using an exact linear formulation, motivated by the use of a CARA-Normal framework, which is the workhorse model in the learning literature, see e.g., Vives (2008) and Veldkamp (2011). In this section, we reproduce our identification results using these exact linear formulations in the case of difference stationary and stationary linear payoffs, and we provide microfoundations in the context of CARA-Normal models.

C.2.1 Difference stationary linear payoff

General framework and identification We consider a discrete time environment with dates $t = 0, 1, 2, \ldots, \infty$, in which investors trade a risky asset in fixed supply at a price P_t at each date t. We assume that the payoff of the risky asset at date $t+1$, X_{t+1}, follows a difference stationary AR(1) process

$$\Delta X_{t+1} = \mu_{\Delta X} + \rho \Delta X_t + u_t, \tag{59}$$

where $\mu_{\Delta X}$ is a scalar, $|\rho| < 1$, and where the innovations to the payoff, u_t, have mean zero, a finite variance denoted by $\mathbb{V}ar\left[u_t\right] = \sigma_u^2 = \tau_u^{-1}$, and are identically and independently distributed over time. We assume that the equilibrium price difference is given by

$$\Delta P_t = \overline{\phi} + \phi_0 \Delta X_t + \phi_1 \Delta X_{t+1} + \phi_n \Delta n_t, \tag{60}$$

where $\overline{\phi}$, ϕ_0, ϕ_1, and ϕ_n are parameters and where n_t represents the aggregate component of investors' trading motives that are orthogonal to the asset payoff, given by $\Delta n_t = \mu_{\Delta n} + \varepsilon_t^{\Delta n}$, where $\mathbb{E}\left[\varepsilon_t^{\Delta n}\right] = 0$ and $\mathbb{V}ar\left[\varepsilon_t^{\Delta n}\right] = \sigma_n^2 = \tau_{\Delta n}^{-1}$. For simplicity, we assume that u_t and Δn_t are independent.

In this case, the unbiased signal of the innovation to the change in the future payoff u_t contained in the price, which we denote by Π_t, is given respectively by

$$\Pi_t \equiv \frac{\Delta P_t - \left(\overline{\phi} + \phi_1 \mu_{\Delta X} + \phi_n \mu_{\Delta n} + (\phi_0 + \rho \phi_1) \Delta X_t\right)}{\phi_1} = u_t + \frac{\phi_n}{\phi_1} \left(\Delta n_t - \mu_{\Delta n}\right)$$

and absolute and relativeprice informativeness are given by

$$\tau_\Pi \equiv \left(\mathbb{V}\mathrm{ar}\left[\Pi_t \middle| \Delta X_{t+1}, \Delta X_t\right]\right)^{-1} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_{\Delta n} \quad \text{and} \quad \tau_\Pi^R \equiv \frac{\tau_\Pi}{\tau_\Pi + \tau_u}.$$

Proposition 8. (Identifying price informativeness difference stationary linear)

a) Absolute price informativeness. Let $\overline{\beta}$, β_0, and β_1 denote the coefficients of the following regression of prices on realized and future payoffs,

$$\Delta P_t = \overline{\beta} + \beta_0 \Delta X_t + \beta_1 \Delta X_{t+1} + e_t, \tag{R1-Linear-Diff}$$

where ΔP_t denotes the date t price change, ΔX_t and ΔX_{t+1} respectively denote the dates t and $t+1$ payoff change, and where $\sigma_e^2 = \mathbb{V}ar\left[e_t\right]$ denotes the variance of the error. Then, absolute price informativeness, τ_Π, can be recovered by

$$\tau_\Pi = \frac{\beta_1^2}{\sigma_e^2}. \tag{61}$$

The OLS estimation of Regression R1-Linear-Diff yields consistent estimates of β_1 and σ_e^2.

b) Relative Price Informativeness. Let $R^2_{\Delta X,\Delta X'}$ denote the R-squared of Regression R1-Linear-Diff. Let $R^2_{\Delta X}$, ζ, and ζ_0 respectively denote the R-squared and the coefficients of the following regression of price differences on payoff differences,

$$\Delta P_t = \overline{\zeta} + \zeta_0 \Delta X_t + e_t^{\zeta}. \tag{R2-Linear-Dif}$$

Then, relative price informativeness, τ_Π^R, can be recovered by

$$\tau_\Pi^R = \frac{R^2_{\Delta X,\Delta X'} - R^2_{\Delta X}}{1 - R^2_{\Delta X}}. \tag{62}$$

The OLS estimation of Regressions R1-Linear-Diff and R2-Linear-Dif yields consistent estimates of $R^2_{\Delta X,\Delta X'}$ and $R^2_{\Delta X}$.

Proof. a) By comparing Regression R1-Linear-Diff with the structural Equation (60), it follows that $\overline{\beta} = \overline{\phi} + \phi_n \mu_{\Delta n}$, $\beta_0 = \phi_0$, $\beta_1 = \phi_1$, and $e_t = \phi_n \varepsilon_t^{\Delta n}$. Consequently, $\sigma_e^2 = \mathbb{V}\mathrm{ar}\left[e_t\right] = \left(\phi_n\right)^2 \mathbb{V}\mathrm{ar}\left[\varepsilon_t^{\Delta n}\right] = \left(\phi_n\right)^2 \tau_{\Delta n}^{-1}$. Therefore, we can recover absolute price informativeness as follows

$$\tau_\Pi = \frac{\left(\beta_1\right)^2}{\sigma_e^2} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_{\Delta n}.$$

Given Equations (59) and (60), as well as the assumptions on u_t and n_t, it is straightforward to show that the OLS estimates of Regressions R1-Linear-Diff and R2-Linear-Dif are consistent, which implies that price informativeness can be consistently estimated as $\widehat{\tau_\Pi} = \frac{\left(\widehat{\beta}_1\right)^2}{\widehat{\sigma_e^2}}$. Formally, $\operatorname{plim}\left(\widehat{\tau_\Pi}\right) = \operatorname{plim}\left(\frac{\left(\widehat{\beta}_1\right)^2}{\widehat{\sigma_e^2}}\right) = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_{\Delta n} = \tau_\Pi$.

b) Note that the R-squareds of Regressions R1-Linear-Diff and R2-Linear-Dif can be expressed as follows

$$R^2_{\Delta X,\Delta X'} = 1 - \frac{\mathbb{V}\mathrm{ar}\left(e_t\right)}{\mathbb{V}\mathrm{ar}\left(\Delta P_t\right)} \quad \text{and} \quad R^2_{\Delta X} = \frac{\mathbb{V}\mathrm{ar}\left(\zeta_0 \Delta X_t\right)}{\mathbb{V}\mathrm{ar}\left(\Delta P_t\right)}.$$

After substituting Equation (59) in Regression R1-Linear-Diff, the following relation holds

$$\Delta P_t = \overline{\phi} + \phi_1 \mu_{\Delta X} + \phi_n \mu_{\Delta n} + \left(\phi_0 + \rho \phi_1\right) \Delta X_t + \phi_1 u_t + \phi_n \varepsilon_t^{\Delta n}. \tag{63}$$

By comparing Regression R2-Linear-Dif with the structural Equation (63), it follows that $\overline{\zeta} = \overline{\phi} + \phi_1 \mu_{\Delta X} + \phi_n \mu_{\Delta n}$, $\zeta_0 = \phi_0 + \rho \phi_1$, and $\varepsilon_t^{\zeta} = \phi_1 u_t + \phi_n \varepsilon_t^{\Delta n}$.

From Equation (63), the following variance decomposition must hold

$$\begin{aligned}\mathbb{V}\text{ar}\left(\Delta P_t\right) &= \mathbb{V}\text{ar}\left(\zeta_0 \Delta X_t\right) + \mathbb{V}\text{ar}\left(\phi_1 u_t + \phi_n \varepsilon_t^{\Delta n}\right) \\ &= \mathbb{V}\text{ar}\left(\zeta_0 \Delta X_t\right) + \left(\phi_1\right)^2 \mathbb{V}\text{ar}\left(u_t\right) + \mathbb{V}\text{ar}\left(e_t\right)\end{aligned}$$

which can be rearranged to express $\frac{\tau_\Pi}{\tau_u}$ as follows

$$1 = \underbrace{\frac{\mathbb{V}\text{ar}\left(\zeta_0 x_t\right)}{\mathbb{V}\text{ar}\left(\Delta P_t\right)}}_{R^2_{\Delta X}} + \underbrace{\frac{\mathbb{V}\text{ar}\left(e_t\right)}{\mathbb{V}\text{ar}\left(\Delta P_t\right)}}_{1-R^2_{\Delta X,\Delta X'}}\left(\underbrace{\frac{\left(\phi_1\right)^2}{\mathbb{V}\text{ar}\left(e_t\right)}\mathbb{V}\text{ar}\left(u_t\right)}_{\frac{\tau_\Pi}{\tau_u}} + 1\right) \Rightarrow \frac{\tau_\Pi}{\tau_u} = \frac{R^2_{\Delta X,\Delta X'} - R^2_{\Delta X}}{1 - R^2_{\Delta X,\Delta X'}}$$

Therefore, relative price informativeness can be written as

$$\tau_\Pi^R = \frac{\tau_\Pi}{\tau_\Pi + \tau_u} = \frac{1}{1 + \frac{1}{\frac{\tau_\Pi}{\tau_u}}} = \frac{R^2_{\Delta X,\Delta X'} - R^2_{\Delta X}}{1 - R^2_{\Delta X}}.$$

□

Microfoundation Time is discrete, with periods denoted by $t = 0, 1, 2, \ldots, \infty$. Each period t, there is a continuum of investors, indexed by $i \in I$. Each generation lives two periods and has exponential utility over their last period wealth. An investor born at time t has preferences given by

$$U\left(W_{t+1}\right) = -e^{-\gamma W_{t+1}},$$

where γ is the coefficient of absolute risk aversion and W_{t+1} is the investor's wealth in his final period. There are two long-term assets in the economy: A risk-free asset in perfectly elastic supply, with return $R > 1$, and a risky asset in fixed supply Q that trades at price P_t in period t. The process for the payoff of the risky asset each period t is given by

$$\Delta X_{t+1} = \mu_{\Delta X} + u_t,$$

where $\Delta X_t = X_t - X_{t-1}$,$\mu_{\Delta X}$ is a scalar and $X_0 = 0$. The payoff X_t is realized and becomes common knowledge at the end of period $t-1$. The innovation in the payoff process, u_t, and, hence, X_{t+1} are realized and observed at the end on period t. The innovations to the payoff are independently distributed over time.

To preserve tractability, we assume that investors' private trading needs arise from random heterogeneous priors– see Davila and Parlatore (2017) for a thorough analysis of this formulation. Formally, each investor i in generation t has a prior over the innovation at time t given by

$$u_t \sim_i N\left(\overline{n}_t^i, \tau_u^{-1}\right),$$

where

$$\overline{n}_t^i = n_t + \varepsilon_{\overline{n}t}^i \quad \text{with} \quad \varepsilon_{\overline{n}t}^i \overset{\text{iid}}{\sim} N\left(0, \tau_{\overline{n}}^{-1}\right)$$

and $\Delta n_t = \mu_{\Delta n} + \varepsilon_t^n$ with $\varepsilon_t^n \sim N\left(0, \tau_{\Delta n}^{-1}\right)$. The term n_t can be interpreted as the aggregate sentiment in the economy, where $n_t \perp \varepsilon_{\overline{n}t}^i$ for all t and all i. The aggregate sentiment n_t is not observed and acts as a source of aggregate noise in the economy, preventing the price from being fully revealing. For simplicity we assume $n_t \perp u_{t+s}$ for all t and all s. Moreover, we assume investors think of their prior as the correct one and do not learn about the aggregate sentiment from it.[12]

Each investor i in generation t receives a signal about the innovation in the asset payoff u_t given by

$$s_t^i = u_t + \varepsilon_{st}^i \quad \text{with} \quad \varepsilon_{st}^i \sim N\left(0, \tau_s^{-1}\right)$$

and $\varepsilon_{st}^i \perp \varepsilon_{st}^j$ for all $i \perp j$, and $u_t \perp \varepsilon_{st}^i$ for all t and all i.

The asset demand submitted by investor i born in period t is given by the solution to the following problem

$$\max_{q_t^i} \left(\mathbb{E}\left[X_{t+1} + R^{-1} p_{t+1} | \mathcal{I}_t^i\right] - P_t\right) Q_t^i - \frac{\gamma}{2} \mathbb{V}ar\left[X_{t+1} + R^{-1} P_{t+1} | \mathcal{I}_t^t\right] \left(Q_t^i\right)^2,$$

where $\mathcal{I}_t^i = \left\{X_t, s_t^i, \overline{n}_t^i, P_t\right\}$ is the information set of an investor i in period t.

The optimality condition for an investor i in period t satisfies

$$Q_t^i = \frac{\mathbb{E}\left[X_{t+1} + R^{-1} P_{t+1} | \mathcal{I}_t^i\right] - P_t}{\gamma \mathbb{V}ar\left[X_{t+1} + R^{-1} P_{t+1} | \mathcal{I}_t^i\right]}.$$

In a stationary equilibrium in linear strategies, we assume and subsequently verify that the equilibrium demand of investor i can be expressed as

$$\Delta Q_t^i = \alpha_X^i X_t + \alpha_s^i s_t^i + \alpha_n^i \overline{n}_t^i - \alpha_P^i P_t + \psi^i, \tag{64}$$

where α_θ^i, α_s^i, α_n^i, α_p^i, and ψ^i are individual equilibrium demand coefficients. Market clearing and the Strong Law of Large Numbers (SLLN) allows us to express the equilibrium price in period t as

$$P_t = \frac{\overline{\alpha_X}}{\overline{\alpha_P}} X_t + \frac{\overline{\alpha_s}}{\overline{\alpha_P}} u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_P}} n_t + \frac{\overline{\psi}}{\overline{\alpha_P}},$$

where we define cross sectional averages $\overline{\alpha_X} = \int \alpha_X^i di$, $\overline{\alpha_s} = \int \alpha_s^i di$, $\overline{\alpha_P} = \int \alpha_P^i di$, and $\overline{\psi} = \int \psi^i di - Q$.

The unbiased signal of the innovation in the payoff contained in the price is

$$\Pi_t = \frac{\overline{\alpha_P}}{\overline{\alpha_s}} \left(P_t - \frac{\overline{\alpha_n}}{\overline{\alpha_s}} \mu_{\Delta n} - \frac{\overline{\alpha_X}}{\overline{\alpha_P}} X_t - \frac{\overline{\psi}}{\overline{\alpha_P}}\right) = u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_s}} \left(n_t - \mu_{\Delta n}\right),$$

[12] Davila and Parlatore (2017) show that the equilibrium structure is preserved if this assumption is relaxed.

where

$$\Pi_t | X_{t+1}, X_t \sim N\left(u_t, \tau_\Pi^{-1}\right),$$

with price informativeness given by

$$\tau_\Pi = \left(\mathbb{V}\text{ar}\left[\Pi_t | X_{t+1}, X_t\right]\right)^{-1} = \left(\frac{\overline{\alpha_s}}{\overline{\alpha_n}}\right)^2 \tau_{\Delta n}.$$

Given our guesses for the demand functions and the linear structure of prices we have

$$X_{t+1} + R^{-1} P_{t+1} = X_{t+1} + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}} X_{t+1} + R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}} u_{t+1} + R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}} n_{t+1} + R^{-1} \frac{\overline{\psi}}{\overline{\alpha_P}},$$

$$\begin{aligned}
\mathbb{E}\left[X_{t+1} + R^{-1} P_{t+1} | \mathcal{I}_t^i\right] &= \left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right) \mathbb{E}\left[X_{t+1} | \mathcal{I}_t^i\right] + R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}} \mathbb{E}\left[u_{t+1}\right] + R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}} \mathbb{E}\left[n_{t+1}\right] + R^{-1} \frac{\overline{\psi}}{\overline{\alpha_P}} \\
&= \left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right) \left(X_t + \mathbb{E}\left[u_t | \mathcal{I}_t^i\right]\right) + R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_p}} \mathbb{E}\left[u_{t+1}\right] + R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}} \mu_{\Delta n} + R^{-1} \frac{\overline{\psi}}{\overline{\alpha_P}},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{V}ar\left[X_{t+1} + R^{-1} P_{t+1} | \mathcal{I}_t^i\right] &= \left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[X_{t+1} | \mathcal{I}_t^i\right] + \left(R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_{t+1}\right] + \left(R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[n_{t+1}\right] \\
&= \left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_t | \mathcal{I}_t^i\right] + \left(R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_{t+1}\right] + \left(R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[n_{t+1}\right].
\end{aligned}$$

Moreover, given the Gaussian structure of the signals in the information set, Bayesian updating implies

$$\mathbb{E}\left[u_t | s_t^i, \overline{n}_t^i, P_t\right] = \frac{\tau_s s_t^i + \tau_u \overline{n}_t^i + \tau_\Pi \Pi_t}{\tau_s + \tau_u + \tau_\Pi} = \frac{\tau_s s_t^i + \tau_{\Delta n} \overline{n}_t^i + + \tau_\Pi \frac{\overline{\alpha_P}}{\overline{\alpha_s}} \left(P_t - \frac{\overline{\alpha_n}}{\overline{\alpha_s}} \mu_{\Delta n} - \frac{\overline{\alpha_X}}{\overline{\alpha_P}} X_t - \frac{\overline{\psi}}{\overline{\alpha_P}}\right)}{\tau_s + \tau_u + \tau_\Pi},$$

and

$$\mathbb{V}ar\left[u_t | \mathcal{I}_t^i\right] = \mathbb{V}ar\left[t_t | s_t^i, \overline{n}_t^i, P_t\right] = \left(\tau_s + \tau_u + \tau_\Pi\right)^{-1}.$$

Then, the first order condition is given by

$$Q_t^i = \frac{1}{\gamma} \frac{\left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right) \left(X_t + \mathbb{V}ar\left[u_t | \mathcal{I}_t^i\right] \left(\tau_s s_t^i + \tau_u \overline{n}_t^i + \tau_\Pi \Pi\right)\right) + R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}} \mathbb{E}\left[u_{t+1}\right] + R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}} \mu_{\Delta n} + R^{-1} \frac{\overline{\psi}}{\overline{\alpha_P}} - P_t}{\left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_t | \mathcal{I}_{it}\right] + \left(R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_{t+1}\right] + \left(R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}}\right)^2 \tau_{\Delta n}^{-1}}.$$

Matching coefficients we have

$$
\begin{aligned}
\alpha_s^i &= \frac{\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa}\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_s \\
\alpha_n^i &= \frac{\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa}\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_\eta \\
\alpha_X^i &= \frac{\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa}\left(1-\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_\Pi\frac{\overline{\alpha_X}}{\overline{\alpha_s}}\right) \\
\alpha_P^i &= \frac{1}{\kappa}\left(1-\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_\Pi\frac{\overline{\alpha_p}}{\overline{\alpha_s}}\right) \\
\psi^i &= -\frac{1}{\kappa}\left(\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_\Pi\left(\frac{\overline{\alpha_n}}{\overline{\alpha_s}}\mu_{\Delta n}+\frac{\overline{\psi}}{\overline{\alpha_s}}\right)-R^{-1}\left(\frac{\overline{\alpha_n}}{\overline{\alpha_P}}\mu_{\Delta n}+\frac{\overline{\psi}}{\overline{\alpha_P}}\right)\right),
\end{aligned}
\tag{65}
$$

where

$$
\kappa \equiv \gamma\left(\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)^2\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]+\left(R^{-1}\frac{\overline{\alpha_s}}{\overline{\alpha_P}}\right)^2\mathbb{V}ar\left[u_{t+1}\right]+\left(R^{-1}\frac{\overline{\alpha_n}}{\overline{\alpha_P}}\right)^2\tau_{\Delta n}^{-1}\right),
$$

since $\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]=(\tau_s+\tau_u+\tau_\Pi)^{-1}$ for all i.

Then, an equilibrium in linear strategies always exists if the system above has a solution. Note that the demand sensitivities are the same for all i. Then, there exists a unique solution to the system in Equations (65), that is given by

$$
\begin{aligned}
\alpha_s^i &= \frac{1}{\kappa}\frac{1}{1-R^{-1}}\frac{\tau_s}{\tau_u+\tau_s+\tau_\Pi}, \quad \alpha_n^i = \frac{1}{\kappa}\frac{1}{1-R^{-1}\rho}\frac{\tau_\eta}{\tau_u+\tau_s+\tau_\Pi} \\
\alpha_X^i &= \frac{1}{\kappa}\frac{\rho}{1-R^{-1}}\frac{\tau_s}{\tau_s+\tau_\Pi}, \quad \alpha_P^i = \frac{1}{\kappa}\frac{\tau_s}{\tau_s+\tau_\Pi}, \quad \text{and} \\
\psi^i &= -\frac{\frac{1}{\kappa}\frac{1}{1-R^{-1}}\left(\left(1-R^{-1}\right)\tau_\Pi-R^{-1}\tau_s\right)\frac{\tau_u}{\tau_u+\tau_s+\tau_{\hat{\Pi}}}\mu_{\Delta n}}{1+\left(1-R^{-1}\right)\tau_\Pi-R^{-1}\tau_s},
\end{aligned}
$$

where

$$
\tau_\Pi = \left(\frac{\tau_s}{\tau_u}\right)^2\tau_{\Delta n}
$$

and

$$
\kappa = \gamma\left(\left(\frac{1}{1-R^{-1}}\right)^2\frac{1}{\tau_u+\tau_s+\tau_\Pi}+\left(R^{-1}\frac{1}{1-R^{-1}}\frac{\tau_s+\tau_\Pi}{\tau_u+\tau_s+\tau_\Pi}\right)^2\tau_u^{-1}+\left(\frac{R^{-1}}{1-R^{-1}}\frac{\tau_s+\tau_\Pi}{\tau_u+\tau_s+\tau_\Pi}\frac{\tau_u}{\tau_s}\right)^2\tau_{\Delta n}^{-1}\right).
$$

In this equilibrium, our guess in Eq. (64) is verified and the equilibrium price is linear and can be expressed as in Eq. (67).

C.2.2 Stationary linear payoff

General framework and identification

Consider a discrete time environment with dates $t = 0, 1, 2, \ldots, \infty$, in which investors trade a risky asset in fixed supply at a price P_t at each date t. We assume that the payoff of the risky asset at date $t+1$, X_{t+1}, follows a stationary AR(1) process

$$X_{t+1} = \mu_X + \rho X_t + u_t, \tag{66}$$

where μ_X is a scalar, $|\rho| < 1$, and where the innovations to the payoff, u_t, have mean zero, a finite variance denoted by $\mathbb{V}ar\left[u_t\right] = \sigma_u^2 = \tau_u^{-1}$, and are identically and independently distributed over time. We assume that the equilibrium price is given by

$$P_t = \overline{\phi} + \phi_0 X_t + \phi_1 X_{t+1} + \phi_n n_t, \tag{67}$$

where $\overline{\phi}$, ϕ_0, ϕ_1, and ϕ_n are parameters and where n_t represents the aggregate component of investors' trading motives that are orthogonal to the asset payoff, given by $n_t = \mu_n + \varepsilon_t^n$, where $\mathbb{E}\left[\varepsilon_t^n\right] = 0$ and $\mathbb{V}ar\left[\varepsilon_t^n\right] = \sigma_n^2 = \tau_n^{-1}$. For simplicity, we assume that u_t and n_t are independent.

In this case, the unbiased signal of the innovation to the change in the future payoff u_t contained in the price, which we denote by π_t, is given by

$$\hat{\Pi}_t \equiv \frac{P_t - \left(\overline{\phi} + \phi_1 \mu_X + \phi_n \mu_n + \left(\phi_0 + \rho\phi_1\right) X_t\right)}{\phi_1} = u_t + \frac{\phi_n}{\phi_1}\left(n_t - \mu_n\right)$$

and absolute and relative price informativeness are given respectively by

$$\tau_{\hat{\Pi}} \equiv \left(\mathbb{V}\mathrm{ar}\left[\hat{\Pi}_t \middle| X_{t+1}, X_t\right]\right)^{-1} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_n \quad \text{and} \quad \tau_{\hat{\Pi}}^R \equiv \frac{\tau_{\hat{\Pi}}}{\tau_{\hat{\Pi}} + \tau_u}.$$

Proposition 9. (Identifying price informativeness difference stationary linear)

a) Absolute price informativeness. Let $\overline{\beta}$, β_0, and β_1 denote the coefficients of the following regression of prices on realized and future payoffs,

$$P_t = \overline{\beta} + \beta_0 X_t + \beta_1 X_{t+1} + e_t, \tag{R1-Linear}$$

where P_t denotes the date t price, X_t and X_{t+1} respectively denote the dates t and $t+1$ payoff, and where $\sigma_e^2 = \mathbb{V}ar\left[e_t\right]$ denotes the variance of the error. Then, absolute price informativeness, $\tau_{\hat{\Pi}}$, can be recovered by

$$\tau_{\hat{\Pi}} = \frac{\beta_1^2}{\sigma_e^2}. \tag{68}$$

The OLS estimation of Regression R1-Linear yields consistent estimates of β_1 and σ_e^2.

b) Relative Price Informativeness. Let $R^2_{X,X'}$ denote the R-squared of Regression R1-Linear. Let R^2_X, ζ, and ζ_0 respectively denote the R-squared and the coefficients of the following regression *of price differences on payoff differences,*

$$\Delta P_t = \overline{\zeta} + \zeta_0 \Delta X_t + e_t^{\zeta}. \tag{R2-Linear}$$

Then, relative price informativeness, $\tau_{\hat{\Pi}}^R$, can be recovered by

$$\tau_{\hat{\Pi}}^R = \frac{R_{X,X'}^2 - R_X^2}{1 - R_X^2}. \tag{69}$$

The OLS estimation of Regressions R1-Linear and R2-Linear yields consistent estimates of $R_{X,X'}^2$ and R_X^2.

Proof. a) By comparing Regression R1-Linear with the structural Equation (48), it follows that $\overline{\beta} = \overline{\phi} + \phi_n \mu_n$, $\beta_0 = \phi_0$, $\beta_1 = \phi_1$, and $e_t = \phi_n \varepsilon_t^n$. Consequently, $\sigma_e^2 = \mathbb{V}\mathrm{ar}\left[e_t\right] = \left(\phi_n\right)^2 \mathbb{V}\mathrm{ar}\left[\varepsilon_t^n\right] = \left(\phi_n\right)^2 \tau_n^{-1}$. Therefore, we can recover absolute price informativeness as follows

$$\tau_{\hat{\Pi}} = \frac{\left(\beta_1\right)^2}{\sigma_e^2} = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_n.$$

Given Equations (66) and (67), as well as the assumptions on u_t and n_t, it is straightforward to show that the OLS estimates of Regressions R1-Linear and R2-Linear are consistent, which implies that price informativeness can be consistently estimated as $\widehat{\tau_{\hat{\Pi}}} = \frac{\left(\widehat{\beta_1}\right)^2}{\widehat{\sigma_e^2}}$. Formally, $\operatorname{plim}\left(\widehat{\tau_{\Pi}}\right) = \operatorname{plim}\left(\frac{\left(\widehat{\beta_1}\right)^2}{\widehat{\sigma_e^2}}\right) = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_n = \tau_{\hat{\Pi}}$.

b) Note that the R-squareds of Regressions R1-Linear and R2-Linear can be expressed as follows

$$R_{X,X'}^2 = 1 - \frac{\mathbb{V}\mathrm{ar}\left(e_t\right)}{\mathbb{V}\mathrm{ar}\left(P_t\right)} \quad \text{and} \quad R_X^2 = \frac{\mathbb{V}\mathrm{ar}\left(\zeta_0 X_t\right)}{\mathbb{V}\mathrm{ar}\left(P_t\right)}.$$

After substituting Equation (66) in Equation (67), the following relation holds

$$P_t = \overline{\phi} + \phi_1 \mu_X + \phi_n \mu_n + \left(\phi_0 + \rho \phi_1\right) X_t + \phi_1 u_t + \phi_n \varepsilon_t^n. \tag{70}$$

By comparing Regression R2-Linear with the structural Equation (70), it follows that $\overline{\zeta} = \overline{\phi} + \phi_1 \mu_X + \phi_n \mu_n$, $\zeta_0 = \phi_0 + \rho \phi_1$, and $\varepsilon_t^{\zeta} = \phi_1 u_t + \phi_n \varepsilon_t^n$.

From Equation (70), the following variance decomposition must hold

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\left(P_t\right) &= \mathbb{V}\mathrm{ar}\left(\zeta_0 X_t\right) + \mathbb{V}\mathrm{ar}\left(\phi_1 u_t + \phi_n \varepsilon_t^n\right) \\
&= \mathbb{V}\mathrm{ar}\left(\zeta_0 X_t\right) + \left(\phi_1\right)^2 \mathbb{V}\mathrm{ar}\left(u_t\right) + \mathbb{V}\mathrm{ar}\left(e_t\right)
\end{aligned}$$

which can be rearranged to express $\frac{\tau_{\hat{\Pi}}}{\tau_u}$ as follows

$$1 = \underbrace{\frac{\mathbb{V}\text{ar}\left(\zeta_0 x_t\right)}{\mathbb{V}\text{ar}\left(P_t\right)}}_{R_X^2} + \underbrace{\frac{\mathbb{V}\text{ar}\left(e_t\right)}{\mathbb{V}\text{ar}\left(P_t\right)}}_{1-R_{X,X'}^2} \left(\underbrace{\frac{\left(\phi_1\right)^2}{\mathbb{V}\text{ar}\left(e_t\right)} \mathbb{V}\text{ar}\left(u_t\right)}_{\frac{\tau_{\hat{\Pi}}}{\tau_u}} + 1 \right) \Rightarrow \frac{\tau_{\Pi}}{\tau_u} = \frac{R_{X,X'}^2 - R_X^2}{1 - R_{X,X'}^2}$$

Therefore, relative price informativeness can be written as

$$\tau_{\hat{\Pi}}^R = \frac{\tau_{\hat{\Pi}}}{\tau_{\hat{\Pi}} + \tau_u} = \frac{1}{1 + \frac{1}{\frac{\tau_{\hat{\Pi}}}{\tau_u}}} = \frac{R_{X,X'}^2 - R_X^2}{1 - R_X^2}.$$

□

Microfoundation

Time is discrete, with periods denoted by $t = 0, 1, 2, \ldots, \infty$. Each period t, there is a continuum of investors, indexed by $i \in I$. Each generation lives two periods and has exponential utility over their last period wealth. An investor born at time t has preferences given by

$$U\left(W_{t+1}\right) = -e^{-\gamma W_{t+1}},$$

where γ is the coefficient of absolute risk aversion and W_{t+1} is the investor's wealth in his final period. There are two long-term assets in the economy: A risk-free asset in perfectly elastic supply, with return $R > 1$, and a risky asset in fixed supply Q that trades at price P_t in period t. The payoff of the risky asset each period t is given by

$$X_{t+1} = \mu_X + \rho X_t + u_t,$$

where μ_X is a scalar, $|\rho| < 1$, and $X_0 = 0$. The payoff X_t is realized and becomes common knowledge at the end of period $t-1$. The innovation in the payoff, u_t, and, hence, X_{t+1} are realized and observed at the end on period t. The innovations to the payoff are independently distributed over time.

To preserve tractability, we assume that investors' private trading needs arise from random heterogeneous priors– see Davila and Parlatore (2017) for a thorough analysis of this formulation. Formally, each investor i in generation t has a prior over the innovation at time t given by

$$u_t \sim_i N\left(\overline{n}_t^i, \tau_u^{-1}\right),$$

where

$$\overline{n}_t^i = n_t + \varepsilon_{\overline{n}t}^i \quad \text{with} \quad \varepsilon_{\overline{n}t}^i \overset{\text{iid}}{\sim} N\left(0, \tau_{\overline{n}}^{-1}\right)$$

and $n_t = \mu_n + \varepsilon_t^n$ with $\varepsilon_t^n \sim N\left(0, \tau_n^{-1}\right)$. n_t can be interpreted as the aggregate sentiment in the economy, where $n_t \perp \varepsilon^i_{\overline{n}t}$ for all t and all i. The aggregate sentiment n_t is not observed and acts as a source of aggregate noise in the economy, preventing the price from being fully revealing. For simplicity we assume $n_t \perp u_{t+s}$ for all t and all s. Moreover, we assume investors think of their prior as the correct one and do not learn about the aggregate sentiment from it.[13]

Each investor i in generation t receives a signal about the innovation in the asset payoff u_t given by

$$s^i_t = u_t + \varepsilon^i_{st} \quad \text{with} \quad \varepsilon^i_{st} \sim N\left(0, \tau_s^{-1}\right)$$

and $\varepsilon^i_{st} \perp \varepsilon^j_{st}$ for all $i \perp j$, and $u_t \perp \varepsilon^i_{st}$ for all t and all i.

Definition. The asset demand submitted by investor i born in period t is given by the solution to the following problem

$$\max_{q^i_t} \left(\mathbb{E}\left[X_{t+1} + R^{-1}p_{t+1}|\mathcal{I}^i_t\right] - P_t\right) Q^i_t - \frac{\gamma}{2}\mathbb{V}ar\left[X_{t+1} + R^{-1}P_{t+1}|\mathcal{I}^t_t\right]\left(Q^i_t\right)^2,$$

where $\mathcal{I}^i_t = \left\{X_t, s^i_t, \overline{n}^i_t, P_t\right\}$ is the information set of an investor i in period t.

The optimality condition for an investor i in period t satisfies

$$Q^i_t = \frac{\mathbb{E}\left[X_{t+1} + R^{-1}P_{t+1}|\mathcal{I}^i_t\right] - P_t}{\gamma \mathbb{V}ar\left[X_{t+1} + R^{-1}P_{t+1}|\mathcal{I}^i_t\right]}.$$

In a stationary equilibrium in linear strategies, we assume and subsequently verify that the equilibrium demand of investor i can be expressed as

$$\Delta Q^i_t = \alpha^i_X X_t + \alpha^i_s s^i_t + \alpha^i_n \overline{n}^i_t - \alpha^i_P P_t + \psi^i, \tag{71}$$

where α^i_θ, α^i_s, α^i_n, α^i_p, and ψ^i are individual equilibrium demand coefficients. Market clearing and the Strong Law of Large Numbers allows us to express the equilibrium price in period t as

$$P_t = \frac{\overline{\alpha_X}}{\overline{\alpha_P}}X_t + \frac{\overline{\alpha_s}}{\overline{\alpha_P}}u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_P}}n_t + \frac{\overline{\psi}}{\overline{\alpha_P}},$$

where we define cross sectional averages $\overline{\alpha_X} = \int \alpha^i_X di$, $\overline{\alpha_s} = \int \alpha^i_s di$, $\overline{\alpha_P} = \int \alpha^i_P di$, and $\overline{\psi} = \int \psi^i di - Q$.

The unbiased signal of the innovation in the payoff contained in the price is

$$\hat{\Pi}_t = \frac{\overline{\alpha_P}}{\overline{\alpha_s}}\left(P_t - \frac{\overline{\alpha_n}}{\overline{\alpha_s}}\mu_n - \frac{\overline{\alpha_X}}{\overline{\alpha_P}}X_t - \frac{\overline{\psi}}{\overline{\alpha_P}}\right) = u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_s}}\left(n_t - \mu_n\right),$$

where

$$\hat{\Pi}_t | X_{t+1}, X_t \sim N\left(u_t, \tau_{\hat{\Pi}}^{-1}\right),$$

[13]Davila and Parlatore (2017) show that the equilibirum structure is preserves if this assumption is relaxed.

with price informativeness given by

$$\tau_{\hat{\Pi}} = \left(\mathbb{V}\text{ar}\left[\hat{\Pi}_t | X_{t+1}, X_t\right]\right)^{-1} = \left(\frac{\overline{\alpha_s}}{\overline{\alpha_n}}\right)^2 \tau_n.$$

Given our guesses for the demand functions and the linear structure of prices we have

$$X_{t+1} + R^{-1} P_{t+1} = X_{t+1} + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}} X_{t+1} + R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}} u_{t+1} + R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}} n_{t+1} + R^{-1} \frac{\overline{\psi}}{\overline{\alpha_P}},$$

$$\begin{aligned}
\mathbb{E}\left[X_{t+1} + R^{-1} P_{t+1} | \mathcal{I}_t^i\right] &= \left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right) \mathbb{E}\left[X_{t+1} | \mathcal{I}_t^i\right] + R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}} \mathbb{E}\left[u_{t+1}\right] + R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}} \mathbb{E}\left[n_{t+1}\right] + R^{-1} \frac{\overline{\psi}}{\overline{\alpha_P}} \\
&= \left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right) \left(\rho X_t + \mathbb{E}\left[u_t | \mathcal{I}_t^i\right]\right) + R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_p}} \mathbb{E}\left[u_{t+1}\right] + R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}} \mu_n + R^{-1} \frac{\overline{\psi}}{\overline{\alpha_P}},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{V}ar\left[X_{t+1} + R^{-1} P_{t+1} | \mathcal{I}_t^i\right] &= \left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[X_{t+1} | \mathcal{I}_t^i\right] + \left(R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_{t+1}\right] + \left(R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[n_{t+1}\right] \\
&= \left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_t | \mathcal{I}_t^i\right] + \left(R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_{t+1}\right] + \left(R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[n_{t+1}\right].
\end{aligned}$$

Moreover, given the Gaussian structure of the signals in the information set, Bayesian updating implies

$$\mathbb{E}\left[u_t | s_t^i, \overline{n}_t^i, P_t\right] = \frac{\tau_s s_t^i + \tau_u \overline{n}_t^i + \tau_{\hat{\Pi}} \hat{\Pi}_t}{\tau_s + \tau_u + \tau_{\hat{\Pi}}} = \frac{\tau_s s_t^i + \tau_n \overline{n}_t^i + + \tau_{\hat{\Pi}} \frac{\overline{\alpha_P}}{\overline{\alpha_s}} \left(P_t - \frac{\overline{\alpha_n}}{\overline{\alpha_s}} \mu_n - \frac{\overline{\alpha_X}}{\overline{\alpha_P}} X_t - \frac{\overline{\psi}}{\overline{\alpha_P}}\right)}{\tau_s + \tau_u + \tau_{\hat{\Pi}}},$$

and

$$\mathbb{V}ar\left[u_t | \mathcal{I}_t^i\right] = \mathbb{V}ar\left[t_t | s_t^i, \overline{n}_t^i, P_t\right] = \left(\tau_s + \tau_u + \tau_{\Pi}\right)^{-1}.$$

Then, the first order condition is the given by is

$$Q_t^i = \frac{1}{\gamma} \frac{\left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right) \left(\rho X_t + \mathbb{V}ar\left[u_t | \mathcal{I}_t^i\right] \left(\tau_s s_t^i + \tau_u \overline{n}_t^i + \tau_{\hat{\Pi}} \hat{\Pi}_t\right)\right) + R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}} \mathbb{E}\left[u_{t+1}\right] + R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}} \mu_n + R^{-1} \frac{\overline{\psi}}{\overline{\alpha_P}} - P_t}{\left(1 + R^{-1} \frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_t | \mathcal{I}_{it}\right] + \left(R^{-1} \frac{\overline{\alpha_s}}{\overline{\alpha_P}}\right)^2 \mathbb{V}ar\left[u_{t+1}\right] + \left(R^{-1} \frac{\overline{\alpha_n}}{\overline{\alpha_P}}\right)^2 \tau_n^{-1}}.$$

Matching coefficients we have

$$
\begin{aligned}
\alpha_s^i &= \frac{\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa}\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_s \\
\alpha_n^i &= \frac{\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa}\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_\eta \\
\alpha_X^i &= \frac{\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa}\left(\rho-\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_{\hat{\Pi}}\frac{\overline{\alpha_X}}{\overline{\alpha_s}}\right) \\
\alpha_P^i &= \frac{1}{\kappa}\left(1-\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_{\hat{\Pi}}\frac{\overline{\alpha_p}}{\overline{\alpha_s}}\right) \\
\psi^i &= -\frac{1}{\kappa}\left(\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]\tau_{\hat{\Pi}}\left(\frac{\overline{\alpha_n}}{\overline{\alpha_s}}\mu_n+\frac{\overline{\psi}}{\overline{\alpha_s}}\right)-R^{-1}\left(\frac{\overline{\alpha_n}}{\overline{\alpha_P}}\mu_n+\frac{\overline{\psi}}{\overline{\alpha_P}}\right)\right),
\end{aligned}
\tag{72}
$$

where

$$
\kappa \equiv \gamma\left(\left(1+R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)^2\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]+\left(R^{-1}\frac{\overline{\alpha_s}}{\overline{\alpha_P}}\right)^2\mathbb{V}ar\left[u_{t+1}\right]+\left(R^{-1}\frac{\overline{\alpha_n}}{\overline{\alpha_P}}\right)^2\tau_n^{-1}\right),
$$

since $\mathbb{V}ar\left[u_t|\mathcal{I}_t^i\right]=\left(\tau_s+\tau_u+\tau_{\hat{\Pi}}\right)^{-1}$ for all i.

Then, an equilibrium in linear strategies always exists if the system above has a solution. Note that the demand sensitivities are the same for all i. Then, there exists a unique solution to the system in Equations (72), that is given by

$$
\begin{aligned}
\alpha_s^i &= \frac{1}{\kappa}\frac{1}{1-R^{-1}\rho}\frac{\tau_s}{\tau_u+\tau_s+\tau_{\hat{\Pi}}}, \quad \alpha_n^i = \frac{1}{\kappa}\frac{1}{1-R^{-1}\rho}\frac{\tau_\eta}{\tau_u+\tau_s+\tau_{\hat{\Pi}}} \\
\alpha_X^i &= \frac{1}{\kappa}\frac{\rho}{1-R^{-1}\rho}\frac{\tau_s}{\tau_s+\tau_{\hat{\Pi}}}, \quad \alpha_P^i = \frac{1}{\kappa}\frac{\tau_s}{\tau_s+\tau_{\hat{\Pi}}}, \quad \text{and} \\
\psi^i &= -\frac{\frac{1}{\kappa}\frac{1}{1-R^{-1}\rho}\left(\left(1-R^{-1}\right)\tau_{\hat{\Pi}}-R^{-1}\tau_s\right)\frac{\tau_u}{\tau_u+\tau_s+\tau_{\hat{\Pi}}}\mu_n}{1+\left(1-R^{-1}\right)\tau_{\hat{\Pi}}-R^{-1}\tau_s},
\end{aligned}
$$

where

$$
\tau_{\hat{\Pi}} = \left(\frac{\tau_s}{\tau_u}\right)^2\tau_n
$$

and

$$
\kappa = \gamma\left(\left(\frac{1}{1-R^{-1}\rho}\right)^2\frac{1}{\tau_u+\tau_s+\tau_{\hat{\Pi}}}+\left(R^{-1}\frac{1}{1-R^{-1}\rho}\frac{\tau_s+\tau_{\hat{\Pi}}}{\tau_u+\tau_s+\tau_{\hat{\Pi}}}\right)^2\tau_u^{-1}+\left(\frac{R^{-1}}{1-R^{-1}\rho}\frac{\tau_s+\tau_{\hat{\Pi}}}{\tau_u+\tau_s+\tau_{\hat{\Pi}}}\frac{\tau_u}{\tau_s}\right)^2\tau_n^{-1}\right).
$$

In this equilibrium, our guess in Eq. (71) is verified and the equilibrium price is linear and can be expressed as in Eq. (67).

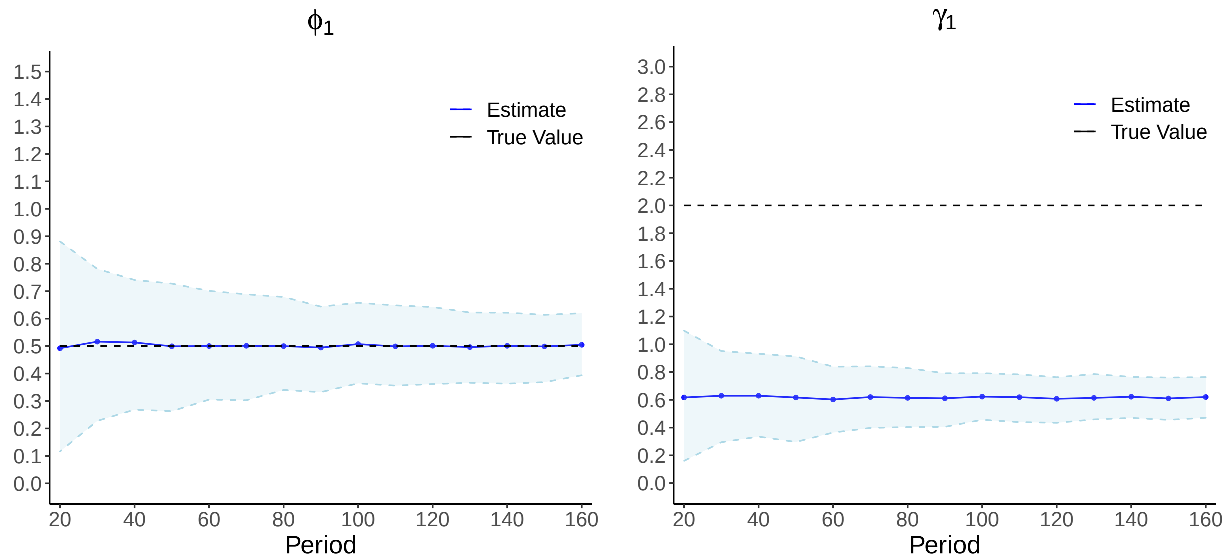

Figure OA-15: Predictability vs. Informativeness: biased estimates (1)

Note: The left panel in Figure OA-15 shows the distribution of estimates ϕ_1 and its true value when estimating Regression R1 via OLS. The right panel in Figure OA-15 shows the distribution of estimates $\frac{1}{\gamma_1}$ and its true value when running the reverse (predictive) Regression R3. In both panels, for every number of periods between 20 to 160, we simulate the model $N = 500$ times, and report the mean estimate and the 5% and 95% estimates. The dashed black lines respectively show the true values of ϕ_1 and γ_1. The solid blue lines show the mean of the OLS estimates in each case, while the shaded area includes estimates within the 5% and 95% percentile.

D Relation to Existing Work

D.1 Predictability vs. Informativeness

As explained in the paper, price informativeness and price/return predictability are conceptually different notions. In this section, we provide a simple illustration of the source of the bias that would arise if one were to run Regression R3.[14] We simulate the model of Section 2 using the following parameters: $\mu_{\Delta x} = \mu_n = \rho = \overline{\phi} = \phi_0 = 0$, $\tau_u = \tau_n = 1$, $\phi_1 = 0.5$, and $\phi_n = 0.75$.

Figure OA-15 shows the distribution of the OLS estimates of β_1 in Regression R1 (equivalently, ϕ_1 in Equation 2) for simulation with different time periods. As shown in Proposition 1, the estimates of ϕ_1 are consistent. Figure OA-15 also shows the distribution of the OLS estimates of $\gamma_1 = \frac{1}{\beta_1}$ in Regression R3. As described in the text, the estimates of γ_1 are downward biased and non-consistent.

Figure OA-16 provides an alternative illustration of the bias in this scenario. The left panel in Figure OA-16 shows a simulation of Regression R1 with 160 observations. The estimated slope of this regression provides a consistent estimate of the estimate of ϕ_1. The right panel in Figure OA-16 shows a simulation of Regression R3 for the exact same simulation. The estimated slope of this regression provides a biased estimate of γ_1 in Regression R3, as shown theoretically in the text.

[14]Note that the bias identified here, which arises due to the presence of noise, is different from Stambaugh (1999) bias, which arises whenever the predictive variables is persistent.

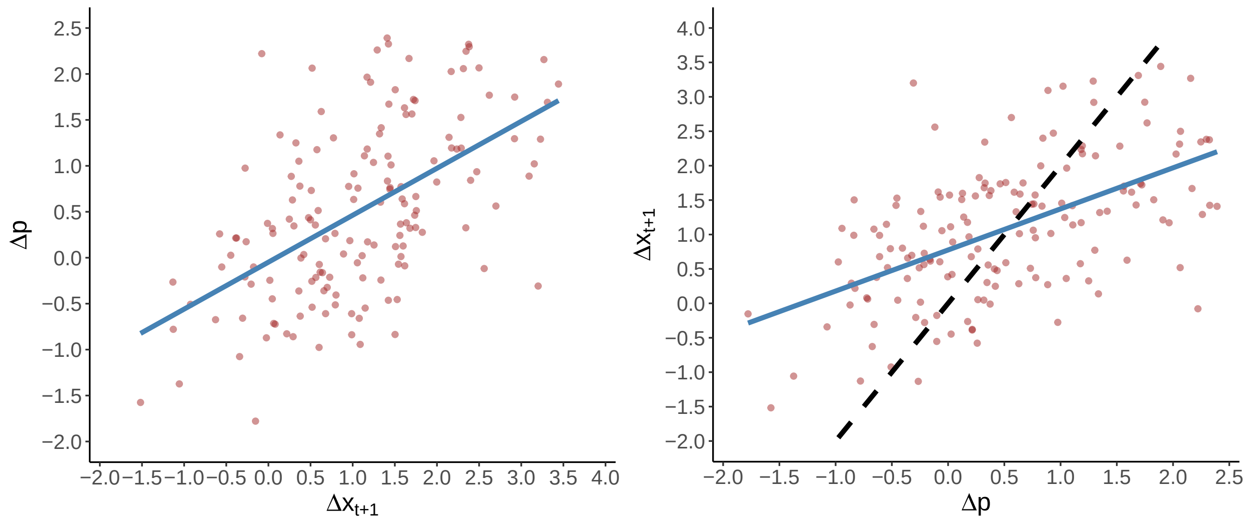

Figure OA-16: Predictability vs. Informativeness: biased estimates (2)

Note: The left panel in Figure OA-16 shows a simulation of Regression R1 with 160 observations. The estimated slope of this regression provides a consistent estimate of the estimate of ϕ_1. The right panel in Figure OA-16 shows a simulation of Regression R3 for the exact same simulation. The estimated slope of this regression provides a biased estimate of γ_1 in Regression R3. The OLS regression is represented with a solid blue line in both panels. The dashed black line in the right panel shows the correct relation between both variables.

D.2 Forecasting Price Efficiency

In this section, we compare forecasting price efficiency (FPE), i.e., the unconditional variance of the expected value of the fundamental conditional on the price, with price informativeness. While higher price informativeness will lead to higher FPE, higher FPE may not reflect an increase in price informativeness. More specifically, we show that FPE confounds changes in the volatility of the fundamental with changes in the ability of markets to aggregate dispersed information. Hence, FPE can increase because price informativeness increases or because the fundamental becomes more volatile and harder to predict. Alternatively, absolute price informativeness is the precision of the unbiased signal about the fundamental contained in the price. This precision is a direct measure of the ability of financial markets to aggregate dispersed information and it is independent of the volatility of the fundamental.

To illustrate this point, we explicitly re-derive the model in Bai, Philippon and Savov (2015) (BPS) as a special case of our general framework.[15] Consistent with our approach, we abstract from investment decisions, and exclusively focus on the role of financial markets aggregating information. We first describe the environment in BPS using our notation to show how it is nested in our general

[15] As described in the introduction, BPS estimate FPE running cross-sectional regressions at specific points in time and report the time-series evolution of their cross-sectional estimates of FPE. This approach implicitly assumes that the data generating process (including the distribution of payoffs, signals, and noise) is the same for all stocks at a given point in time. We instead recover a panel of stock-specific measures of price informativeness by using rolling regressions. While they run predictive regressions of future fundamentals on current market values, we show that in order to have consistent estimates of price informativeness, one must regress price changes on future payoffs – see previous section.

specification. Then, we show that while FPE is relevant for welfare, it does not disentangle the ability of markets to aggregate information from how easy it is to forecast the fundamental.

Environment There are two periods, $t = 0, 1$. There is one asset with a payoff $u \sim N\left(\overline{u}, \tau_u^{-1}\right)$. There are $i = 1, ..., I$ informed traders who choose their demand q_{1i} to maximize mean variance preferences with imperfect information about u. The asset payoff u is not observable. However, investors observe a private signal

$$s = u + \varepsilon_s$$

and a public signal

$$\chi = u + \varepsilon_\chi,$$

where $\varepsilon_s \sim N\left(0, \tau_s^{-1}\right)$, $\varepsilon_\chi \sim N\left(0, \tau_\chi^{-1}\right)$, and $\varepsilon_s \perp \varepsilon_\pi$. Note that all informed investors observe the same set of signals. There are N noise traders whose total demand is random and given by $n \sim N\left(0, \tau_n^{-1}\right)$.

The informed traders' problem is

$$\max_{\theta_{1i}} \left(\mathbb{E}\left[u|s,\chi\right] - p\right)\theta_{1i} - \frac{\gamma}{2}\mathbb{V}ar\left[u|s,\chi\right]\theta_{1i}^2 + p\theta_{0i}$$

which leads to the following demand curve

$$\theta_{1i} = \frac{\mathbb{E}\left[u|s,\chi\right] - p}{\gamma \mathbb{V}ar\left[u|s,\chi\right]},$$

where

$$\mathbb{E}\left[u|s,\chi\right] = \frac{\tau_u \overline{u} + \tau_s s + \tau_\chi \chi}{\tau_u + \tau_s + \tau_\chi} \quad \text{and} \quad \mathbb{V}ar\left[u|s,\chi\right] = \frac{1}{\tau_u + \tau_s + \tau_\chi}.$$

Since all informed investors share the same information set, there is no learning from the price.

In an equilibrium in linear strategies demands for informed traders and uninformed traders are respectively given by

$$\begin{aligned} \theta_{1i}^I &= \alpha_s^I s + \alpha_\chi^I \chi + \alpha_n^I n - \alpha_p^I p + \psi^I \\ \theta_{1i}^U &= \alpha_s^U s + \alpha_\chi^U \chi + \alpha_n^U n - \alpha_p^U p + \psi^U. \end{aligned}$$

Matching coefficients we have that

$$\alpha_s^I = \frac{\tau_s}{\gamma}, \quad \alpha_\chi^I = \frac{\tau_\chi}{\gamma}, \quad \alpha_n^I = 0, \quad \alpha_p^I = \frac{1}{\gamma}\left(\tau_u + \tau_s + \tau_\chi\right), \quad \psi^I = \frac{\tau_u}{\gamma}\overline{u} - \theta_{0i},$$

and $\alpha_s^U = \alpha_\chi^U = \alpha_p^U = \psi^U = 0$, and $\alpha_n^U = \frac{1}{N}$.

Market clearing implies

$$\sum_{i=1}^{I} \theta_{si}^I + n = Q,$$

which is the same as

$$p = \frac{\overline{\alpha_s}}{\overline{\alpha_p}} s + \frac{\overline{\alpha_\chi}}{\overline{\alpha_p}} \chi + \frac{\overline{\psi}}{\overline{\alpha_p}} + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} n,$$

where $\overline{\alpha_s} = I\alpha_s^I + N\alpha_s^N$. $\overline{\alpha_\chi} = I\alpha_\chi^I + N\alpha_\chi^N$, $\overline{\alpha_p} = I\alpha_p^I + N\alpha_p^N$, $\overline{\alpha_n} = I\alpha_n^I + N\alpha_n^N$, and $\overline{\psi} = I\psi^I + N\psi^N - Q$.

Price informativeness and forecasting price efficiency Observing the price is equivalent to observing

$$\pi = \frac{\overline{\alpha_p}}{\overline{\alpha_s} + \overline{\alpha_\chi}} \left(p - \frac{\overline{\psi}}{\overline{\alpha_p}} \right) = \theta + \frac{\overline{\alpha_s}}{\overline{\alpha_s} + \overline{\alpha_\chi}} \varepsilon_s + \frac{\overline{\alpha_\chi}}{\overline{\alpha_s} + \overline{\alpha_\chi}} \varepsilon_\chi + \frac{\overline{\alpha_n}}{\overline{\alpha_s} + \overline{\alpha_\chi}} n,$$

where

$$\pi | \theta \sim N \left(\theta, \tau_\pi^{-1} \right)$$

with

$$\tau_\pi = \left(\frac{\overline{\alpha_s}}{\overline{\alpha_s} + \overline{\alpha_\chi}} \right)^2 \tau_s^{-1} + \left(\frac{\overline{\alpha_\chi}}{\overline{\alpha_s} + \overline{\alpha_\chi}} \right)^2 \tau_\chi^{-1} + \left(\frac{\overline{\alpha_n}}{\overline{\alpha_s} + \overline{\alpha_\chi}} \right)^2 \tau_n^{-1}. \tag{73}$$

τ_π in Equation (73) corresponds to our measure of absolute price informativeness when there is a finite number of investors. There are two differences with respect to the baseline model presented in the main text. First, there are multiple sources of aggregate noise: the error of the private signal, ε_s; the error of the public signal, ε_χ; and the demand of noise traders, n. Second, price informativeness is modulated by $\overline{\alpha_s} + \overline{\alpha_\chi}$ instead of by $\overline{\alpha_s}$ because there are two sources of external information about the fundamental θ.

A Bayesian external observer who only observes the price, learns from the price in the following way

$$\mathbb{E}\left[u | \pi \right] = \frac{\tau_u \overline{u} + \tau_\pi \pi}{\tau_u + \tau_\pi}.$$

Forecasting price efficiency (FPE) is then given by

$$\begin{aligned} \mathcal{V}_{FPE} = \mathbb{V}ar\left(\mathbb{E}\left[u | \pi \right] \right) &= \left(\frac{\tau_\pi}{\tau_u + \tau_\pi} \right)^2 \left(\tau_u^{-1} + \tau_\pi^{-1} \right) \\ &= \frac{\tau_\pi}{\tau_u + \tau_\pi} \tau_u^{-1}. \end{aligned} \tag{74}$$

The expression for FPE in Equation (74) is the predicted variance of cash flows u from prices. From this equation, it is easy to see that FPE confounds two effects. FPE can increase due to changes in the ability of prices to aggregate information, τ_π, or due to changes in the variability of the fundamental, τ_u. Hence, conditional on the variance of the fundamental remaining constant, FPE and price informativeness will co-move. However, without controlling for changes in fundamental volatility, one cannot make any inferences about price informativeness by looking at FPE.