Identifying Price Informativeness*

Eduardo Dávila† Cecilia Parlatore‡

December 2023

Abstract

This paper shows how to identify and estimate price informativeness, a necessary step to test theories of information aggregation. Starting from a pricing equation and a stochastic process for payoffs, we show how to recover relative price informativeness from regressions of asset price changes on changes in asset payoffs. Applying our identification results, we estimate a panel of stock-specific measures of informativeness for U.S. stocks. In the cross-section, large, high turnover, high institutional-ownership, and high analyst-coverage stocks have higher informativeness. In the time-series, the median, mean, and standard deviation of the distribution of informativeness have steadily increased since the mid-1980s.

JEL Classification: G14, E44, D82

Keywords: price informativeness, informational efficiency, information aggregation

*This a substantially revised version of earlier drafts. We would like to thank our discussant Xavier Vives, as well as Yakov Amihud, Philip Bond, Ricardo Caballero, John Campbell, Jennifer Carpenter, John Cochrane, Darrell Duffie, Emmanuel Farhi, Itay Goldstein, Gary Gorton, Piero Gottardi, Marcin Kacperczyk, Ralph Koijen (editor), Arvind Krishnamurthy, Guido Lorenzoni, Harry Mamaysky, Stephen Morris, Stefan Nagel, Guillermo Ordoñez, Monika Piazzesi, Martin Schneider, David Sraer, David Thesmar, Aleh Tsyvinski, Harald Uhlig, Laura Veldkamp, Annette Vissing-Jørgensen, Wei Xiong, and Moto Yogo, 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, Virtual Finance Theory, HEC, Bank of Canada, London Business School, and Imperial College for helpful comments and discussions. We are also thankful to Alexi Savov for comments and discussions, and for sharing his code. Dean Parker, Luke Min and Jack Kelly provided outstanding research assistance.
First Draft: December 2017.

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

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

1 Introduction

Financial markets play an important role by aggregating 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 that treats asset prices as a signal about future fundamentals is typically traced back to Hayek (1945).

Despite the substantial theoretical literature that studies how prices aggregate information, the connection between the theoretical and empirical research on price informativeness remains understudied. In particular, most existing empirical work on the informational content of prices typically focuses on predictability/forecasting measures, which fail to isolate the role of prices as signals about fundamentals. Moreover, the existing literature lacks formal identification results. This paper seeks to fill both gaps.

In this paper, we show how to identify and estimate relative price informativeness, a notion of informativeness that formally corresponds to the relative precision of the signal about future payoffs contained in asset prices. To derive our results, we only need to postulate i) an asset pricing equation and ii) a stochastic process for asset payoffs. Our main results show that a specific combination of R-squareds from linear regressions of changes in asset prices on changes in asset payoffs exactly identifies relative price informativeness within a class of models that may feature heterogeneity across investors' preferences, endowments, private signals, and private trading needs; competitive or strategic market structures; symmetric or asymmetric information; and that require minimal distributional assumptions. Our approach has the potential to provide model-based tests of predictions in widely used dispersed information models.

The main contribution of this paper is an identification result and an estimation procedure that determine how much information about future payoffs is contained in asset prices. Our approach allows us to answer questions that cannot be answered with the existing measures of informational efficiency. For example, suppose one is interested in whether asset prices are a good guide for allocating real capital to firms. In this case, measures related to the predictability of future earnings may give the wrong answer if prices are very noisy but future earnings are very responsive to earnings news. Our identification and estimation results correctly account for the noise in prices, giving an appropriate answer to the question.

Alternatively, one can be interested in the effect that changes in the market structure or financial regulation have on the ability of markets to aggregate information, usually central issues in policy debates. This matter cannot be settled by looking at forecasting measures and requires measuring the informational content of prices directly, as we do in this paper. Perhaps more importantly, the ability to measure price informativeness in a model-consistent way opens the door to testing and disciplining a variety of theories of information aggregation. We hope that showing how to identify and estimate price informativeness leads the way to making empirically-based statements about social welfare in the future.

We begin by formally defining two price informativeness measures with desirable properties: absolute and relative price informativeness. Absolute price informativeness, which formally 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, making it easy to interpret and to compare 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 an 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% on the price signal (and a weight of 80% on the prior) when forming a posterior belief over the future payoff.

We succinctly describe our approach to identifying and estimating (relative) price informativeness about the one-period ahead payoff growth in a simplified environment without public signals. Consider the following two regressions that relate log-price changes, Δp_t, to the contemporary and future differences in log-asset payoffs, denoted by Δx_t and Δx_{t+1}, respectively:

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

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

where we denote the R-squareds of Regressions R1 and R2 by $R^2_{\Delta x, \Delta x'}$ and $R^2_{\Delta x}$, respectively. Our main result shows 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 (OLS) 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 identification results are therefore agnostic about the nature of the noise in asset prices.

While it may seem straightforward to identify and estimate price informativeness in the stylized environment implied by equations R1 and R2, it is harder to do so in more complex scenarios. In fact, in our baseline environment, we consider a framework with an arbitrary set of public signals and a nonzero correlation between payoff- and non-payoff related trading motives (Proposition 1). In extensions to our baseline result, we also allow for the possibility that payoffs have an unlearnable component (Proposition 2) and provide results to recover price informativeness about payoff growths at longer horizons (Proposition 3).[1]

Even though we show that price informativeness can be recovered without fully specifying the model primitives, a microfounded model is necessary to understand the link between price informativeness and the primitives in the economy. For this reason, we develop several microfounded dynamic models of trading that are consistent with the asset pricing equation and the stochastic process for asset payoffs that we use to derive 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). Subsequently, we study a representative agent model similar to those used in the macro-finance literature. Finally, we study a model with informed and uninformed investors, as in the classic literature on information and learning. Analyzing these applications has a dual purpose. First, these applications show that our identification results apply to economies i) with or without dispersed information among investors, ii) with time-varying risk aversion/risk-premia, iii) with investors who may or may not learn from prices, and iv) in which noise may arise from different sources. Second, these applications allow us to provide a structural economic interpretation of the empirical results presented in Section 4.

Next, we use our identification results to estimate relative price informativeness. First, we recover a quarterly panel of stock-specific measures of price informativeness between 1985 and 2017 by running rolling time-series regressions of the form implied by our baseline environment at the stock level using year-on-year changes with overlapping 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 given by 4.47% and 8.54%, respectively. Our estimation exercise allows us to uncover both cross-sectional and time-series patterns in price informativeness. In the cross-section, we find that stocks that i) are larger, ii) turn over more quickly, iii) have a higher institutional ownership share, and iv) have a higher coverage by analysts have higher price informativeness. We also find that when we control for the size of the stocks, the cross-sectional results on institutional ownership and analyst

[1] In previous versions of this paper, we also considered environments with multiple risky assets and strategic investors.

coverage attenuate substantially. In the time series, we find that the median and mean price informativeness have steadily increased since the mid-1980s. The standard deviation of price informativeness has also increased over this period.

We then present estimation results in the context of our model with payoffs that have an unlearnable component. In this case, the distribution of informativeness across stocks is also right-skewed. Time-series averages of the median and mean levels of price informativeness across all stocks and years are given by 4.67% and 9.29%, respectively. Consistent with our baseline results, we find increasing time trends in the mean and median estimates of price informativeness, as in the cross-sectional standard deviation. Finally, we extend our empirical analysis to consider price informativeness about payoff growth rates at longer horizons and find that price informativeness about future payoff growth decreases with the horizon of interest.

Our results by no means settle the question of how to identify and estimate price informativeness, as we discuss in Section 5 before concluding. In particular, there is scope to derive new identification results in more general environments, such as those with feedback effects or significant non-linearities, and to overcome some limitations of our measurement exercise by studying environments with better data availability.

Related Literature

Our theoretical framework takes as starting point 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), and De Long et al. (1990), among others.[2] Building on this literature, our results show how to identify and consistently estimate relative price informativeness, which is a notion of informativeness that formally corresponds to the relative precision of the signal about future payoffs contained in asset prices.

Despite the substantial theoretical literature that has studied learning in financial markets, there has been less interest in measuring the precision of prices as signals over future payoffs. There is a literature that has proposed empirical measures to capture the informational content of prices. These empirical measures have been inspired by economic models to different degrees. Influenced by the predictions of the CAPM/APT frameworks and following the prominent Roll (1988) presidential address, Morck, Yeung and Yu (2000) studies regressions of asset returns on 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 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,

[2]See Vives (2008) and Veldkamp (2011) for systematic reviews of this body of work.

Morck and Yeung (2004) documents a positive correlation between price nonsynchronicity and corporate investment. Chen, Goldstein and Jiang (2006) establishes that there exists a positive relationship 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 in Easley, O'Hara and Paperman (1998), estimates the probability of an informed trade using high-frequency data through the lens of a model with informed and uninformed traders. Hou and Moskowitz (2005) and Weller (2018) also propose alternative empirical measures. In particular, Weller (2018) uses a price jump ratio to measure how much information enters prices relative to how much is potentially acquirable at the stock level, finding that algorithmic trading decreases the amount of information incorporated into prices. When compared to this body of work, a central contribution of our paper lies in providing identification results within a general environment that encompasses the class of the models most frequently used in the theoretical literature on learning. By doing this, our results have the potential to directly discipline theories of information and learning in financial markets.

The interpretation of the results of the empirical literature mentioned above has not gone unchallenged. For instance, Hou, Peng and Xiong (2013) highlights that a measure like Roll's R^2 (price nonsynchronicity) lacks a structural interpretation, questioning the link between return price nonsynchronicity and notions of price informativeness from theoretical and empirical perspectives. Our results in this paper address the Hou, Peng and Xiong (2013) critique by first defining and justifying a theoretical notion of price informativeness, and then showing how to identify it and consistently estimate it.

Our work is also related to that of Bai, Philippon and Savov (2016), who consider 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 or show how to formally identify and estimate price informativeness in the context of a structural model, which is a central contribution of our paper. Our results and the recent contemporaneous work of Farboodi et al. (2020) and Kacperczyk, Sundaresan and Wang (2020) 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. (2020) seeks to understand how changes in data processing over time have altered the amount of information (signal) incorporated in asset prices. Using our measure of price informativeness as an input in their analysis, they conclude 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 (2020) finds a positive relationship between price informativeness and the ownership share of foreign institutional investors, using both empirical measures of informativeness and the identification results developed in this paper. We discuss how our results relate to existing work in more detail in Section 2.7.

2 General Framework

In this section, we first define and justify the notion of relative price informativeness, and then show how to formally identify it and estimate it by specifying an asset pricing equation and a stochastic process for asset payoffs. To align our results to the empirical implementation in Section 4, we derive the main results in the body of the paper in a log-difference-stationary environment, which allows us to sidestep concerns associated with nonstationarity.[3]

2.1 Baseline 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 $\mathbb{V}\mathrm{ar}[u_t] = \sigma_u^2 = \tau_u^{-1}$, 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 $t+1$ — to indicate that investors can potentially learn about the realization of u_t at date t.

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

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

where the coefficients $\overline{\phi}$, ϕ_0, ϕ_1, and ϕ_n are scalar parameters and ϕ_χ is a vector of N parameters. We denote the vector of changes in the N public signals observed by investors by $\Delta \chi_t$, where

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

where $\overline{\omega}$ is a $N \times 1$ vector and $\overline{\varepsilon}_t^{\Delta \chi}$ is a $N \times 1$ random vector that has mean zero, finite variance, and is i.i.d. across time and independent of the innovations u_t. We denote investors' trading

[3] In the Internet 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.

motives that are not coming from their information by Δn_t, and allow them to be correlated with the innovation to the payoff as follows:

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

where $\varepsilon_t^{\Delta n}$ has mean zero and finite variance $\tau_{\Delta n}^{-1}$ and is i.i.d. across time and independent of the innovations u_t. 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.

In Section 3, we show that Equation (2) emerges endogenously as the solution to several fully specified dynamic models of trading. In that case, the model parameters can be mapped to specific combinations of primitives. We conclude the description of the environment with the following remark.

Remark 1. *Features of the environment.* Note that our baseline model allows for an arbitrary set of public signals and for a nonzero correlation between payoff- and non-payoff related trading motives. In Sections 2.4 and 2.5, we further allow for the possibility that payoffs have an unlearnable component and that the price may contain information about longer horizons.

2.2 Price Informativeness: Definition and Justification

Within the environment introduced in Section 2, we now formally define and justify two related measures of informativeness: absolute and relative price informativeness. We focus on these notions because of their desirable properties, as we explain below. Both are commonplace in the theoretical literature on information and learning. Absolute price informativeness is discussed in Section 4 of Vives (2008), while relative price informativeness corresponds to the exact notion of informativeness used in Grossman and Stiglitz (1980).[4] The contribution of this paper is to provide formal identification results of both notions, and to use such results for estimation. In Remark 6 below, we discuss the relation of these measures of price informativeness to others.

Formally, in our context, the unbiased signal of the innovation to future payoffs u_t contained in the price is the key variable of interest from the perspective of understanding how informative asset prices are about future payoffs.[5] This endogenous unbiased signal, which we denote by π_t,

[4] Relative price informativeness, as defined in Equation (4) below, exactly corresponds to Equation (17) in Grossman and Stiglitz (1980).

[5] In principle, one could study whether prices are informative about other variables of interest. Studying informativeness about future payoffs seems like a natural starting point, although in an earlier version of this paper we explained how to adapt our approach to identify informativeness about future prices.

is given by

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

Given Equation (2),

$$\pi_t = u_t + \frac{\phi_n}{\phi_1 + \phi_n \omega} \left(\Delta n_t - \mu_{\Delta n} \right). \tag{3}$$

defines an endogenous unbiased signal about u_t, where $\frac{\phi_n}{\phi_1+\phi_n\omega}(\Delta n_t - \mu_{\Delta n})$ acts as the noise contained in the price signal. This signal π_t is unbiased because $\mathbb{E}[\pi_t | u_t, \Delta x_t, \Delta \chi_t] = u_t$. Both absolute and relative price informativeness, which we define and justify next, are the relevant measures for an external observer who uses the asset price as a signal to learn about future asset payoffs.

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 the vector of public signals $\Delta \chi_t$. Given Equation (2), it is formally given by

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

where $\tau_{\Delta n} = \mathbb{V}\mathrm{ar}[\Delta n_t]^{-1}$.

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 the vector of public signals $\Delta \chi_t$. Given Equation (2), it is formally given by

$$\tau_\pi^R \equiv \frac{\tau_\pi}{\tau_\pi + \tau_{u|\Delta \chi_t}}, \tag{5}$$

where $\tau_{u|\Delta\chi} = \mathbb{V}\mathrm{ar}[u_t | \Delta x_t, \Delta \chi_t]^{-1}$.

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 another 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.

Intuitively, absolute price informativeness measures the signal-to-noise ratio contained in the asset price. If the price is very responsive to Δx_{t+1}, perhaps because investors trade with very precise information about the future payoff, so ϕ_1 is high, or their trading motives not coming from their information are highly correlated to the future payoff, so $\phi_n \omega$ is high, 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 change in the asset price Δp_t. On the contrary, when price informativeness is low, an external observer learns little about future payoffs by observing the change in the asset price Δp_t.

The definition of *relative price informativeness* corrects absolute price informativeness to account for the variability of the payoff, 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 who only learns from the price, 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. If an external observer had additional information about the future payoff in addition to the price, the Kalman gain that we identify would be an upper bound to the one used by such external observer.

Relative price informativeness is an appealing object because it provides a bounded (between 0 and 1), unit-free measure of informativeness that facilitates 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 comparisons of absolute price informativeness do not have a clear interpretation. In Remark 6 below, we further explain how absolute and relative price informativeness relate to other notions of informativeness 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 typically 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 changes in asset prices on realized and future changes in asset payoffs to recover price informativeness.

Proposition 1. (Identifying price informativeness) *Let $\overline{\beta}$, β_0, β_1, and β_2 denote the coefficients of the following regression of log-price differences on realized and future log-payoff* *differences:*

$$\Delta p_t = \overline{\beta} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t+1} + \beta_2 \cdot \Delta \chi_t + e_t, \tag{R1}$$

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,* $\Delta \chi_t$ *denotes the change in public information, and where* $R^2_{\Delta x, \Delta x'}$ *denotes the R-squared of Regression R1. Let* $\overline{\zeta}$, ζ_0, *and* ζ_2 *denote the coefficients of the following regression of log-price differences on realized log-payoff differences:*

$$\Delta p_t = \overline{\zeta} + \zeta_0 \Delta x_t + \zeta_2 \cdot \Delta \chi_t + e^{\zeta}_t, \tag{R2}$$

where $R^2_{\Delta x}$ *denotes the R-squared of Regression R2. Then, relative price informativeness,* τ^R_π, *defined in Equation (5), can be recovered as*

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

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^R_π. 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.

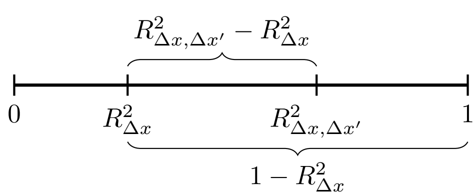


Figure 1: Interpreting relative price informativeness

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

Figure 1 illustrates how to interpret Equation (6). The denominator $1 - R^2_{\Delta x}$ can be interpreted as the 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 percentage reduction in uncertainty about future payoffs after 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 $\tau^R_\pi \in [0,1]$.

As we show in the Appendix, if all random variables in the model are Gaussian, a Bayesian external observer who only learns from the price has the following posterior distribution over u_t:

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

where π_t, τ_π, and τ^R_π 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 who only learns from the price about the innovation to the future payoff is reduced by 15% after learning from the price — this interpretation follows from the fact that $(\tau_\pi + \tau_u)^{-1} = \left(1 - \tau^R_\pi\right) \tau_u^{-1}$.

Even though we emphasize the economic identification of price informativeness, we also address how to recover consistent estimates. When the external observer has access to public information, the coefficient estimates in Regression R2 will be biased. The proof of Proposition 1 takes into account this bias and provides consistent estimates of relative price informativeness as defined in Equation 5.

2.4 Unlearnable Payoff Component

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

$$\Delta x_{t+1} = \mu_{\Delta x} + u_t, \tag{8}$$

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

$$u_t = u^L_t + u^U_t,$$

where u^L_t and u^U_t are the learnable and unlearnable components of the innovation. These innovations have mean zero and variances given by

$$\mathbb{V}\text{ar}\left[u^L_t\right] = \left(\tau^L_u\right)^{-1} \quad \text{and} \quad \mathbb{V}\text{ar}\left[u^U_t\right] \sim \left(\tau^U_u\right)^{-1},$$

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. Moreover, as in Section 2, $\Delta n_t \equiv n_t - n_{t-1}$ represents the change in the aggregate component of investors' trading motives that are not based on information about payoffs, given by $\Delta n_t = \mu_{\Delta n} + \omega u_t^L + \varepsilon_t^{\Delta n}$, where $\mathbb{V}\text{ar}\left[\varepsilon_t^{\Delta n}\right] = \sigma_{\Delta n}^2 = \tau_{\Delta n}^{-1}$.

In this case, the price process is given by

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \phi_1 \Delta u_t^L + \phi_\chi \Delta \chi_t + \phi_n \Delta n_t$$

and the unbiased signal contained in the price about the learnable component of next period's payoff is given by

$$\pi_t^L = u_t^L + \frac{\phi_n}{\phi_1 + \phi_n \omega} \varepsilon_t^{\Delta n}.$$

Then, absolute and relative price informativeness about u_t^L are respectively given by

$$\tau_{\pi^L} \equiv \left(\mathbb{V}\text{ar}\left[\frac{\phi_n}{\phi_1}\varepsilon_t^{\Delta n} \middle| u_t^L, u_{t-1}^L, \Delta\chi_t, \Delta x_t\right]\right)^{-1} \quad \text{and} \quad \tau_{\pi^L}^R \equiv \frac{\tau_{\pi^L}}{\tau_{u_t^L|\Delta\chi} + \tau_{\pi^L}}. \tag{9}$$

These definitions are analogous to the ones in Equations (4) and (5) in the baseline case with the only difference that the future payoff is replaced by the learnable component of the future payoff, u_t^L, and the learnable component of the current payoff, u_{t-1}^L, is part of the information set. This definition of π_t^L assumes that u_{t-1}^L is observed at date t. This is consistent with our empirical implementation in which we map u_t^L with the average forecast of earnings growth at date $t+1$ between t and $t+1$. From this observation it follows that Proposition 1 can be extended to the case in which there is an unlearnable component of the payoff as follows.

Proposition 2. (Identifying price informativeness with unlearnable component) *Let $\overline{\beta}^u$, β_0^u, β_1^u, β_2^u, and β_3^u denote the coefficients of the following regression of log-price differences on realized log-payoff differences and log-forecast earnings differences:*

$$\Delta p_t = \overline{\beta}^u + \beta_0^u \Delta x_t + \beta_1^u u_{t-1}^L + \beta_2^u u_t^L + \beta_3^u \cdot \Delta\chi_t + e_t^u, \tag{R1U}$$

where $\Delta p_t = p_t - p_{t-1}$ denotes the date t change in log-price, $\Delta x_t = x_t - x_{t-1}$ denotes the date t log-payoff difference, u_t^L denotes the date t log-forecast earnings growth, and where $R_{u,u'}^2$ denotes the R-squared of Regression R1U. Let $\overline{\zeta}^u$, ζ_0^u, ζ_1^u, and ζ_3^u denote the coefficients of *the following regression of log-price differences on realized log-payoff differences and lagged log-forecast earnings:*

$$\Delta p_t = \overline{\zeta}^u + \zeta_0^u \Delta x_t + \zeta_1^u u_{t-1}^L + \zeta_3^u \cdot \Delta \chi_t + e_t^{\zeta u}, \tag{R2U}$$

where R_u^2 *denotes the R-squared of Regression R2U. Then, relative price informativeness,* $\tau_{\pi^L}^R$*, defined in Equation (9), can be recovered as*

$$\tau_{\pi^L}^R = \frac{R_{u,u'}^2 - R_u^2}{1 - R_u^2}. \tag{10}$$

Estimating Regressions R1U and R2U via OLS yields consistent estimates of $R_{u,u'}^2$ *and* R_u^2*.*

2.5 Longer Horizons

Alternatively, one may be interested in understanding the behavior of price informativeness over different horizons. In particular, when investors may have information over changes in future log asset payoffs, the process for the (log) price difference can be modeled as

$$\Delta p_t = \overline{\phi} + \phi_0 \Delta x_t + \sum_{m=1}^{M} \phi_m \Delta x_t^m + \phi_\chi \cdot \Delta \chi_t + \phi_n \Delta n_t, \tag{11}$$

for $M \geq 1$, where

$$\Delta x_t^m \equiv x_{t+m} - x_t = \sum_{l=1}^{m-1} \Delta x_t^l + u_{t+m-1}$$

is the cumulative growth between t and $t+m$ and the process for Δx_{t+l} is given by Equation 1. In this case, the price contains information about payoff changes at different horizons. More specifically, the unbiased signal about the change in the fundamental m periods ahead, Δx_t^m, contained in the price p_t is given by

$$\pi_t^m = \frac{1}{\phi_m + \sum_{l=1,l\neq m}^{M} \phi_l K_m^l} \left(\Delta p_t - \overline{\phi} + \phi_0 \Delta x_t - \phi_n \mu_{\Delta n} - \sum_{l=1,l\neq m}^{M} \phi_l \left(\Delta x_t^l - K_m^l \Delta x_t^m \right) \right)$$

where $K_m^l = \mathbb{C}ov\left[\Delta x_t^m, \Delta x_t^l\right] \mathbb{V}ar\left[\Delta x_t^m\right]^{-1} = \mathbb{V}ar\left[\Delta x_t^{\min\{m,l\}}\right] \mathbb{V}ar\left[\Delta x_t^m\right]^{-1}$. Then, absolute price informativeness about the growth in the payoff m periods ahead is given by

$$\tau_\pi^m \equiv \left(\mathbb{V}\text{ar}\left[\pi_t^m \middle| \Delta x_t^m, \Delta x_t, \Delta \chi_t\right]\right)^{-1}$$

and relative price informativeness about Δx_t^m is given by

$$\tau_\pi^{mR} \equiv \frac{\tau_\pi^m}{\tau_\pi^m + \tau_{\Delta x_t^m | \Delta\chi}},$$

where $\tau_{\Delta x_t^m | \Delta\chi} \equiv \mathbb{V}\mathrm{ar}\left[\Delta x_t^m \middle| \Delta x_t, \Delta\chi_t\right]$.

Proposition 3. (Identifying price informativeness at longer horizons) *Let $\overline{\beta}^m$, β_0^m, β_1^m, and β_2^m denote the coefficients of the following regression of log-price differences on realized and future log-payoff differences and public signals:*

$$\Delta p_t = \overline{\beta}^m + \beta_0^m \Delta x_t + \beta_1^m \Delta x_t^m + \beta_2^m \cdot \Delta\chi_t + e_t^m, \tag{R1m}$$

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 $\Delta x_t^m = x_{t+m} - x_t$ respectively denote the date t and the cumulative $t+m$ log-payoff differences, $\Delta\chi_t$ denotes the change in public information, and where $R^2_{\Delta x, \Delta x^m}$ denotes the R-squared of Regression 3. Let $\overline{\zeta}$, ζ_0, and ζ_2 denote the coefficients of the following regression of log-price differences on realized log-payoff differences and change in public information:

$$\Delta p_t = \overline{\zeta} + \zeta_0 \Delta x_t + \zeta_2 \cdot \Delta\chi_t + e_t^\zeta, \tag{R2}$$

where $R^2_{\Delta x}$ denotes the R-squared of Regression R2. Then, relative price informativeness, τ_π^{mR}, defined in Equation (5), can be recovered as

$$\tau_\pi^{mR} = \frac{R^2_{\Delta x, \Delta x^m} - R^2_{\Delta x}}{1 - R^2_{\Delta x}}. \tag{12}$$

An important observation from Proposition 3 is that whether investors have information about payoff growth rates further out in the future does not invalidate of our approach to learn about the one-period ahead payoff growth. This is easily shown by noting that Proposition 1 can be obtained from Proposition 3 by setting $m = 1$.

2.6 Price Informativeness: Remarks

We qualify and interpret our results in the following four remarks.

Remark 2. *Observability of public signals.* If the external observer from whose perspective one is computing price informativeness only learns from the price and does not observe the public signals, absolute and relative price informativeness are respectively given by given by

$$\tau_\pi \equiv \left(\mathbb{V}\mathrm{ar}\left[\pi_t \middle| \Delta x_{t+1}, \Delta x_t\right]\right)^{-1} = \left(\frac{\phi_1 + \phi_n \omega + \phi_\chi \cdot \overline{\omega}}{\phi_n}\right)^2 \tau_{\Delta n} \quad \text{and} \quad \tau_\pi^R = \frac{\tau_\pi}{\tau_\pi + \tau_u}.$$

In this case, price informativeness can be recovered by dropping the public signals $\Delta \chi_t$ from Regressions R1 and R2 and following the procedure in Proposition 1.

Remark 3. *Usefulness of price informativeness measures.* From a practical standpoint, our results are useful to test predictions in models with dispersed information. For instance, our results open the door to provide model-consistent tests of the results in Kacperczyk, Nosal and Sundaresan (2020), who theoretically characterize the relation between institutional ownership and price informativeness, or Dávila and Parlatore (2021), who theoretically characterize the relation between trading costs and price informativeness. Moreover, while the precise relation between price informativeness and welfare remains an open question in general — despite some recent attempts in particular environments, see e.g., Angeletos and Pavan (2007, 2009), Vives (2017) and Pavan, Sundaresan and Vives (2022) — we hope that showing how to identify and estimate price informativeness leads the way to making empirically-based statements about social welfare in the future.

Remark 4. *Informativeness vs. predictability.* Even though price informativeness and price/return predictability may seem closely connected, they are conceptually different notions. Given our assumptions, Proposition 1 shows that running regressions of prices, which are endogenous, on future payoffs, which are exogenous, allows us to recover price informativeness consistently. 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).

Remark 5. *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 4. This observation may open the door to a higher frequency implementation of our results as data become increasingly available.

2.7 Comparison to Existing Literature

There are two significant differences between our approach and the approach in, for instance, Bai, Philippon and Savov (2016): i) we focus on relative price informativeness, in contrast to forecasting price efficiency (FPE) introduced in Bond, Edmans and Goldstein (2012), and ii) we provide formal identification results, which shapes our estimation approach based on time series regressions. The following two remarks elaborate on these differences.

Remark 6. *Alternative measures of informativeness.* The notions of absolute and relative price informativeness defined above can be related to other variables that have been used to make inferences about the informational content of prices. In particular, our measures of price informativeness are linked to i) the posterior variance of the future payoff conditional on the price and the current payoff, given by $\mathcal{V}_P \equiv \mathbb{V}\mathrm{ar}\left[u_t|\,\Delta p_t, \Delta x_t\right]$ and ii) forecasting price efficiency (FPE), given by $\mathcal{V}_{\mathrm{FPE}} \equiv \mathbb{V}\mathrm{ar}\left[\mathbb{E}\left[u_t|\,\Delta p_t, \Delta x_t\right]\right]$, as defined in Bond, Edmans and Goldstein (2012), through the Law of Total Variance, as follows

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

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

$$\mathcal{V}_P = \frac{1}{\tau_\pi + \tau_{u|\Delta\chi_t}} = \frac{1-\tau^R_\pi}{\tau_{u|\Delta\chi_t}} \qquad \text{and} \qquad \mathcal{V}_{\mathrm{FPE}} = \frac{\tau_\pi}{\tau_\pi + \tau_{u|\Delta\chi_t}}\frac{1}{\tau_{u|\Delta\chi_t}} = \frac{\tau^R_\pi}{\tau_{u|\Delta\chi_t}}. \tag{13}$$

Equation (13) illustrates how both $\mathcal{V}_P$ and $\mathcal{V}_{\mathrm{FPE}}$ inherit the units and scale of the volatility about future payoffs, τ_u^{-1}. In contrast, as we establish above, relative price informativeness, τ^R_π, is a bounded, unit-free measure of informativeness. These properties give relative price informativeness a clear economic interpretation that is independent of the environment analyzed. To see this, note that

$$\tau^R_\pi = \frac{\mathbb{V}\mathrm{ar}\left[\mathbb{E}\left[u_t|\,\Delta p_t, \Delta x_t, \Delta\chi_t\right]\right]}{\mathbb{V}\mathrm{ar}\left[u_t|\,\Delta x_t, \Delta\chi_t\right]}.$$

Therefore, relative price informativeness measures the expected reduction in uncertainty about the innovation to the payoff u_t after observing the change in price Δp_t.

If the objective is to derive theoretical predictions, the difference in units of these measures should not be a concern, since there is a one-to-one mapping between all these notions for a given τ_u. However, when looking at the data, changes in $\mathcal{V}_P$ or in $\mathcal{V}_{\mathrm{FPE}}$ can be driven either by changes in the information contained in the price τ^R_π or by the volatility of the payoff τ_u. Therefore, when making comparisons across assets or time periods, it seems desirable to use a notion of informativeness that does not depend on the volatility of payoffs, hence our preference for τ^R_π.

Equation (13) also highlights that linking $\mathcal{V}_P$ or $\mathcal{V}_{\mathrm{FPE}}$ to the precision of the information contained in prices requires making assumptions on distributions of priors, signals, and updating procedures. For instance, to compute a posterior variance ($\mathcal{V}_P$) it is necessary to take a stance on whether updating is Bayesian or not. Hence, the upshot of working with absolute and relative price informativeness is that these notions do not require us to make assumptions on how an external observer updates or on the shape of the underlying distributions.[6]

Remark 7. *Identification results: time-series vs. cross-sectional regressions.* Since the definition of relative price informativeness in Section 2.2 is based on the precision of the price as a signal over payoffs at different states of nature, and given the assumption that all underlying parameters are time-invariant, our identification results in Propositions 1, 2, and 3 are based on time-series regressions. Therefore, our identification results apply to time-series regressions, which we then implement for a specific stock over a specific time period, as described in Section 4. To account for the potential of time-varying parameters, we run rolling time-series regressions in our empirical implementation.

One could conceive deriving identification results based on cross-sectional regressions of changes in asset prices on changes in asset payoffs using data for multiple stocks at a given point in time. For this alternative approach to be valid, it would be necessary to assume that all underlying parameters (including the distributions of payoffs, signals, and noise) are the same for all stocks at a given point in time. In that case, a cross-sectional regression would recover a single measure of price informativeness — identical across all stocks since that is the identification assumption. This is a highly implausible scenario, both conceptually, since there is no prior reason for informativeness to be identical across all stocks, but also empirically, as we we show direct evidence that the distributions of payoffs across stocks differ.

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 study a representative agent model similar to those used in the macro-finance literature. Finally, we study a model with informed and uninformed investors, as in the classic literature on information and learning.[7] To simplify the analysis and to keep it closer to the existing theoretical literature, the price is the only public signal observed by the investors in this section.

[6] Formally, while finding the precision of the unbiased signal about the innovation to the asset payoff contained in the asset price, $\mathbb{V}\mathrm{ar}\left[\pi_t \middle| u_t, \Delta x_t, \Delta \chi_t\right]$, does not require making distributional assumptions beyond the existence of second moments, finding a posterior variance like $\mathbb{V}\mathrm{ar}\left[u_t \middle| \Delta p_t, \Delta x_t, \Delta \chi_t\right]$ does require such assumptions.

[7] In the Internet Appendix, we present conditions on investors' asset demands that are sufficient to generate an asset pricing equation of the form assumed in Equation (2).

The results in this section have a dual purpose. First, these applications show that our identification results apply to economies i) with or without dispersed information among investors, ii) with time-varying risk aversion/risk-premia, iii) in which investors may or may not learn from prices, and iv) in which noise may arise from different sources. These applications are particularly useful to highlight that our approach does not take a stance on the source of the aggregate noise. Second, these applications allow us to provide a structural economic interpretation of the empirical results presented in Section 4.

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. Each investor i is born with wealth w_0^i and has well-behaved expected utility preferences over his terminal wealth w_1^i, with flow utility given by $U_i\left(w_1^i\right)$, where $U_i'(\cdot) > 0$ and $U_i''(\cdot) < 0$. We assume that the distribution of initial wealth is bounded and i.i.d. across time and investor types.

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

$$\Delta x_{t+1} = \mu_{\Delta x} + u_t, \tag{14}$$

where $\Delta x_{t+1} = x_{t+1} - x_t$, $\mu_{\Delta x}$ 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 revealed to investors at date $t+1$. Note that Equation (14) is a special case of Equation (1) when $\rho = 0$. We focus on the $\rho = 0$ case 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 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 Dávila and Parlatore (2021) for a thorough analysis of this formulation, which extends the classic DSSW model (De Long et al., 1990) to incorporate 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

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

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

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

where $n_0 = 0$, $\mu_{\Delta n}$ is a scalar, and where $\varepsilon_t^{\Delta n} \perp \varepsilon_{\overline{n}t}^i$ 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. Without loss of generality, we assume that $u_{t+s} \sim_{i,t} N\left(0, \tau_u^{-1}\right)$ for all $s > 0$.[8]

Each investor i born at date t optimally chooses a portfolio share in the risky asset, denoted by θ_t^i, to solve

$$\max_{\theta_t^i} \mathbb{E}_t^i\left[U_i\left(w_1^i\right)\right] \tag{15}$$

subject to a wealth accumulation constraint

$$w_1^i = \left(R^f + \theta_t^i\left(\frac{X_{t+1} + P_{t+1}}{P_t} - R^f\right)\right) w_0^i, \tag{16}$$

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, $\int \theta_t^i w_0^i di = Q$.[9]

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).

[8] To simplify the analysis, we assume that investors do not learn from their priors and that the signals and priors are identically distributed across investors. Our results can be easily extended to allow for heterogeneity in τ_s, τ_u, and $\tau_{\overline{n}}$.

[9] It is well known that dynamic rational expectation models may feature multiple equilibria. Our approach is valid for any given equilibrium that may arise.

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 at date t can be approximated as

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

where $\gamma^i \equiv -\frac{w_0^i U_i''\left(w_0^i\right)}{U_i'\left(w_0^i\right)}$, $r^f = \ln\left(R^f\right)$, and k_0 and k_1 are scalars defined in the Appendix.

As we show in the Appendix, taking a first-order log-linear approximation of the first-order condition, 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 a positive or negative scalar that incorporates the risk premium. These coefficients are time invariant since we have assumed that the distribution of investor types is time invariant and the wealth distribution across time and investor type is i.i.d. Using the market clearing condition with this approximation and the information structure described above yields a log-linear approximated price given by

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

where $\overline{\alpha_h} \equiv \int \alpha_h^i w_0^i di$ denotes the wealth-weighted cross-sectional average of a given coefficient α_h^i and $\overline{\psi} = \int \psi^i w_0^i di - Q$. 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 $U'\left(w_1^i\right)$ around the initial wealth level w_0^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 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 recovered 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 will be (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, all 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 asset pricing 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 with sentiment 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, $w_0^i = w_0$, and utility, $\gamma^i = \gamma$.

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

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

where the coefficients α_x, α_s, α_n, and α_p are demand sensitivities and ψ 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 learns from 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' prior $\overline{n}_t$. It is easy to map the representative agent model into the framework developed 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 $\overline{\phi} = 0$, $\phi_0 = \frac{\alpha_x}{\alpha_p} - \frac{\alpha_s}{\alpha_p}$, $\phi_1 = \frac{\alpha_s}{\alpha_p}$, *and* $\phi_n = \frac{\alpha_n}{\alpha_p}$ *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 from 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.

Finally, it is worth highlighting that it is easy to introduce time-varying risk aversion in this framework — this would imply assuming that γ and consequently ψ vary over time, as γ_t and ψ_t. In that case, time-varying risk aversion manifests itself as another source of noise.

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. The general framework in Section 2 applies to settings with noise traders. This application highlights that our identification results accommodate 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 being 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 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 — these 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

$$p_t \approx \frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\alpha_s}}{\overline{\alpha_p}} u_t + \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \overline{n}_t + \frac{\overline{\psi}}{\overline{\alpha_p}} + \frac{\delta_t}{\overline{\alpha_p}}, \tag{17}$$

where $\overline{\alpha_h} \equiv \int_{I \cup U} \alpha_h^i w_0^i di$ denotes the wealth-weighted cross-sectional average of α_h^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 $\overline{\psi} \equiv \int_{I \cup U} \psi^i w_0^i di - Q$.

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 $\Delta \tilde{n}_t \equiv \Delta \overline{n}_t + \frac{1}{\overline{\alpha_n}} \Delta \delta_t$.

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 $\Delta\tilde{n}_t \equiv \Delta\overline{n}_t + \frac{1}{\alpha_n}\Delta\delta_t$ allows us to clearly map this model into the general framework developed 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.

4 Empirical Implementation: Stock-Specific Price Informativeness

In this section, we use our identification results to construct and analyze measures of stock-specific relative price informativeness. We exclusively report estimates of relative price informativeness since — as shown in Section 2 — these allow for 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 at the stock level. In the next three subsections, we present our estimates based on Propositions 1, 2, and 3, respectively.

4.1 Baseline Environment

This subsection implements the results in Proposition 1.

Data Description and Empirical Specification We initially provide a brief description of the data and the sample selection procedure. The Internet Appendix includes a more detailed description. 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 and Murray (2016). From 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. From Institutional Brokers' Estimate System (IBES), we obtain measures of analyst coverage.

Our analysis uses quarterly data, available from 1966 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 — using year-on-year changes with overlapping data over rolling windows of 40 quarters. By working with the model in log-differences, we sidestep concerns associated with failures of stationary — see e.g., Campbell (2017). Considering year-on-year changes deals with seasonality concerns. The use of rolling regressions makes the underlying assumption of parameter stability over a given estimation window more plausible.

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 for different reasons. As discussed in Section 2, our model can be flexibly interpreted to use earnings as the payoff measure. Since earnings can be negative, we compute Δx_t^j as the log year-on-year growth rate scaled by book equity (difference operators Δ are year-on-year) . Formally, in a given rolling window, we run time-series regressions of the form

$$\Delta p_t^j = \overline{\beta}^j + \beta_0^j \Delta x_t^j + \beta_1^j \Delta x_{t+4}^j + \beta_c^j \cdot \Delta w_t^{j,q} + \varepsilon_t^j \quad \Rightarrow R^{2,j}_{\Delta x, \Delta x'} \tag{18}$$

$$\Delta p_t^j = \overline{\zeta}^j + \zeta_0^j \Delta x_t^j + \zeta_c^j \cdot \Delta w_t^{j,q} + \hat{\varepsilon}_t^j \quad \Rightarrow R^{2,j}_{\Delta x} , \tag{19}$$

where Δp_t^j is a measure of capital gains, Δx_t^j and its one-year ahead counterpart Δx_{t+4}^j are measures of earnings growth, and $w_t^{j,q}$ denotes a given set of controls/public signals. We use the following firm-specific variables as public signals for all stocks: i) profitability, ii) dividend ratio, iii) asset growth, and iv) market beta, following Koijen and Yogo (2019). Profitability corresponds to the total operating profits divided by book equity and the dividend ratio to the sum of total dividends over the past year divided by book equity. We estimate the regression coefficients and errors using OLS. We respectively denote the R-squareds of the regressions (18) and (19) by $R^{2,j}_{\Delta x, \Delta x'}$ and the $R^{2,j}_{\Delta x}$. Hence, Regression R1 maps to Equation (18), while Regression R2 maps to Equation (19).

Consistent with Proposition 1, we recover relative price informativeness for stock j in a given period/window from Equations (18) and (19) as follows:

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

After restricting our results to stocks with contiguous observations and whose maximum leverage score across observations is lower than 0.95, we end up with a panel of price informativeness measures for 4063 unique stocks.[10] We have explored alternative criteria to deal with outliers or abnormal observations — for instance, restricting the set of observations

[10]We compute leverage scores as the i^{th} diagonal element of the projection matrix of the observations. Leverage scores describe the influence that each value of the dependent variable has on the fitted value for that same observation.

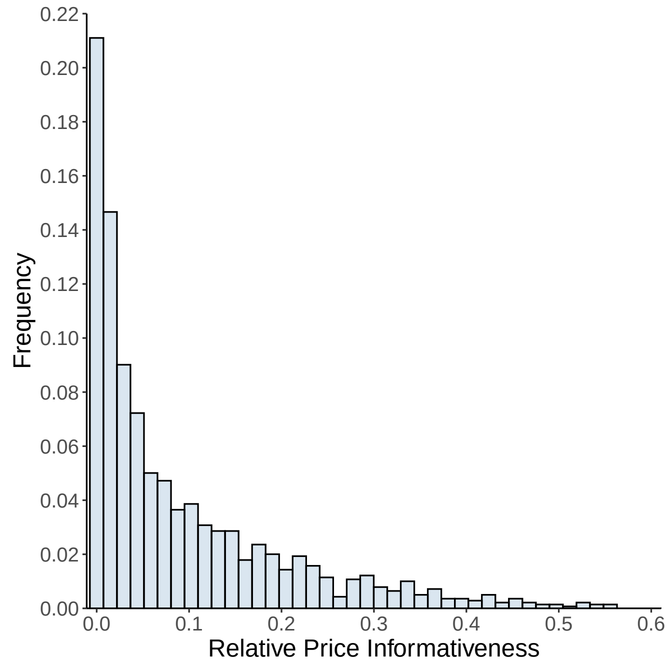


Figure 2: Price informativeness: relative-frequency histogram

Note: Figure 2 shows a relative-frequency histogram of price informativeness for a representative time period, the last quarter of 2015. The histogram features 1,398 stocks.

to those with $\beta_1^j \in [0, 1]$ — but this does not change our conclusions.[11]

Table 1 reports year-by-year summary statistics (every five years) of the distribution of stock-specific price informativeness, starting in 1985. Throughout the paper, informativeness in year t is computed over a rolling window of 40 quarters prior. We illustrate our results graphically in Figure 2, which presents a relative-frequency histogram of price informativeness for a specific time period (last quarter of 2015). The shape of this histogram is representative of other periods.

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 4.47% and 8.54%. Because the distribution of informativeness is skewed, the median is often perceived as a better measure of central tendency. The 95% percentile of the distribution stays below 0.34, which means that an external observer who only learns from the price would rarely put more than a one-third weight on the price when updating his beliefs to form a posterior over future payoffs. Since we have included additional controls in the regressions, our results should be interpreted as the price informativeness for an external observer who observes prices, past payoffs, and the public signals included as controls. We present results excluding the additional controls in the Internet Appendix, which correspond

[11] The fraction of stocks with negative β_1^j coefficients is between 35% and 45% in any given rolling window. Although there are models that can rationalize this, negative β_1^j coefficients can also be a sign of model misspecification.

to estimates of informativeness from the perspective of an external observer who only observes prices and past payoffs.

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 six 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 between trading volume and 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, v) institutional ownership, measured as the proportion of shares held by institutional investors, and vi) analyst coverage, measured as the number of analysts covering the stock in IBES data. Since the last two variables are heavily correlated with size, we also provide results on institutional ownership and analyst coverage orthogonalized to size.

In Table 2, we report the estimates of panel regressions of relative price informativeness (in twentiles) on each of the six 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, analyst coverage). Figures IA-2 through IA-6 in the Internet Appendix provide an alternative graphical illustration of our results. These figures show that the cross-sectional relations identified in Table 2 are stable over time. The last two rows of Table 2 report the estimates of panel regressions of the residual relative price informativeness after controlling for size on institutional ownership and analyst coverage.

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 cross-sectional relation between a stock's book-to-market ratio and price informativeness; that is, value stocks have lower price informativeness. 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 cross-sectional relation between a stock's idiosyncratic return volatility and price informativeness; that is, stocks whose returns are more volatile have higher price informativeness.[12] We also find a strong positive cross-sectional

[12] Dávila and Parlatore (2023) provides a systematic analysis of the relation between volatility and price informativeness.

relation between a stock's institutional ownership share and price informativeness, that is, stocks owned mostly by institutional investors have higher price informativeness. Finally, we find that price informativeness is strongly positively correlated with analyst coverage, that is, stocks that are covered by more analysts have higher price informativeness.

Institutional ownership and analyst coverage are highly correlated with the size of the stock. When we look at the correlation between price informativeness and institutional ownership and analyst coverage beyond that implied by size, we find that the positive correlation between price informativeness and institutional ownership and the correlation between analyst coverage and price informativeness decrease substantially. These results are consistent with institutional ownership being driven by passive investors or investors that are focused on the long run and it suggests that while analysts may help to incorporate information into the price, most investors trading large stocks do not rely on analysts as their main source of information. Finally, if one thinks of large and high turnover stocks as being cheaper to trade, in the sense of price impact being lower, it makes sense that informed investors would gravitate towards them when choosing which stocks to learn about and hence, increase the price informativeness. At the same time, highly liquid stocks may have a different investor base that is more focused on the short run, which would also lead us to expect the correlations we observe.

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 year fixed effects. Even though the distributions of informativeness differ across characteristics, 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 as informative as those listed in the AMEX. Second, we study whether price informativeness varies among stocks that belong to the S&P500 and those that do not. Consistent with our findings on size, we find that stocks outside of the S&P have lower price informativeness on average. Finally, we study the behavior of price informativeness across sectors. Median price informativeness is highest in the manufacturing and wholesale/retail sectors and lowest in the service sector.

Price Informativeness over Time: Aggregate Results An advantage of computing stock-specific measures of price informativeness is that it allows us to study how the distribution of stock-specific price informativeness evolves over time.[13] Table 1 can be used to infer the time

[13] To keep the paper focused, we exclusively study the behavior of the panel of stock-specific price informativeness measures. There is scope to apply our approach to aggregate data in order to generate a time-series of aggregate price informativeness. There is also scope to further explore the time series evolution of informativeness after grouping stocks by characteristics, as we discuss in the next section.

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

t	$Median$	$Mean$	SD	$Skew$	$Kurt$	$P5$	$P25$	$P75$	$P95$	n
1985	0.0492	0.0885	0.1105	2.0898	4.7510	0.0010	0.0116	0.1146	0.3110	360
1990	0.0407	0.0779	0.0952	2.0233	4.5793	0.0020	0.0127	0.1063	0.2777	728
1995	0.0304	0.0648	0.0845	2.3804	7.8812	0.0010	0.0092	0.0903	0.2358	1118
2000	0.0501	0.0878	0.1021	1.9424	4.7537	0.0020	0.0140	0.1235	0.2913	1095
2005	0.0470	0.0852	0.1045	2.1060	5.4419	0.0012	0.0112	0.1200	0.2957	1256
2010	0.0474	0.0924	0.1105	1.7832	3.3741	0.0014	0.0125	0.1335	0.3399	1458
2015	0.0484	0.0918	0.1101	1.8070	3.5715	0.0013	0.0133	0.1314	0.3284	1476

Note: Table 1 reports year-by-year summary statistics (every five years) on the panel of price informativeness estimates. It provides information on the median; mean; standard deviation; skewness; excess kurtosis; and 5th, 25th, 75th, and 95th percentiles of each yearly distribution, as well as the number of stocks in each year. This table averages the quarterly measures of price informativeness for a given year before computing the summary statistics.

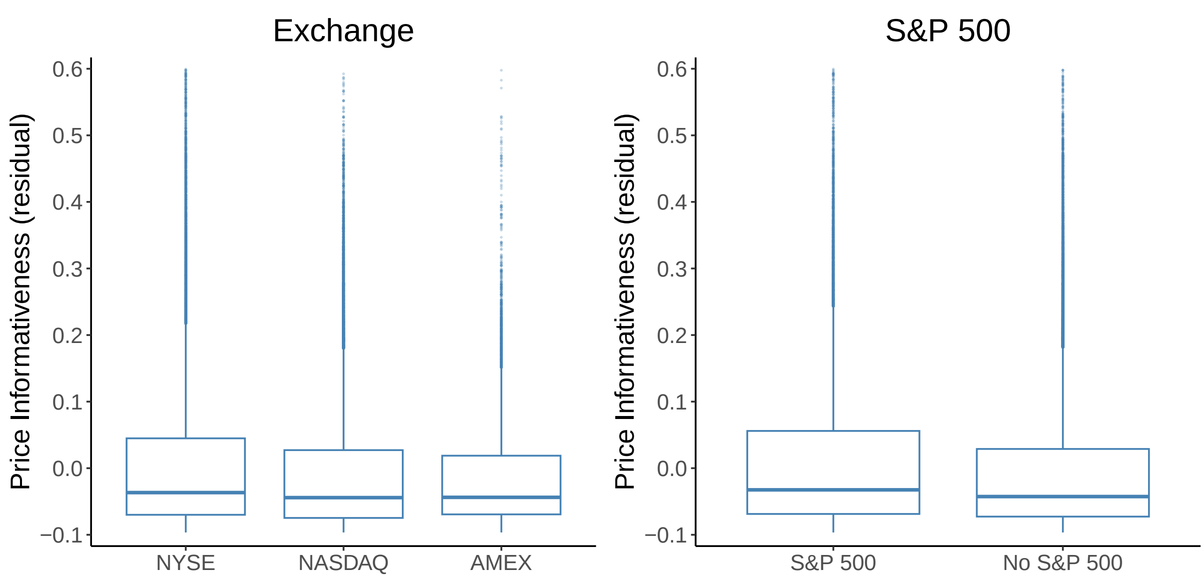


Figure 3: Cross-sectional results (2)

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 75th and 25th percentiles. The whiskers extend up to 1.5 times the interquartile range.

Table 2: Cross-sectional results (1)

	Estimate	Std. Error	t-stat
Size	0.004667	0.000284	16.45
Value	-0.010120	0.000820	-12.34
Turnover	0.001103	0.000042	26.04
Idiosyncratic Volatility	0.038921	0.012887	3.02
Institutional Ownership	0.058937	0.001906	30.92
Analysts Covering	0.002218	0.000121	18.26
Institutional Ownership (Residualized)	0.007705	0.000621	12.41
Analysts (Residualized)	0.000061	0.000030	2.06

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 informativeness 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 IA-2 through IA-6 provide the graphical counterpart of the results in this table. Size is measured as the natural log of stock market capitalization, value is measured as the ratio between a stock's book value and its market capitalization, turnover is measured as the ratio between trading volume and 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, institutional ownership is measured as the proportion of a stock held by institutional investors, and analyst coverage measured as the number of analysts covering the stock in a given quarter. The last two rows provide the estimates of the residual of price informativeness on size on institutional ownership and analyst coverage.

evolution of the distribution of informativeness. To better illustrate the results, we show the behavior of the median, mean, and standard deviation of the cross-sectional distribution of informativeness between 1985 and 2017 graphically in Figure 5.[14]

We find that both the median and the mean of the distribution of informativeness feature increasing trends. The median moves from roughly 4% to 4.5% between 1985 and 2017, while the mean moves from roughly 7.5% in 1986 to roughly 9% by 2017. We also find that the standard deviation of informativeness has a positive long-run trend in our sample. In this case, there are spikes in the cross-sectional standard deviation of informativeness around the early 2000's and the global financial crisis of 2008 — other measures of dispersion have similar behavior.

[14]We start our analysis in 1985 due to small sample sizes in prior years.

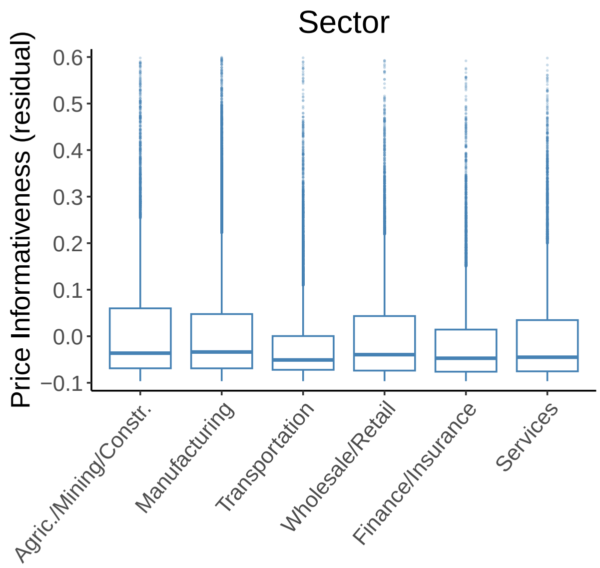


Figure 4: Cross-sectional results (3)

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 75th and 25th percentiles. The whiskers extend up to 1.5 times the interquartile range.

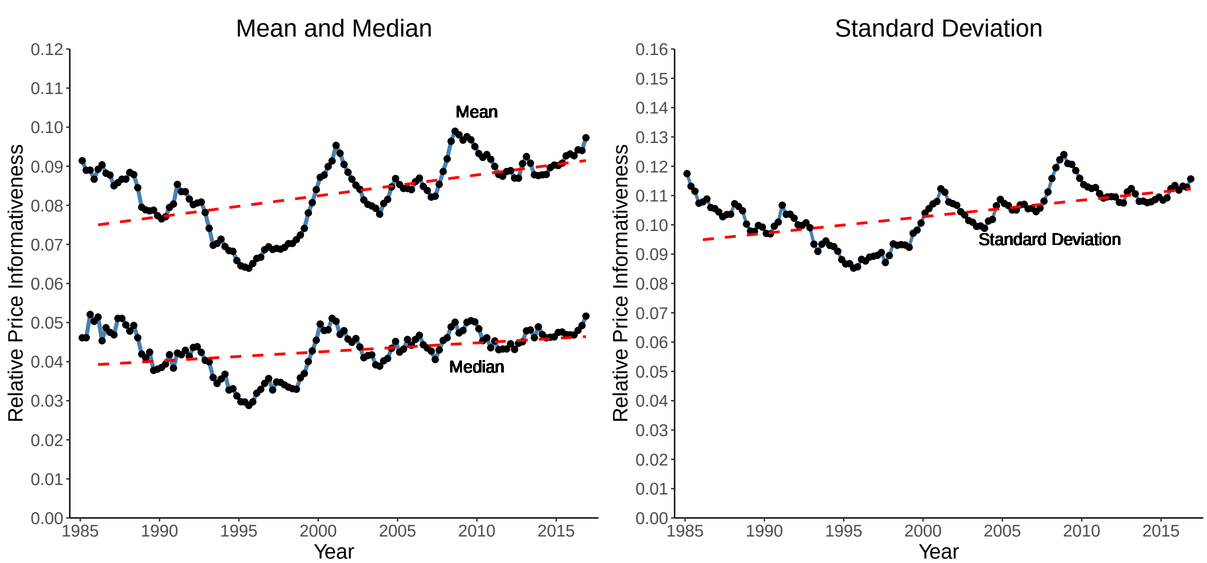


Figure 5: Price informativeness over time: aggregate results

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

Price Informativeness over Time: Cross-Sectional Characteristics We also study the evolution of the distribution of stock-specific price informativeness over time across the six stock characteristics described above: size, turnover, value, idiosyncratic volatility, institutional ownership, and analyst coverage. Figure 6 shows the average and the standard deviation of the top and bottom halves of the distribution of price informativeness along each of the characteristics.

We find that the gap in price informativeness among large and small stocks has increased over time. This fact, together with the higher increase in dispersion in price informativeness for larger stocks, suggests that price informativeness for large and small stocks is diverging. The top and bottom of the distribution of informativeness according to turnover, value, institutional ownership, and analyst coverage have evolved in parallel over the period studied here, while price informativeness for stocks with higher idiosyncratic volatility is lower in recent years.

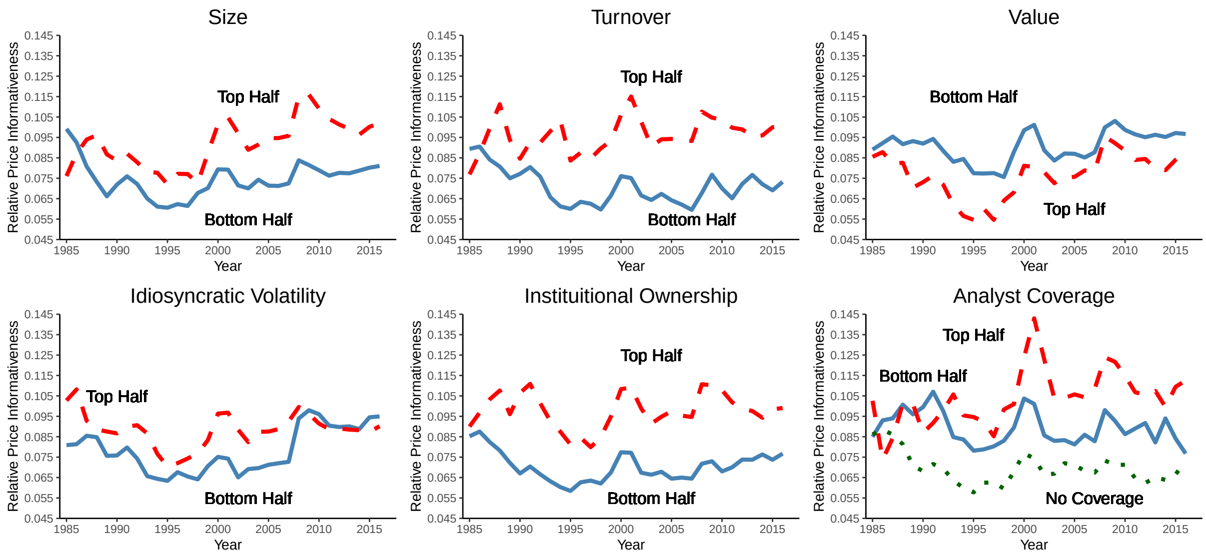
Interpretation of Empirical Findings through Structural Models Finally, note that it is possible to interpret the empirical findings presented in this section through the lens of the structural models developed in Section 3. In particular, if one were merely interested in knowing the precision of the signal contained in asset prices about future payoffs, the empirical results we have just presented directly conclude that such signal is more precise for large, high turnover, high idiosyncratic volatility, and high institutional ownership stocks, and has become more precise on average over the last few decades.

In terms of deeper primitives, our cross-sectional empirical findings are consistent with models in which investors have relatively more precise private information about large, high turnover, and high institutional ownership stocks, while our time-series findings are consistent with an interpretation in which, over the last decades, on average private information has increased relative to the noise in prices. That said, cross-sectional and time-series interpretations of our empirical findings can be consistent with other changes in primitives in alternative models. We provide a more detailed discussion of possible interpretations of these findings in Section F in the Internet Appendix.

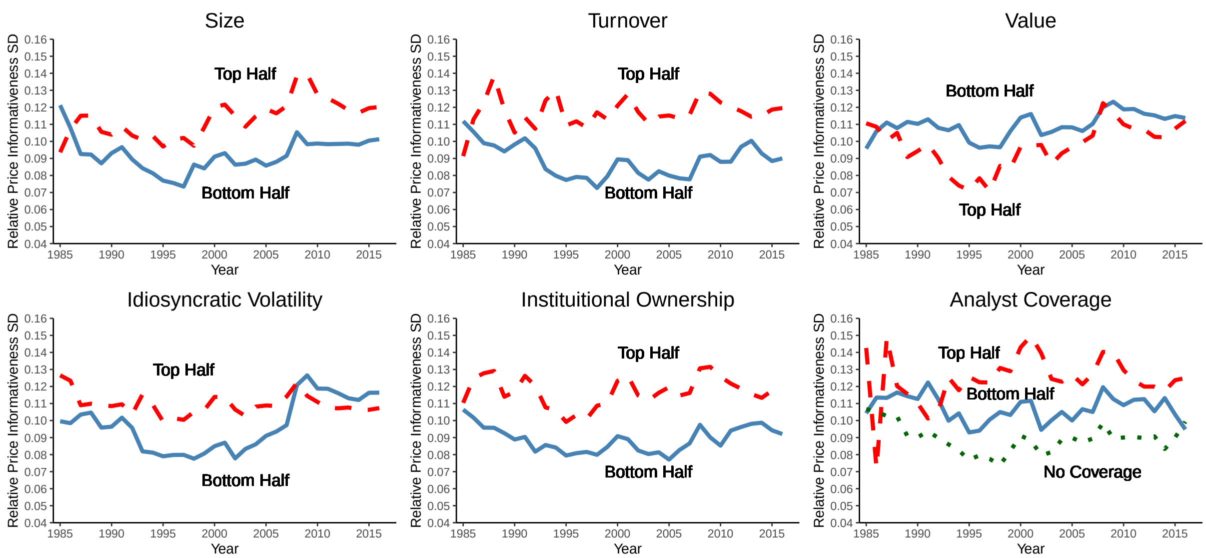
4.2 Unlearnable Payoff Component

In this subsection, we implement the results derived in Proposition 2 to recover stock-specific measures of price informativeness one period ahead when there is an unlearnable component of the future payoff. We parallel the analysis of Section 4, but now implement Regressions R1U and R2U mapping the average analyst forecast to the learnable component of the payoff.

In addition to the data described in the previous section for prices, earnings and controls, we need a measure of the learnable component of earnings to implement the results in Proposition



(a) Mean price informativeness over time across characteristics



(b) Standard deviation of price informativeness over time across characteristics

Figure 6: Price informativeness over time: cross-sectional characteristics

Note: Each of the panels in Figures 6a and 6b shows the time-series evolution of the mean relative price informativeness (Figures 6a) and its standard deviation (Figure 6b) for the top and bottom of the distribution of price informativeness for each characteristic (size, turnover, value, idiosyncratic volatility, institutional ownership, and analyst coverage). The red dashed lines correspond to the top half of the distribution, while the solid blue lines correspond to the bottom half of the distribution. The green dotted line in the last panel corresponds to stocks with no analyst coverage. In this figure, stocks are assigned to each half every quarter and observations are aggregated at the annual level.

2. For this purpose, we use four quarters ahead IBES analyst forecasts of earnings growth rates for the period 1984-2017.[15] We construct our measure of the learnable component of earnings at t by taking the average four quarters ahead analyst growth forecast for period t.

Analogous to our implementation of Proposition 1, we implement Proposition 2 by running time-series regressions for each individual stock over rolling windows of 40 quarters using year-on-year changes and overlapping quarterly data. We denote by f_t^j our measure of the learnable component of earnings computed as the logarithm of the average four quarters ahead analyst growth rate forecast for period t. Formally, in a given rolling window, we run the following time-series regressions:

$$\Delta p_t^j = \overline{\beta}^j + \beta_0^j \Delta x_t^j + \beta_1^{uj} f_t^j + \beta_2^{uj} f_{t+4}^j + \beta_c^j \cdot \Delta w_t^{j,q} + \varepsilon_t^{uj} \quad \Rightarrow R^{2,j}_{f,f'} \tag{20}$$

$$\Delta p_t^j = \overline{\zeta}^{uj} + \zeta_0^{uj} \Delta x_t^j + \zeta_1^{uj} f_t^j + \zeta_c^j \cdot \Delta w_t^{j,q} + \hat{\varepsilon}_t^{uj} \quad \Rightarrow R^{2,j}_{f} , \tag{21}$$

where Δp_t^j is a measure of capital gains, Δx_t^j is a measure of earnings growth, f_t^j is the log of the average analyst forecast, issued at $t-4$, of the change in earnings between $t-4$ and t normalized by book equity at $t-4$ plus one, f_{t+4}^j is its one year ahead counterpart, and $\Delta w_t^{j,q}$ are publicly observed signals. As in the previous section, we use the following firm-specific variables as public signals: i) profitability ii) dividend ratio iii) asset growth, and iv) market beta.

Consistent with Proposition 2, we recover relative price informativeness for stock j in a given window from Equations 20 and 21

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

Figure 7 shows the distribution of price informativeness under the assumption that there is an unlearnable component of earnings for a representative period (last quarter of 2015). The shape of the distribution is similar to the one in Figure 2, with the main difference that there are fewer stocks with price informativeness close to zero. This is reflected in a mean price informativeness in 2015 of 9.29% estimated under unlearnable component model compared to 8.54% estimated under the assumptions of Proposition 1. This pattern is consistent throughout all our rolling window estimates and can be appreciated by comparing Figure 5 with Figure 8. The correlation between our estimates with and without an unlearnable component of earnings is 0.234.

[15] As is usual with this data, we adjust per-share based forecast for splits using the CRSP split factor data. We then construct summary estimates by following guidance from WRDS to most closely match IBES summary data.

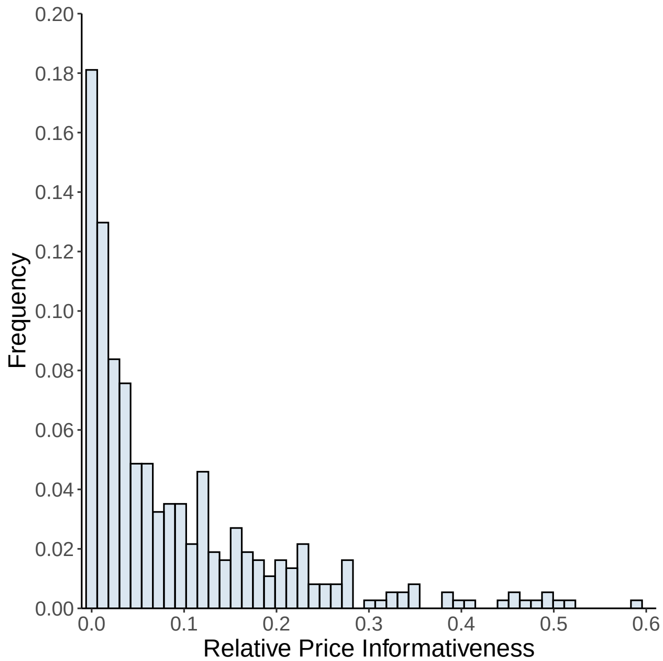


Figure 7: Price informativeness about learnable component: relative-frequency histogram

Note: Figure 7 shows a relative-frequency histogram of price informativeness about the learnable component of future earnings for a representative time period, the last quarter of 2015. The histogram features 370 stocks.

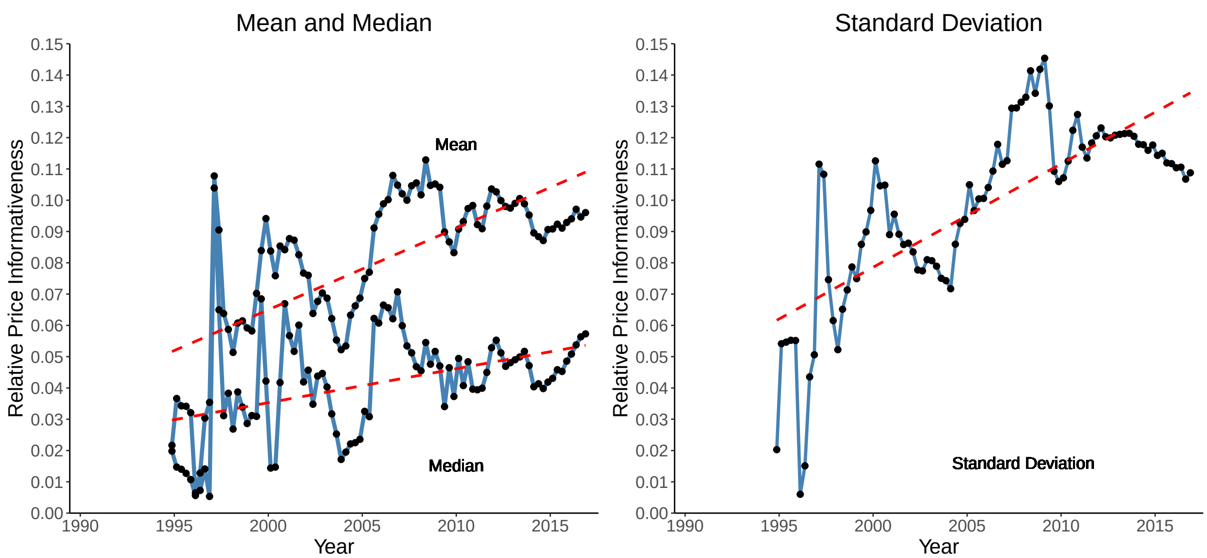


Figure 8: Price informativeness about learnable component over time: aggregate results

Note: The left panel in Figure 8 shows the time-series evolution of the cross-sectional mean and median relative price informativeness about the learnable component of future earnings. The right panel in Figure 8 shows the time-series evolution of the cross-sectional standard deviation of price informativeness. The red dashed lines show linear trends starting in 1994. In both panels, the dots correspond to the average within a quarter of the price informativeness measures computed using year-on-year changes and overlapping quarterly data.

Figure 8 shows the behavior of the median, mean, and standard deviation of the cross-sectional distribution of price informativeness between 1993 and 2017. We find that the mean, median and standard deviation increase during this time period. The median moves from roughly 3% in 1994 to 5% in 2017, while the mean moves from almost 5% to almost 11% between 1994 and 2017. In turn, the standard deviation of informativeness has a positive long-run trend in our sample. These trends, and the cross sectional results presented in the Appendix, are qualitatively consistent with our findings in the previous section.[16]

Our measure of the learnable component of the payoff is by no means perfect. The main assumption behind the choice of proxy for the learnable component of earnings is that analysts actually try to predict future earnings with their forecasts. A potential issue with this measure is the response of analyst forecasts to prices described in Chaudhry (2023). If this response is associated with the information prices contain about future earnings, this dependence is not a concern for our approach. However, any mechanical response of forecasts to prices makes our choice of proxy noisy while still inducing a high R-squared in our regression.

4.3 Longer Horizons

Our empirical analysis so far has focused on price informativeness one year ahead. However, one can be interested in learning about earnings growth at longer horizons. In this subsection, we implement the results in Proposition 3 to estimate price informativeness about earnings at longer horizons using time-series regressions in rolling windows. More specifically, in a given rolling window, we run the following regressions:

$$\Delta p_t^j = \overline{\beta}^j + \beta_0^j \Delta x_t^j + \beta_1^j \Delta x_t^{mj} + \beta_c^j \cdot \Delta w_t^{j,q} + \varepsilon_t^j \quad \Rightarrow R^{2,j}_{\Delta x, \Delta x^m} \tag{22}$$

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

where Δp_t^j is a measure of year-on-year capital gains, Δx_t^j is a measure of year-on-year earnings growth, Δx_t^{mj} measures earnings growths between t and $t+m$, and $w_t^{j,q}$ denotes a given set of controls/public signals. As in our main specification, we use the following aggregate variables as public signals for all stocks: i) profitability, ii) dividend ratio, iii) asset growth, and iv) market beta. We estimate the regression coefficients and errors using OLS. We respectively denote the R-squareds of the regressions (22) and (23) by $R^{2,j}_{\Delta x, \Delta x^m}$ and the $R^{2,j}_{\Delta x}$. Hence, Regression R1m maps to Equation (22), while Regression R2 maps to Equation (23).

Table 3 shows the distribution of the average price informativeness across the different rolling windows at different yearly horizons. We find that price informativeness decreases as

[16] In the Appendix, we provide cross-sectional results for the implementation of Proposition 2, as well as summary statistics year-by-year under this specification.

Table 3: Distribution of average price informativeness at different horizons

Tenor	Mean	Std. Dev	P5	P95
1	0.0841	0.1053	0.000379	0.308
2	0.0638	0.0843	0.000248	0.243
3	0.0536	0.0702	0.000242	0.202
4	0.0553	0.0724	0.000226	0.210
5	0.0542	0.0703	0.000246	0.202

Note: Table 3 shows results for price informativeness about growth rates 1, 2, 3, 4, and 5 years ahead. The table reports the mean, the standard deviation and the 5th and 95th percentiles of average price informativeness across different rolling windows for the different horizons.

one increases the number of periods ahead. This seems natural, since prices are likely to be less informative over payoffs further in the future.

5 Limitations

As in any structural model, our results identifying price informativeness are linked to model assumptions. Moreover, the frequency and type of available data used in the estimation call for caution when drawing conclusions from the estimates of informativeness we present in this paper, both in the cross section and over time. With the goal of aiding future research, we highlight several limitations of our approach in this section.

On the theoretical side, the central assumption behind our results is the linearity of prices (i.e., log price differences). Finding new identification results for notions of informativeness in nonlinear models, like the one in Albagli, Hellwig and Tsyvinski (2015), is a natural progression. Moreover, our analysis purposefully abstracts from feedback between prices and fundamentals, summarized in Bond, Edmans and Goldstein (2012) and tested in Chen, Goldstein and Jiang (2006). Deriving identification results in models with two-way feedback between asset prices and payoffs seems to be another fruitful avenue for future research. There is also scope to connect informativeness measures with social welfare, as discussed in Remark 3, and to better account for the role played by discount rates. Finally, as explained at the end of Section 4.1, drawing conclusive statements about changes in primitives requires additional modeling assumptions (and additional data).

On the measurement side, a key challenge for our approach is that we do not directly observe the learnable component of earnings. This lack of observability leads us to use the average analyst forecast as a proxy in the empirical implementation of our identification results when allowing for an unlearnable component of earnings. While this is a reasonable approach, it is far from ideal, and there is value in exploring alternative empirical designs to address this issue.

Our measure of informativeness is horizon dependent and focuses on informativeness about earnings. Building on our approach, there is scope to develop and estimate measures of informativeness that account for the role of prices as signals over multiple horizons and about different notions of payoffs. It is important to note that, despite conducting multiple robustness checks, looking at different horizons, varying the definition of payoffs/fundamentals, and modifying the set of public signals could lead to finding different conclusions.

Given the rapid pace of information flow in financial markets, using higher frequency data could also provide more precise estimates of informativeness in alternative contexts. Using data available at higher frequencies can also ameliorate concerns associated with parameter instability, which the current draft addresses via rolling windows. The central challenge associated with a higher frequency approach is the need to have valid measures of payoffs/fundamentals.

6 Conclusion

We have shown that the outcomes of regressions of changes in asset prices on changes in asset payoffs can be combined to recover exact measures of price informativeness within a large class of linear/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 fluctuate around levels of 4.5% and 8.5%, respectively. These values, which can be interpreted as the weight that an external observer who learns from the price puts on the price signal when forming a posterior belief about future payoffs, measure the precision of the public signal contained in prices about future payoffs. Cross-sectionally, we find that price informativeness is higher for stocks with higher market capitalization, that trade more frequently, and that have a higher institutional share and higher analyst coverage. Over time, we find that mean and median price informativeness have steadily increased since the mid-1980s. Our framework allows for public signals, noise that is correlated to payoff innovations, an unlearnable payoff component, and makes it to possible to study informativeness at different horizons.

Our identification results open the door to answering a broad set of questions. Empirically, there is scope to explore further 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. Our methodology can also be used to empirically assess the effect of policies on the ability of markets to aggregate information. Theoretically, our results can be used to discipline theories of information and learning in financial markets. There is also scope to export our approach to identification to other environments in which structurally recovering the informativeness of endogenous signals is important, for instance, auctions, macroeconomic environments, or labor markets.

References

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

Angeletos, George-Marios, and Alessandro Pavan. 2007. "Efficient Use of Information and Social Value of Information." *Econometrica*, 75(4): 1103–1142.

Angeletos, George-Marios, and Alessandro Pavan. 2009. "Policy with Dispersed Information." *Journal of the European Economic Association*, 7(1): 11–60.

Bai, Jennie, Thomas Philippon, and Alexi Savov. 2016. "Have Financial Markets Become More Informative?" *Journal of Financial Economics*, 122(3): 625–654.

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*, 24(2): 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.

Chaudhry, Aditya. 2023. "The impact of prices on analyst cash flow expectations." *working paper.*

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.* Princeton University Press.

Dávila, Eduardo, and Cecilia Parlatore. 2021. "Trading Costs and Informational Efficiency." *The Journal of Finance*, 76(3): 1471–1539.

Dávila, Eduardo, and Cecilia Parlatore. 2023. "Volatility and Informativeness." *Journal of Financial Economics*, 147(3): 550–572.

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." *Journal of Finance*, 59(1): 65–105.

Easley, David, Maureen O'Hara, and Joseph Paperman. 1998. "Financial Analysts and Information-Based Trade." *Journal of Financial Markets*, 1: 175–201.

Farboodi, Maryam, Adrien Matray, Laura Veldkamp, and Venky Venkateswaran. 2020. "Where Has All the Data Gone?" *NBER Working Paper.*

Grossman, Sandford J., and Joseph E. Stiglitz. 1980. "On the Impossibility of Informationally Efficient Markets." *American Economic Review*, 393–408.

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

Hellwig, Martin F. 1980. "On the Aggregation of Information in Competitive Markets." *Journal of Economic Theory*, 22(3): 477–498.

Hou, Kewei, and Tobias J. Moskowitz. 2005. "Market Frictions, Price Delay, and the Cross-Section of Expected Returns." *The Review of Financial Studies*, 18: 981–1020.

Hou, Kewei, Lin Peng, and Wei Xiong. 2013. "Is R^2 a Measure of Market Inefficiency?" *Working Paper.*

Kacperczyk, Marcin, Jaromir Nosal, and Savitar Sundaresan. 2020. "Market Power and Price Informativeness." *Working Paper.*

Kacperczyk, Marcin, Savitar Sundaresan, and Tianyu Wang. 2020. "Do Foreign Institutional Investors Improve Price Efficiency?" *Working Paper.*

Koijen, Ralph S. J., and Motohiro Yogo. 2019. "A Demand System Approach to Asset Pricing." *Journal of Political Economy*, 127(4): 1475–1515.

Morck, Randall, Bernard Yeung, and Wayne Yu. 2000. "The Information Content of Stock Markets: Why do Emerging Markets have Synchronous Stock Price Movements?" *Journal of Financial Economics*, 58(1-2): 215–260.

Pavan, Alessandro, Savitar Sundaresan, and Xavier Vives. 2022. "(In) efficiency in Information Acquisition and Aggregation through Prices." *Working Paper.*

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

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.

Vives, Xavier. 2017. "Endogenous public information and welfare in market games." *The Review of Economic Studies*, 84(2): 935–963.

Weller, Brian M. 2018. "Does Algorithmic Trading Reduce Information Acquisition?" *The Review of Financial Studies*, 31(6): 2184–2226.

Wurgler, Jeffrey. 2000. "Financial Markets and the Allocation of Capital." *Journal of Financial Economics*, 58(1-2): 187–214.

Appendix

A Proofs and Derivations: Section 2

Proof of Proposition 1. (Identifying price informativeness)

For completeness, we reproduce here Regressions R1 and R2:

$$\Delta p_t = \overline{\beta} + \beta_0 \Delta x_t + \beta_1 \Delta x_{t+1} + \beta_2 \cdot \Delta \chi_t + e_t \tag{R1}$$
$$\Delta p_t = \overline{\zeta} + \zeta_0 \Delta x_t + \zeta_2 \cdot \Delta \chi_t + e_t^{\zeta}. \tag{R2}$$

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

$$R^2_{\Delta x, \Delta x'} = 1 - \frac{\mathbb{V}\text{ar}\left[e_t\right]}{\mathbb{V}\text{ar}\left[\Delta p_t\right]} \quad \text{and} \quad R^2_{\Delta x} = \frac{\mathbb{V}\text{ar}\left[\zeta_0 \Delta x_t + \zeta_2 \cdot \Delta \chi_t\right]}{\mathbb{V}\text{ar}\left[\Delta p_t\right]}.$$

After substituting Equation (1) in Equation (2), the following relation holds

$$\Delta p_t = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n} + \left(\phi_0 + \rho \phi_1\right) \Delta x_t + \phi_\chi \cdot \Delta \chi_t + \left(\phi_1 + \phi_n \omega\right) u_t + \phi_n \varepsilon_t^{\Delta n}. \tag{24}$$

By comparing Regression R1 with the structural Equation (2), it follows that $\overline{\beta} = \overline{\phi} + \phi_n \mu_{\Delta n}$, $\beta_0 = \phi_0$, $\beta_1 = \phi_0 + \phi_n \omega$, and $e_t = \phi_n \varepsilon_t^{\Delta n}$. By comparing Regression R2 with the structural Equation (24), it follows that $\overline{\zeta} = \overline{\phi} + \phi_1 \mu_{\Delta x} + \phi_n \mu_{\Delta n}$, $\zeta_0 = \phi_0 + \rho \phi_1$, $\zeta_2 = \phi_\chi$ and $e_t^{\zeta} = \beta_1 u_t + \phi_n \varepsilon^{\Delta n_t}$.

On the one hand, note that $\frac{\beta_1^2}{\mathbb{V}\text{ar}[e_t]} = \tau_\pi$. Given the assumptions on u_t and $\varepsilon_t^{\Delta n}$, it is straightforward to show that the OLS estimates of Regressions R1 are consistent, which implies that price informativeness can be consistently estimated as $\frac{\hat{\beta}_1^2}{\widehat{\mathbb{V}\text{ar}}[e_t]}$. Formally, $\operatorname{plim}\left(\widehat{\tau_\pi}\right) = \operatorname{plim}\left(\frac{\hat{\beta}_1^2}{\widehat{\mathbb{V}\text{ar}}[e_t]}\right) = \tau_\pi$.

On the other hand, when estimating R2 via OLS, the error term e_t^{ζ} is correlated with $\Delta \chi_t$ since $\Delta \chi_t = \varpi u_t + \varepsilon_t^{\Delta \chi}$. Then, the estimated coefficient for $\Delta \chi_t$ will be given by

$$\hat{\zeta}_2 = \phi_\chi + \left(\phi_1 + \rho \phi_n\right) \hat{\lambda}_\chi = \zeta_2 + \beta_1 \hat{\lambda}_\chi,$$

where $\hat{\lambda}_\chi = \left(\mathbb{V}ar\left[\Delta \chi_t\right]\right)^{-1} \mathbb{C}ov\left[\Delta \chi_t, u_t\right]$ is the coefficient in the regression R3 below

$$u_t = \overline{\lambda} + \lambda_\chi \cdot \Delta \chi_t + e_t^{\lambda}. \tag{R3}$$

From Regression (R2) we have

$$\mathbb{V}\text{ar}\left[\Delta p_t\right] = \mathbb{V}\text{ar}\left[\hat{\zeta}_0 \Delta x_t + \hat{\zeta}_2 \cdot \Delta \chi_t\right] + \mathbb{V}\text{ar}\left[\hat{e}_t^{\zeta}\right],$$

which can be expressed as

$$1 = \underbrace{\frac{\mathbb{V}\text{ar}\left[\hat{\zeta}_0 \Delta x_t + \hat{\zeta}_2 \cdot \Delta \chi_t\right]}{\mathbb{V}\text{ar}\left[\Delta p_t\right]}}_{=R^2_{\Delta x}} + \frac{\mathbb{V}\text{ar}\left[\hat{e}_t^{\zeta}\right]}{\mathbb{V}\text{ar}\left[e_t\right]} \underbrace{\frac{\mathbb{V}\text{ar}\left[e_t\right]}{\mathbb{V}\text{ar}\left[\Delta p_t\right]}}_{=1-R^2_{\Delta x, \Delta x'}}. \tag{25}$$

The errors in Regression (R2) are given by

$$\begin{aligned}
\hat{e}_t^{\zeta} &= \beta_1 u_t + \phi_n \varepsilon^{\Delta n_t} - \beta_1 \hat{\lambda}_\chi \Delta \chi_t \\
\hat{e}_t^{\zeta} &= \beta_1 \left(1 - \hat{\lambda}_\chi' \omega\right) u_t + \phi_n \varepsilon^{\Delta n_t} - \beta_1 \hat{\lambda}_\chi \cdot \varepsilon_t^{\Delta \chi_t}.
\end{aligned}$$

Then

$$\mathbb{V}\text{ar}\left[\hat{e}_t^{\zeta}\right] = \beta_1^2 \mathbb{V}\text{ar}\left[u_t\right] + \mathbb{V}\text{ar}\left[e_t\right] - \beta_1^2 \mathbb{C}\text{ov}\left[\Delta \chi_t, u_t\right]' \mathbb{V}\text{ar}\left[\Delta \chi_t\right]^{-1} \mathbb{C}\text{ov}\left[\Delta \chi_t, u_t\right].$$

Moreover, using the Law of Total Variance, we have that

$$\begin{aligned}
\mathbb{V}\text{ar}\left[u_t\right] &= \mathbb{E}\left[\mathbb{V}\text{ar}\left[u_t | \Delta \chi_t\right]\right] + \mathbb{V}\text{ar}\left[\mathbb{E}\left[u_t | \chi_t\right]\right] \\
&= \mathbb{E}\left[\mathbb{V}\text{ar}\left[u_t | \Delta \chi_t\right]\right] + \mathbb{V}\text{ar}\left[\hat{\lambda}_\chi' \Delta \chi_t \hat{\lambda}_\chi\right] \\
&= \mathbb{E}\left[\mathbb{V}\text{ar}\left[u_t | \Delta \chi_t\right]\right] + \hat{\lambda}_\chi' \mathbb{V}\text{ar}\left[\Delta \chi_t\right] \hat{\lambda}_\chi \\
&= \mathbb{E}\left[\mathbb{V}\text{ar}\left[u_t | \Delta \chi_t\right]\right] + \mathbb{C}\text{ov}\left[\Delta \chi_t, u_t\right]' \mathbb{V}\text{ar}\left[\Delta \chi_t\right]^{-1} \mathbb{C}\text{ov}\left[\Delta \chi_t, u_t\right].
\end{aligned}$$

Then, the expected posterior variance of u_t given χ_t is $\tau_{u|\Delta\chi}^{-1} \equiv \mathbb{E}\left[\mathbb{V}\text{ar}\left[u_t | \chi_t\right]\right]$, so

$$\mathbb{V}\text{ar}\left[u_t\right] - \tau_{u|\Delta\chi}^{-1} = \mathbb{C}\text{ov}\left[\Delta \chi_t, u_t\right]' \mathbb{V}\text{ar}\left[\Delta \chi_t\right]^{-1} \mathbb{C}\text{ov}\left[\Delta \chi_t, u_t\right].$$

Therefore, we can write

$$\begin{aligned}
\frac{\mathbb{V}\text{ar}\left[\hat{e}_t^{\zeta}\right]}{\mathbb{V}\text{ar}\left[e_t\right]} &= \frac{\beta_1^2 \mathbb{V}\text{ar}\left[u_t\right] + \mathbb{V}\text{ar}\left[e_t\right] - \beta_1^2 \left(\mathbb{V}\text{ar}\left[u_t\right] - \tau_{u|\Delta\chi}^{-1}\right)}{\mathbb{V}\text{ar}\left[e_t\right]} \\
&= \frac{\mathbb{V}\text{ar}\left[e_t\right] + \beta_1^2 \tau_{u|\Delta\chi}^{-1}}{\mathbb{V}\text{ar}\left[e_t\right]} = 1 + \frac{\beta_1^2 \tau_{u|\Delta\chi}^{-1}}{\mathbb{V}\text{ar}\left[e_t\right]} = 1 + \frac{\tau_\pi}{\tau_{u|\Delta\chi}}
\end{aligned}$$

Using this in Equation (25), we get

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

Note that if variables are normally distributed, $\tau^{-1}_{u|\Delta\chi} = \mathbb{V}\text{ar}\left[u_t|\chi_t\right]$ and the expression above is the Kalman gain associated with the unbiased signal contained in the price π_t for an external observer who has access to the public signals χ_t before seeing the price.

Proof of Proposition 2. (Identifying price informativeness)

The proof follows the proof of Proposition 1 replacing the future payoff by the future learnable component of the payoff and including the realized learnable component of the payoff as a public signal.

Proof of Proposition 3 (Identifying price informativeness at longer horizons)

The proof follows the same structure as the proof of Proposition 1. First note that absolute price informativeness about the m-period ahead change in the payoff, Δx^m_t, is given by

$$\begin{aligned}
\tau^m_\pi &\equiv \left(\mathbb{V}\text{ar}\left[\pi^m_t \middle| \Delta x_{t+m}, \Delta x_t, \Delta\chi_t\right]\right)^{-1} \\
&= \left(\mathbb{V}\text{ar}\left[\frac{\phi_n}{\phi_m + \sum_{l=1,l\neq m}^M \phi_l K^l_m}\varepsilon^{\Delta n}_t + \sum_{l=1,l\neq m}^M \frac{\phi_l\left(\Delta x^l_t - K^l_m \Delta x^m_t\right)}{\phi_m + \sum_{l=1,l\neq m}^M \phi_l K^l_m}\right| \Delta x^m_t, \Delta x_t, \Delta\chi_t\Bigg]\Bigg)^{-1} \\
&= \left(\left(\frac{\phi_n}{\phi_m + \sum_{l=1,l\neq m}^M \phi_l K^l_m}\right)^2 \tau^{-1}_{\Delta n} + \mathbb{V}\text{ar}\left[\sum_{l=1,l\neq m}^M \frac{\phi_l\left(\Delta x^l_t - K^l_m \Delta x^m_t\right)}{\phi_m + \sum_{l=1,l\neq m}^M \phi_l K^l_m}\right| \Delta x^m_t, \Delta x_t, \Delta\chi_t\Bigg]\Bigg)^{-1} \\
&= \left(\left(\frac{\phi_n}{\phi_m + \sum_{l=1,l\neq m}^M \phi_l K^l_m}\right)^2 \tau^{-1}_{\Delta n} + \mathbb{V}\text{ar}\left[\sum_{l=1,l\neq m}^M \frac{\phi_l\left(\Delta x^l_t - K^l_m \Delta x^m_t\right)}{\phi_m + \sum_{l=1,l\neq m}^M \phi_l K^l_m}\right| \Delta x_t, \Delta\chi_t\Bigg]\Bigg)^{-1}
\end{aligned}$$

since $\frac{\sum_{l=1,l\neq m}^M \phi_l\left(\Delta x^l_t - K^l_m \Delta x^m_t\right)}{\phi_m + \sum_{l=1,l\neq m}^M \phi_l K^l_m}$ is orthogonal to Δx^m_t by construction. Moreover, not that $\phi_m + \sum_{l=1,l\neq m}^M \phi_l K^l_m$ can be consistently estimated by β^m_2 through OLS.

Moreover, from Regression R2 we have

$$1 = \underbrace{\frac{\mathbb{V}\text{ar}\left[\hat{\zeta}_0 \Delta x_t + \hat{\zeta}_2 \cdot \Delta\chi_t\right]}{\mathbb{V}\text{ar}\left[\Delta p_t\right]}}_{=R^2_{\Delta x}} + \frac{\mathbb{V}\text{ar}\left[\hat{e}^\zeta_t\right]}{\mathbb{V}\text{ar}\left[\hat{e}^\lambda_t\right]} \underbrace{\frac{\mathbb{V}\text{ar}\left[\hat{e}^\lambda_t\right]}{\mathbb{V}\text{ar}\left(\Delta p_t\right)}}_{=1-R^2_{\Delta x,\Delta x^m}}, \tag{26}$$

where $\hat{\zeta}_0$, $\hat{\zeta}_2$, $\hat{e}_t^{\zeta}$, and $\hat{e}_t^{\lambda}$ are OLS estimates with

$$\frac{\mathbb{V}\mathrm{ar}\left[\hat{e}_t^{\zeta}\right]}{\mathbb{V}\mathrm{ar}\left[\hat{e}_t^{\lambda}\right]} = \frac{\mathbb{V}\mathrm{ar}\left[\hat{e}_t^{\zeta}\right]}{\left(\beta_2^m\right)^2 \mathbb{V}\mathrm{ar}\left[\pi_t^m \middle| \Delta x_t^m, \Delta x_t, \Delta \chi_t\right]} = \frac{\tau_\pi^m}{\lambda_m^2} \mathbb{V}\mathrm{ar}\left[\hat{e}_t^{\zeta}\right]$$

and

$$\mathbb{V}\mathrm{ar}\left[\hat{e}_t^{\zeta}\right] = \phi_n^2 \tau_{\Delta n}^{-1} + \mathbb{V}\mathrm{ar}\left[\sum_{l=1}^{M} \phi_l \Delta x_t^l\right] - \mathbb{C}\mathrm{ov}\left[X_t, \sum_{l=1}^{M} \phi_l \Delta x_t^l\right]' \mathbb{V}\mathrm{ar}\left(X_t\right)^{-1} \mathbb{C}\mathrm{ov}\left[X_t, \sum_{l=1}^{M} \phi_l \Delta x_t^l\right]$$

where $X_t \equiv [\Delta x_t, \Delta \chi_t]$. Note that the estimated error takes into account that X_t is correlated with e_t^{ζ}. Moreover, using the Law of Total Variance we have

$$\mathbb{V}\mathrm{ar}\left[\sum_{l=1}^{M} \phi_l \Delta x_t^l\right] - \mathbb{C}\mathrm{ov}\left[X_t, \sum_{l=1}^{M} \phi_l \Delta x_t^l\right]' \mathbb{V}\mathrm{ar}\left(X_t\right)^{-1} \mathbb{C}\mathrm{ov}\left[X_t, \sum_{l=1}^{M} \phi_l \Delta x_t^l\right] = \mathbb{E}\left[\mathbb{V}\mathrm{ar}\left[\sum_{l=1}^{M} \phi_l \Delta x_t^l \middle| \Delta x_t, \Delta \chi_t\right]\right],$$

where we can rewrite

$$\begin{aligned}
\mathbb{V}\mathrm{ar}\left[\sum_{l=1}^{M} \phi_l \Delta x_t^l \middle| \Delta x_t, \Delta \chi_t\right] &= \mathbb{V}\mathrm{ar}\left[\left(\phi_m + \sum_{l=1, l\neq m}^{M} K_m^l \phi_l\right) \Delta x_t^m + \sum_{l=1, l\neq m}^{M} \phi_l \left(\Delta x_t^l - K_m^l \Delta x_t^m\right) \middle| \Delta x_t, \Delta \chi_t\right] \\
&= \left(\phi_m + \sum_{l=1, l\neq m}^{M} K_m^l \phi_l\right)^2 \mathbb{V}\mathrm{ar}\left[\Delta x_t^m \middle| \Delta x_t, \Delta \chi_t\right] \\
&+ \mathbb{V}\mathrm{ar}\left[\sum_{l=1, l\neq m}^{M} \phi_l \left(\Delta x_t^l - K_m^l \Delta x_t^m\right) \middle| \Delta x_t, \Delta \chi_t\right] \\
&= \left(\beta_2^m\right)^2 \mathbb{V}\mathrm{ar}\left[\Delta x_t^m \middle| \Delta x_t, \Delta \chi_t\right] + \mathbb{V}\mathrm{ar}\left[\sum_{l=1, l\neq m}^{M} \phi_l \left(\Delta x_t^l - K_m^l \Delta x_t^m\right) \middle| \Delta x_t, \Delta \chi_t\right]
\end{aligned}$$

Then, using the definition of τ_π^m and $\tau_{\Delta x_t^m|\Delta\chi}$ we have

$$\begin{aligned}
\frac{\mathbb{V}\mathrm{ar}\left[\hat{e}_t^{\zeta}\right]}{\mathbb{V}\mathrm{ar}\left[\hat{e}_t^{\lambda}\right]} &= \frac{\tau_\pi^m}{\left(\beta_2^m\right)^2}\left(\phi_n^2 \tau_{\Delta n}^{-1} + \mathbb{V}\mathrm{ar}\left[\sum_{l=1, l\neq m}^{M} \phi_l \left(\Delta x_t^l - K_m^l \Delta x_t^m\right) \middle| \Delta x_t, \Delta \chi_t\right] + \left(\beta_2^m\right)^2 \mathbb{V}\mathrm{ar}\left[\Delta x_t^m \middle| \Delta x_t, \Delta \chi_t\right]\right) \\
&= \frac{\tau_\pi^m + \tau_{\Delta x_t^m|\Delta\chi}}{\tau_{\Delta x_t^m|\Delta\chi}}.
\end{aligned}$$

Plugging this in Equation (26) we have

$$\frac{R^2_{\Delta x, \Delta x^m} - R^2_{\Delta x}}{1 - R^2_{\Delta x}} = \frac{\tau_\pi^m}{\tau_\pi^m + \tau_{m|\chi}} \equiv \tau_\pi^{mR},$$

which proves our result.

Internet Appendix

For Online Publication Only

Section B of this Internet Appendix describes in more detail the data used for the empirical implementation of the results in Section 4. Section C of this Internet Appendix reports 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 additional cross-sectional results under the assumption that there is an unlearnable component of earnings. Finally, we present the cross-sectional correlation between price informativeness measured in year t and price informativeness measured in prior years. Section E of this Internet Appendix includes additional results on the identification of absolute price informativeness. Section F provides possible structural interpretations of the empirical findings. Section G of this Internet Appendix considers three alternative 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.

A Proofs and Derivations: Section 3

Portfolio Demand Approximation The optimality condition of an investor who maximizes Equation (15) subject to the wealth accumulation constraint in Equation (16) is given by

$$\mathbb{E}\left[U_i'\left(w_1^i\right)\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\middle|\mathcal{I}_t^i\right]=0. \tag{27}$$

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 $U'\left(w_1^i\right)$ around the current date t wealth level w_0^i. Formally, we approximate $U'\left(w_1^i\right)$ as follows

$$U'\left(w_1^i\right)\approx U'\left(w_0^i\right)+U''\left(w_0^i\right)\Delta w_1^i,$$

which allows us to express Equation (27) as

$$U'\left(w_0^i\right)\mathbb{E}_i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]+U''\left(w_0^i\right)w_0^i\mathbb{E}_i\left[\left(R^f-1+\theta_t^i\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\right)\left(\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right)\right]\approx 0.$$

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 (27) as

$$U'\left(w_0^i\right)\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}-R^f\right]+U''\left(w_0^i\right)w_0^i\theta_t^i\mathbb{V}\mathrm{ar}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}\right]\approx 0.$$

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

$$\theta_t^i \approx \frac{1}{\gamma^i} \frac{\mathbb{E}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t} - R^f\right]}{\mathbb{V}\mathrm{ar}_t^i\left[\frac{X_{t+1}+P_{t+1}}{P_t}\right]}, \tag{28}$$

where $\gamma^i \equiv -\frac{w_0^i U_i''(w_0^i)}{U_i'(w_0^i)}$ denotes the coefficient of relative risk aversion. These coefficients are time invariant since we have assumed that the distribution of investor types is time invariant and the wealth distribution across time and investor type is i.i.d.

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

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

where $y_t = \ln Y_t$ for any given variable Y_t. 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 (28), we can express an investor's risky portfolio share θ_t^i as

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

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 $\mathbb{V}\mathrm{ar}\left[k_1\left(p_{t+1}-x_{t+1}\right)+\Delta x_{t+1}\right]$ is a constant. Under this conjecture, θ_t^i is a linear function of s_t^i, x_t, and $\overline{n}_t^i$, given by

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

These coefficients are time invariant since we have assumed that the distribution of investor types is time invariant and the wealth distribution across time and investor type is i.i.d.

This expression and the market clearing condition $\int \theta_t^i w_0^i di = Q$ imply that

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

where $\overline{\alpha_h} \equiv \int \alpha_h^i w_0^i di$ for $h = \{x, s, n, p\}$ and $\overline{\psi} \equiv \int \psi^i w_0^i di - Q$. As in Vives (2008), we make use of the Strong Law of Large Numbers, since the sequence of independent random variables $\left\{\alpha_s^i w_0^i \varepsilon_{st}^i, \alpha_n^i w_0^i \varepsilon_{\overline{n}t}^i\right\}$

has uniformly bounded variance and mean zero. 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 $\tau_{\hat{\pi}} \equiv \mathbb{V}\mathrm{ar}\left[\hat{\pi}_t | u_t, \{x_s\}_{s \le t}, p_{t-1}\right]^{-1} = \left(\frac{\overline{\alpha_s}}{\overline{\alpha_n}}\right)^2 \tau_{\Delta n}$. Note that we denote by π_t the unbiased signal of u_t contained in the change in log prices Δp_t and by $\hat{\pi}_t$ the unbiased signal about u_t contained in the log price p_t.

Given the information set of the investor, $\mathbb{V}\mathrm{ar}\left[n_t | u_t, \{x_s\}_{s \le t}, p_{t-1}\right] = \mathbb{V}\mathrm{ar}\left[\Delta n_t | u_t, \{x_s\}_{s \le t}, p_{t-1}\right]$. 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 $\mathbb{V}\mathrm{ar}\left[u_t | \mathcal{I}_t^i\right] = (\tau_s + \tau_u + \tau_{\hat{\pi}})^{-1}$, where $\mathcal{I}_t^i = \left\{s_t^i, \overline{n}_t^i, \{p_s\}_{s \le t}, \{x_s\}_{s \le t}\right\}$. Note that these two expressions imply that our conjecture about θ_t^i is satisfied. To see this, note that

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

where we used that $\mathbb{E}_t^i[u_{t+1}] = 0$, that $\mathbb{E}_t^i\left[\varepsilon_{t+1}^{\Delta n}\right] = 0$, and that $\mathbb{E}_t^i[n_t]$ is linear in p_t and x_t. To see this, first note that n_{t-1} is known at time t since the information set of the investor includes all past prices and payoffs. Therefore, the prior mean of investor i about n_t is $\mu_{\Delta n} + n_{t-1}$. Second, the price p_t contains information about n_t. The unbiased signal about n_t contained in the price p_t is given by

$$\pi_t^n \equiv \frac{\overline{\alpha_p}}{\overline{\alpha_n}} \left(p_t - \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\psi}}{\overline{\alpha_p}}\right)\right) = n_t + \frac{\overline{\alpha_s}}{\overline{\alpha_n}} u_t,$$

and its precision is given by $\tau_{\pi^n} \equiv \left(\frac{\overline{\alpha_n}}{\overline{\alpha_s}}\right)^2 \tau_u$. Then,

$$\mathbb{E}_t^i[n_t] = \frac{\tau_{\Delta n}(\mu_{\Delta n} + n_{t-1}) + \tau_{\pi^n} \frac{\overline{\alpha_p}}{\overline{\alpha_n}} \left(p_t - \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} x_t + \frac{\overline{\psi}}{\overline{\alpha_p}}\right)\right)}{\tau_{\Delta n} + \tau_{\pi^n}}$$

and $\mathbb{V}\mathrm{ar}_t^i[n_t] = (\tau_{\Delta n} + \tau_{\pi^n})^{-1}$. Moreover,

$$\mathbb{V}\mathrm{ar}_t^i[k_1(p_{t+1} - x_{t+1}) + \Delta x_{t+1}] = k_1^2 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1}\right)^2 (\tau_s + \tau_u + \tau_{\hat{\pi}})^{-1} + k_1^2 \left(\frac{\overline{\alpha_s}}{\overline{\alpha_p}}\right)^2 \tau_u^{-1} + k_1^2 \left(\frac{\overline{\alpha_n}}{\overline{\alpha_p}}\right)^2 \tau_{\Delta n}^{-1}.$$

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

$$\alpha_x^i = \frac{1}{\kappa_i} k_1 \left(\left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \left(1 - \frac{\tau_{\hat{\pi}} \frac{\overline{\alpha_x}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \right) - \frac{\tau_{\pi^n} \frac{\overline{\alpha_x}}{\overline{\alpha_p}}}{\tau_{\Delta n} + \tau_{\pi^n}} \right) \tag{30}$$

$$\alpha_s^i = \frac{1}{\kappa_i} k_1 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \frac{\tau_s}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \tag{31}$$

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

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

$$\psi^i = \frac{1}{\kappa_i} \left(\begin{array}{c} k_0 + k_1 \left(- \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \left(\frac{\tau_\pi \frac{\overline{\alpha_p}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} - \mu_{\Delta x} \right) + 1 \right) \left(\frac{\overline{\alpha_n}}{\overline{\alpha_p}} \mu_{\Delta n} + \frac{\overline{\psi}}{\overline{\alpha_p}} \right) \\ + k_1 \frac{\overline{\alpha_n}}{\overline{\alpha_p}} \frac{\tau_{\Delta n} (\mu_{\Delta n} + n_{t-1}) + \tau_{\pi^n} \frac{\overline{\alpha_p}}{\overline{\alpha_n}} \frac{\overline{\psi}}{\overline{\alpha_p}}}{\tau_{\Delta n} + \tau_{\pi^n}} - r^f \end{array} \right), \tag{34}$$

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

Proof of Lemma 1 Iterating forward Equation (29) and taking differences, we find that

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

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 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 (30) through (34). 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.

Time-varying risk-aversion interpretation Note that from Equations (30) through (34) one can see that $\frac{\alpha_{ht}}{\alpha_{pt}} = \frac{\alpha_{ht-1}}{\alpha_{pt-1}} = \frac{\alpha_h}{\alpha_p}$ for all t and $h = x, s, n$, and that

$$\frac{\psi_t - Q}{\alpha_{pt}} - \frac{\psi_{t-1} - Q}{\alpha_{pt-1}} = \Delta \gamma_t \frac{\mathbb{V}\mathrm{ar} \left[k_1 \left(p_{t+1} - x_{t+1} \right) + \Delta x_{t+1} \right]}{\left(1 - k_1 \left(\frac{\overline{\alpha_x}}{\overline{\alpha_p}} - 1 + \frac{1}{k_1} \right) \frac{\tau_{\hat{\pi}} \frac{\overline{\alpha_p}}{\overline{\alpha_s}}}{\tau_s + \tau_u + \tau_{\hat{\pi}}} \right)} Q = \frac{\Delta \gamma_t}{\gamma_t \alpha_{pt}} Q.$$

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 $\overline{\phi} = 0$, $\phi_0 = \frac{\alpha_x}{\alpha_p} - \frac{\alpha_s}{\alpha_p}$, $\phi_1 = \frac{\alpha_s}{\alpha_p}$, and $\phi_n = \frac{\alpha_n}{\alpha_p}$, 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. In that case, the demand for informed and uninformed investors is respectively given by

$$\begin{aligned}\theta_t^I &\approx \alpha_x^I x_t + \alpha_s^I u_t - \alpha_p^I p_t + \psi^I\\ \theta_t^U &\approx \alpha_x^U x_t + \alpha_n^U \overline{n}_t^U - \alpha_p^U p_t + \psi^U,\end{aligned}$$

and the demand of noise traders is given by δ. Market clearing and the SLLN imply that the equilibrium price in Equation (17).

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 $\Delta\tilde{n}_t \equiv \Delta\overline{n}_t + \frac{1}{\overline{\alpha_n}}\Delta\delta_t$, which proves our claim.

B Detailed Data Description

This section describes in more detail the data used for the empirical implementation of the results in Section 4. 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 between trading volume and shares outstanding.

From FRED, we obtain 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 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 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 changes in payoffs as the log of one plus the year-on-year change in earnings divided by book equity.

We define our public signals as follows. The profitability ratio is the total operating profits of the trailing four quarter period divided by book equity lagged four quarters. The divided ratio is the total dividends of the trailing four quarter period divided by book equity lagged four quarters. Asset growth is the log growth of assets over the previous four quarters. Market beta is the coefficient of the stock's excess monthly returns against the S&P500's excess monthly returns over a rolling 5 year backward looking window containing at least 24 months of observations.

For analyst forecasts, we use all available, non-excluded 4-quarter ahead EPS forecasts from IBES. From this we construct forecasted earnings growth by subtracting the realized value of EPS for 4 quarters before the fiscal quarter the analyst is forecasting (i.e., if the analyst is forecasting quarter t, then the realized value comes from $t-4$). We then apply the same transformation as with realized earnings by normalizing by book equity, adding 1, and taking the logarithm. Similarly, forecasted future earnings growth uses the realized value from quarter t and forecasts for $t+4$.

Figure IA-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.

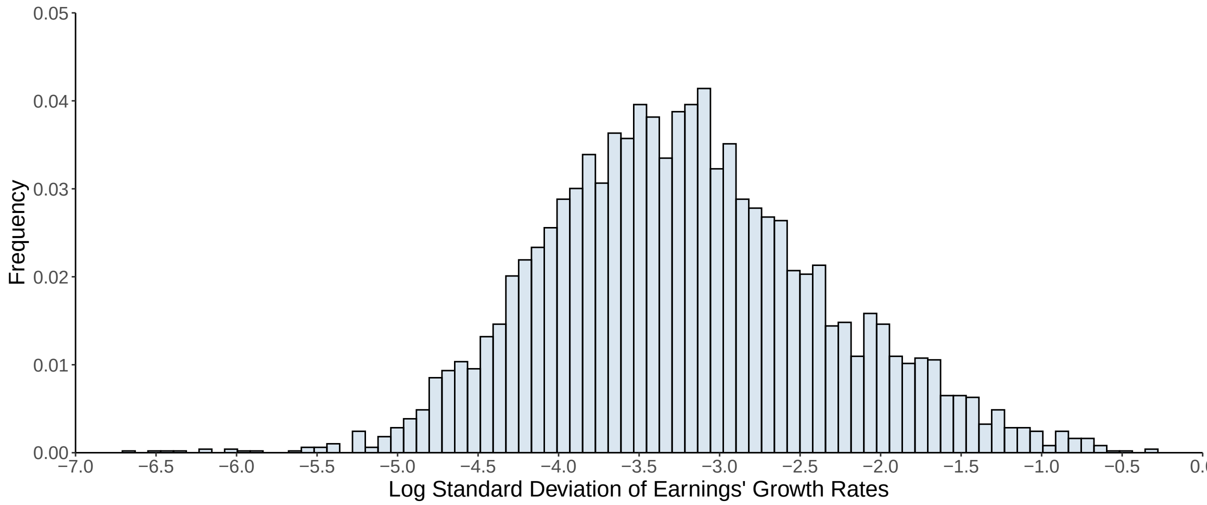


Figure IA-1: Cross-sectional standard deviation of earnings' growth rates

Note: Figure IA-1 shows a relative-frequency histogram of the distribution across stocks of the time-series standard deviation of earnings growth rates. This histogram features 6,803 stocks. For reference, 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.

C 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 additional cross-sectional results under the assumption that there is an unlearnable component of earnings. Finally, we present the cross-sectional correlation between price informativeness measured in year t and price informativeness measured in prior years.

C.1 Cross-sectional Relation: Graphical Illustration

Figures IA-2 through IA-6 are the counterparts 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. These figures make clear that the positive relationships between price informativeness and size, turnover, and institutional ownership are robust across time.

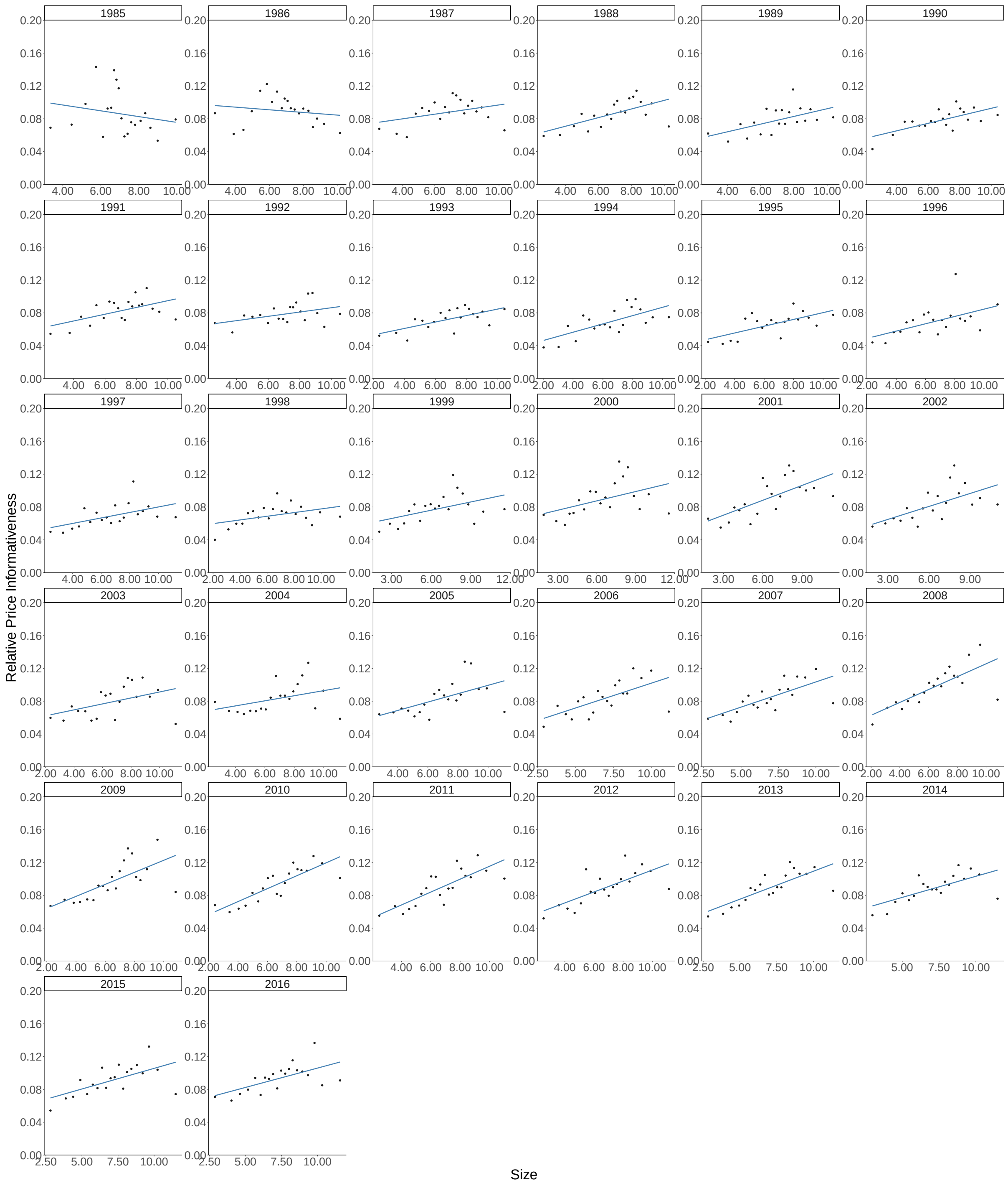


Figure IA-2: Price informativeness and size

Note: Figure IA-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 Table 2 can be interpreted as a weighted averaged of the year-by-year slope coefficient illustrated here.

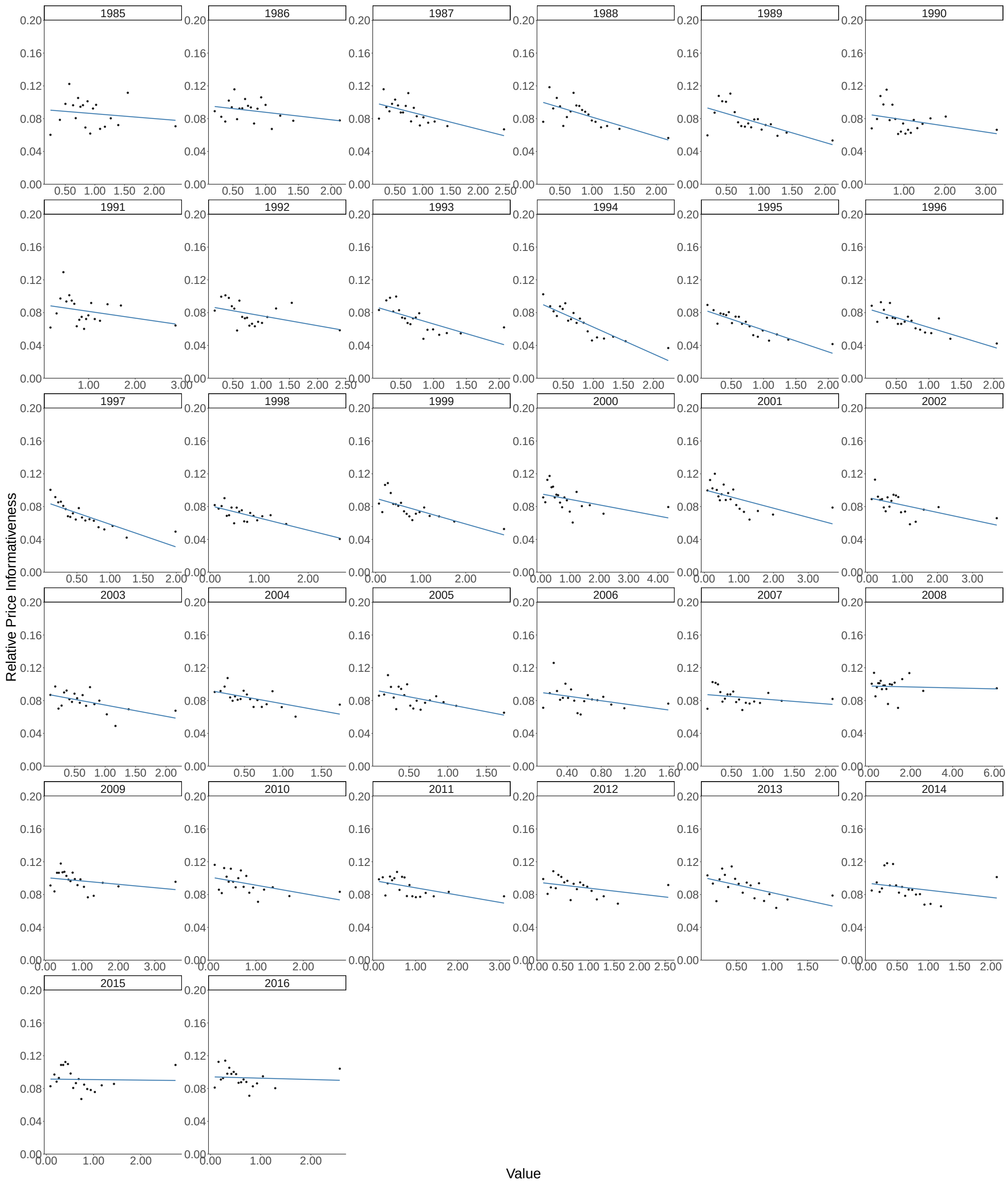


Figure IA-3: Price informativeness and value

Note: Figure IA-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 Table 2 can be interpreted as a weighted averaged of the year-by-year slope coefficient illustrated here.

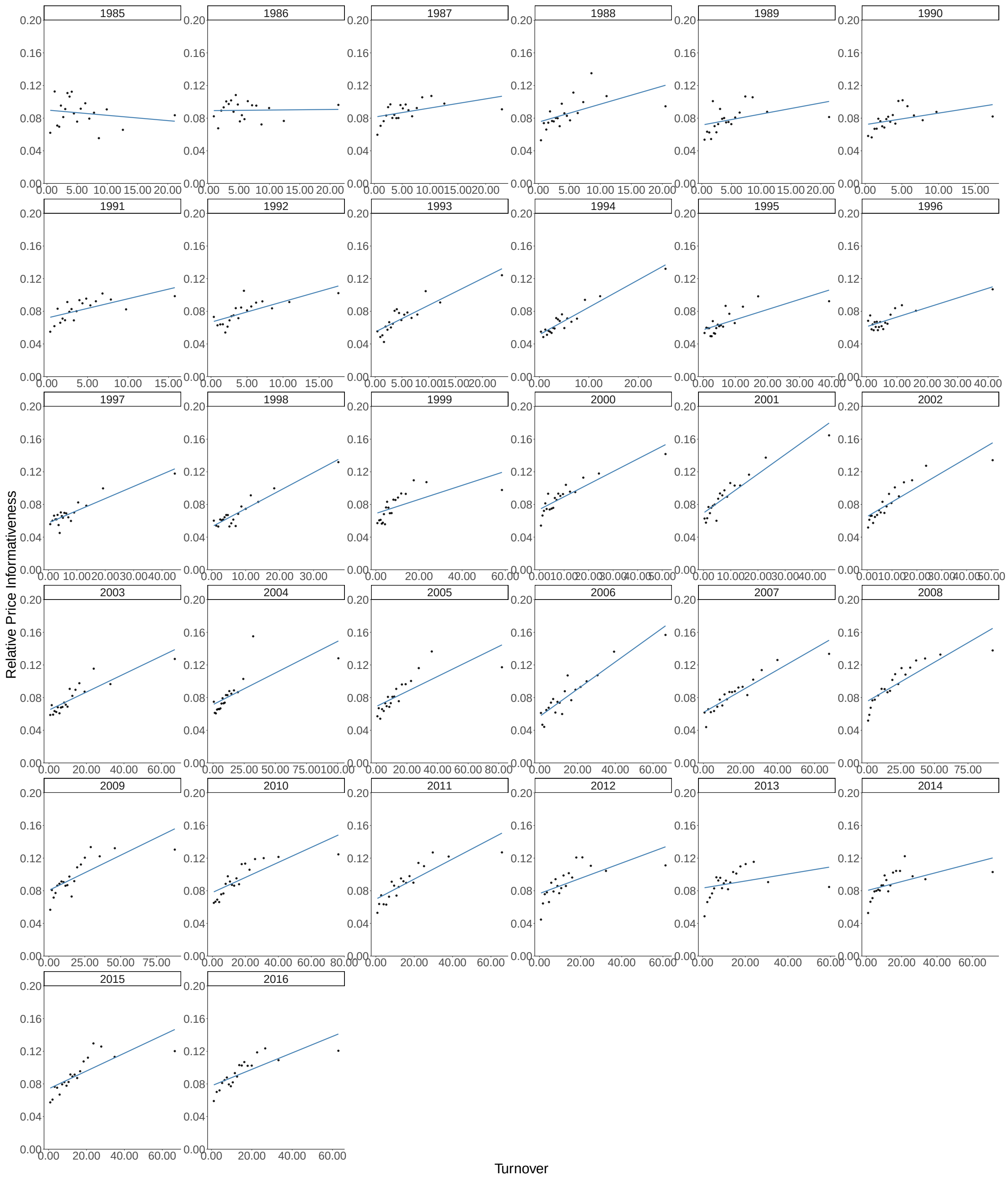


Figure IA-4: Price informativeness and turnover

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

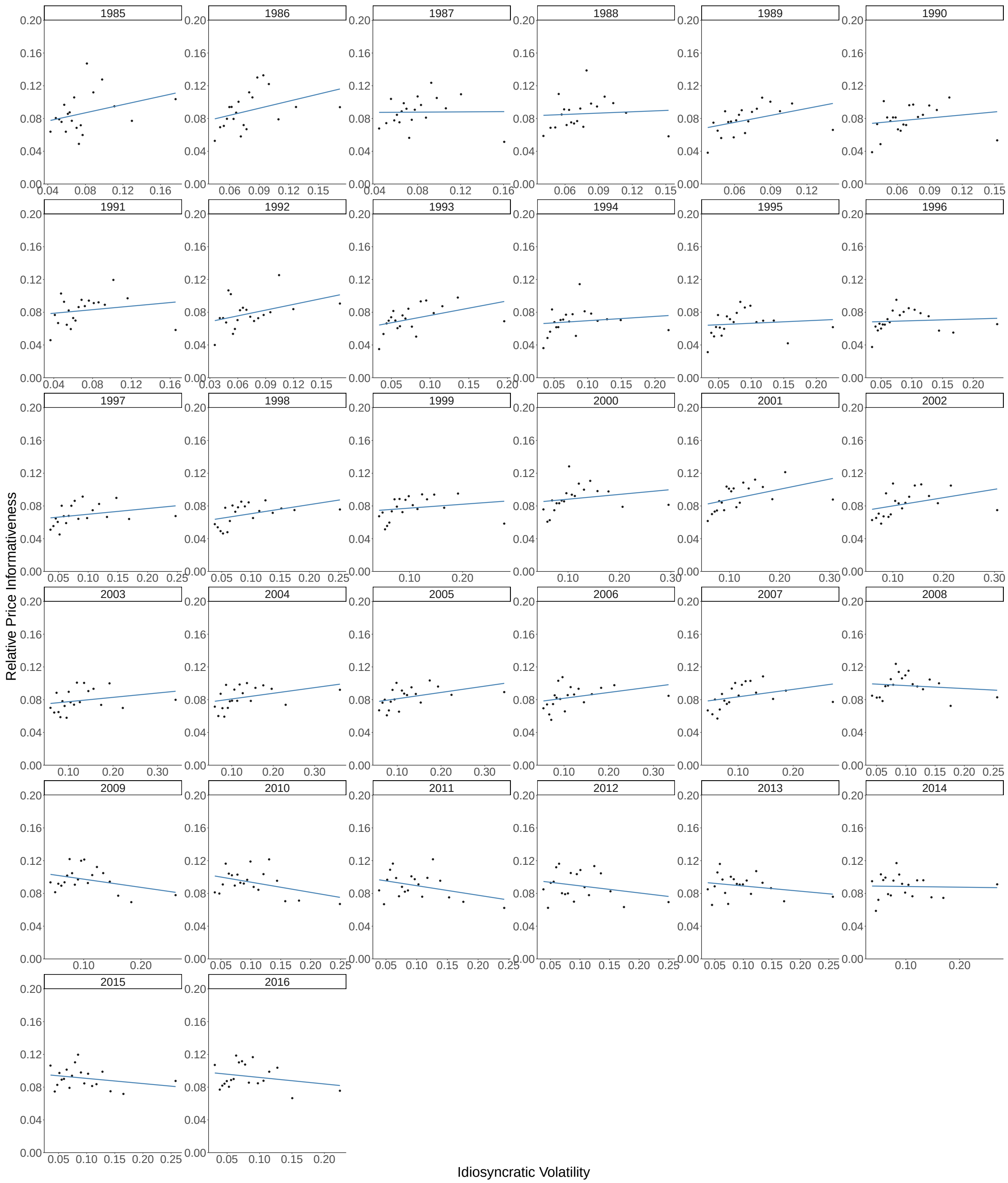


Figure IA-5: Price informativeness and idiosyncratic return volatility

Note: Figure IA-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.

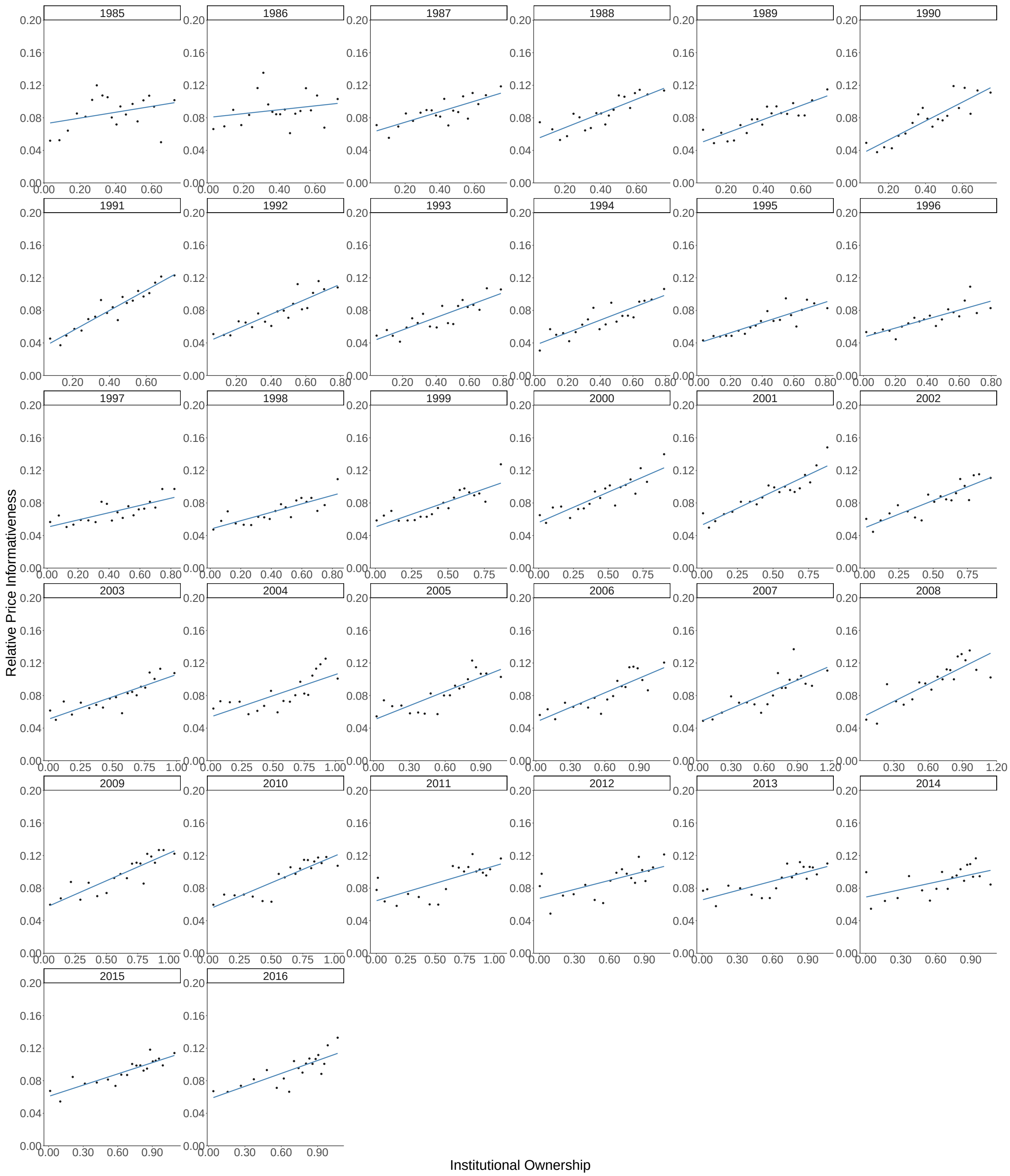


Figure IA-6: Price informativeness and institutional ownership

Note: Figure IA-6 shows year-by-year cross-sectional regressions of relative price informativeness (in twentiles) on institutional ownership, defined as the proportion of a stock 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.

C.2 Unlearnable Payoff Component

In this section, we provide additional results following the identification results in Proposition 10, which considers a payoff process with an unlearnable component.

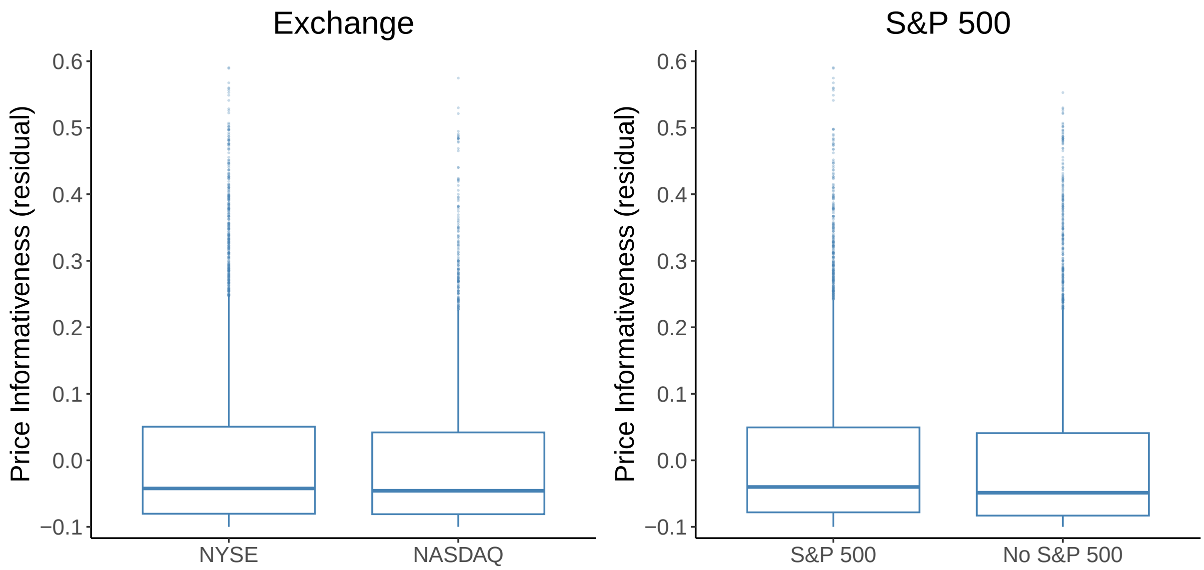


Figure IA-7: Cross-sectional results (2) - unlearnable component of earnings

Note: The left panel in Figure IA-7 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 IA-7 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 75th and 25th percentiles. The whiskers extend up to 1.5 times the interquartile range.

Table IA-1: Price informativeness summary statistics - unlearnable component of earnings

t	$Median$	$Mean$	SD	$Skew$	$Kurt$	$P5$	$P25$	$P75$	$P95$	n
1994	0.0198	0.0217	0.0203	0.2522	-1.3266	0.0019	0.0091	0.0324	0.0441	4
1995	0.0130	0.0343	0.0547	1.8893	1.8290	0.0023	0.0074	0.0277	0.1183	7
1996	0.0094	0.0296	0.0428	1.2605	0.0637	0.0014	0.0019	0.0386	0.0938	6
1997	0.0724	0.0891	0.0905	0.9956	0.6098	0.0011	0.0062	0.1461	0.2119	18
1998	0.0308	0.0543	0.0632	1.2803	0.6297	0.0006	0.0068	0.0844	0.1648	17
1999	0.0652	0.0836	0.0867	0.8544	-0.4005	0.0010	0.0066	0.1171	0.2343	22
2000	0.0387	0.0839	0.0932	1.1043	0.2803	0.0021	0.0098	0.1389	0.2306	21
2001	0.0461	0.0762	0.0823	1.6577	2.2211	0.0022	0.0222	0.0974	0.2660	26
2002	0.0447	0.0690	0.0732	1.9219	4.5972	0.0047	0.0187	0.0901	0.1747	37
2003	0.0299	0.0601	0.0763	1.9483	4.0100	0.0025	0.0075	0.0922	0.1947	46
2004	0.0202	0.0635	0.0842	1.7434	2.4242	0.0009	0.0059	0.0943	0.2599	52
2005	0.0512	0.0857	0.0967	1.6455	2.6865	0.0024	0.0128	0.1226	0.2680	62
2006	0.0629	0.0989	0.1067	1.4966	1.9529	0.0008	0.0162	0.1423	0.3101	75
2007	0.0580	0.1070	0.1226	1.5701	2.4518	0.0023	0.0155	0.1521	0.3621	85
2008	0.0559	0.1048	0.1295	1.8777	3.3902	0.0039	0.0185	0.1338	0.3920	110
2009	0.0432	0.0940	0.1205	1.9953	4.3382	0.0016	0.0123	0.1198	0.3411	136
2010	0.0390	0.0940	0.1149	1.6629	2.4799	0.0012	0.0122	0.1527	0.3258	170
2011	0.0514	0.0992	0.1136	1.5659	2.2905	0.0013	0.0156	0.1529	0.3246	219
2012	0.0545	0.0995	0.1171	1.8642	4.0107	0.0011	0.0181	0.1434	0.3167	259
2013	0.0518	0.0983	0.1178	1.9602	4.6162	0.0010	0.0154	0.1372	0.3077	309
2014	0.0406	0.0879	0.1140	2.1623	5.0558	0.0010	0.0132	0.1168	0.3310	349
2015	0.0454	0.0903	0.1087	1.8839	3.8100	0.0014	0.0137	0.1307	0.3229	403
2016	0.0548	0.0949	0.1062	1.5788	2.5071	0.0014	0.0136	0.1521	0.3183	444
2017	0.0534	0.0942	0.1058	1.6646	2.8589	0.0017	0.0165	0.1359	0.3282	482

Note: Table IA-1 reports year-by-year summary statistics on the panel of price informativeness measures recovered under the assumption of an unlearnable component of earnings. It provides information on the median; mean; standard deviation; skewness; excess kurtosis; and 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. Informativeness in year t is computed over a rolling window of 40 quarters prior.

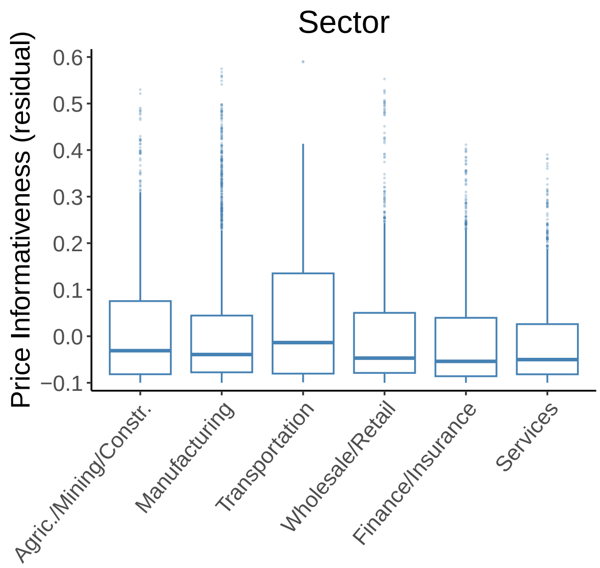


Figure IA-8: Cross-sectional results (3) - unlearnable component of earnings

Note: Figure IA-8 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 75th and 25th percentiles. The whiskers extend up to 1.5 times the interquartile range.

C.3 Persistence of Price Informativeness Estimates

Table IA-2 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 consistently above 0.7, while five-year cross-correlations can be as low as 0.2.

Table IA-2: Persistence of price informativeness

t	$\rho_{t,t-1}(\tau_\pi^R)$	$\rho_{t,t-2}(\tau_\pi^R)$	$\rho_{t,t-3}(\tau_\pi^R)$	$\rho_{t,t-4}(\tau_\pi^R)$	$\rho_{t,t-5}(\tau_\pi^R)$
1985	0.9407	0.9131	0.8807	0.7778	0.1711
1986	0.9117	0.7853	0.7942	0.8509	0.7809
1987	0.9160	0.7708	0.6230	0.6617	0.8335
1988	0.9060	0.7640	0.6627	0.5003	0.5155
1989	0.9128	0.7678	0.6233	0.5312	0.3969
1990	0.8922	0.7783	0.6381	0.5036	0.4044
1991	0.8871	0.7280	0.6374	0.4996	0.3721
1992	0.8750	0.7198	0.5976	0.5173	0.3799
1993	0.8172	0.6571	0.5138	0.4499	0.3846
1994	0.8857	0.6002	0.5085	0.3963	0.3294
1995	0.8901	0.7335	0.4908	0.4348	0.3563
1996	0.8641	0.7171	0.6054	0.3728	0.3151
1997	0.8646	0.6883	0.5868	0.4821	0.2138
1998	0.8423	0.6439	0.5253	0.4618	0.3766
1999	0.8461	0.6378	0.4923	0.3912	0.3709
2000	0.8293	0.6222	0.4690	0.3651	0.2760
2001	0.8681	0.6556	0.5171	0.4203	0.3254
2002	0.8876	0.7256	0.5468	0.4535	0.3801
2003	0.8929	0.7440	0.6160	0.4805	0.3949
2004	0.9109	0.7681	0.6328	0.5130	0.3923
2005	0.9233	0.7915	0.6752	0.5503	0.4397
2006	0.9259	0.8264	0.7038	0.6116	0.5044
2007	0.9043	0.8017	0.7212	0.6083	0.5449
2008	0.8497	0.7168	0.6481	0.5759	0.4787
2009	0.8742	0.6493	0.5318	0.4812	0.4165
2010	0.8617	0.6742	0.4946	0.4140	0.3746
2011	0.8959	0.7100	0.5310	0.4037	0.3566
2012	0.9027	0.7524	0.6045	0.4432	0.3151
2013	0.8850	0.7467	0.6306	0.5163	0.3803
2014	0.8990	0.7409	0.6380	0.5257	0.4282
2015	0.9068	0.7916	0.6460	0.5551	0.4596
2016	0.9241	0.7855	0.6865	0.5787	0.5082

Note: Table IA-2 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 40 quarters prior.

C.4 Correlation of Price Informativeness Estimates

Table IA-3 reports the correlation matrix among the different estimates of price informativeness. This table shows that there is a robust correlation between informativeness estimates with a without controls.

The relation between the baseline and those using average analyst forecast as a proxy for the learnable component of the payoff, as in Section 4.2, is also positive but less strong.

Table IA-3: Correlation of price informativeness estimates

	Baseline	Unlearnable	No Controls	Unlearnable No Controls
Baseline	1	–	–	–
Unlearnable	0.234	1	–	–
No Controls	0.838	0.225	1	–
Unlearnable No Controls	0.163	0.839	0.210	1

Note: Table IA-3 reports the correlation among the different estimated measures of informativeness.

D Empirical Implementation: Price as only signal

While in the body of the paper we report measures of price informativeness that include additional controls — public signals, in the language of Section 2 — in this section we report the results without controls, from the perspective of an external observer who only sees the price and past payoffs. Formally, the results reported here are the outcome of running the following two regressions

$$\begin{aligned}
\Delta p_t^j &= \overline{\beta}^j + \beta_0^j \Delta x_t^j + \beta_1^j \Delta x_{t+4}^j + \varepsilon_t^j &&\Rightarrow R^{2,j}_{\Delta x, \Delta x'} \\
\Delta p_t^j &= \overline{\zeta}^j + \zeta_0^j \Delta x_t^j \qquad\qquad\qquad + \hat{\varepsilon}_t^j &&\Rightarrow R^{2,j}_{\Delta x} \,.
\end{aligned}$$

Figure IA-9, 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 IA-10, which is the counterpart of Figure 5, shows the time-series evolution of the cross-sectional mean, median, and standard deviation of relative price informativeness. Table IA-4 and Figures IA-11 and IA-12, 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.

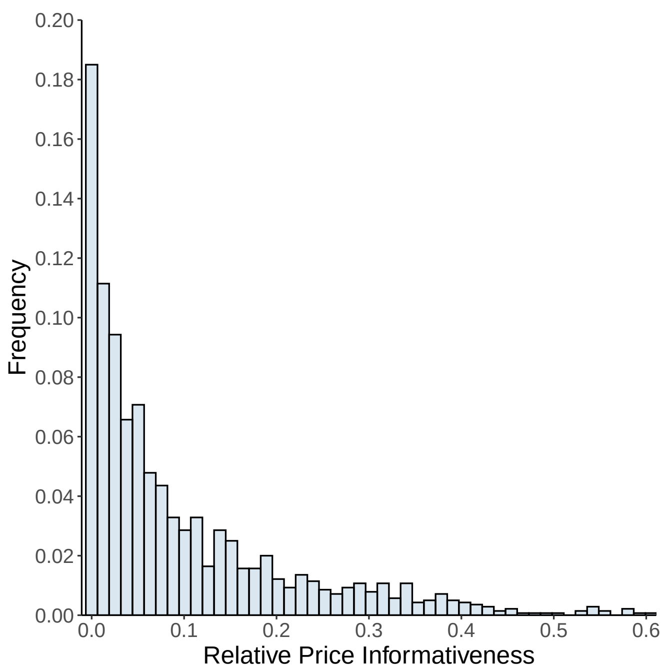


Figure IA-9: Price informativeness: relative-frequency histogram, no public signals

Note: Figure IA-9 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 40 quarters prior.

Table IA-4: Cross-sectional results, no public signals

	Estimate	Std. Error	t-stat
Size	0.002270	0.000186	12.19
Value	0.000083	0.000519	0.16
Turnover	0.000454	0.000029	15.75
Idiosyncratic Volatility	0.040384	0.008264	4.89
Institutional Ownership	0.024064	0.001449	16.61
Analysts Covering	0.001251	0.000084	14.95

Note: Table IA-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): $\tau_{\pi,t}^{R,b} = a_0^c + a_1^c c_t^b + \xi_t + \epsilon_{b,t}$, 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 IA-2 through IA-6 provide the graphical counterpart of the results in this table. Size is measured as the natural log of stock market capitalization, value is measured as the ratio between a stock's book value and its market capitalization, turnover is measured as the ratio between trading volume and 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 a stock held by institutional investors.

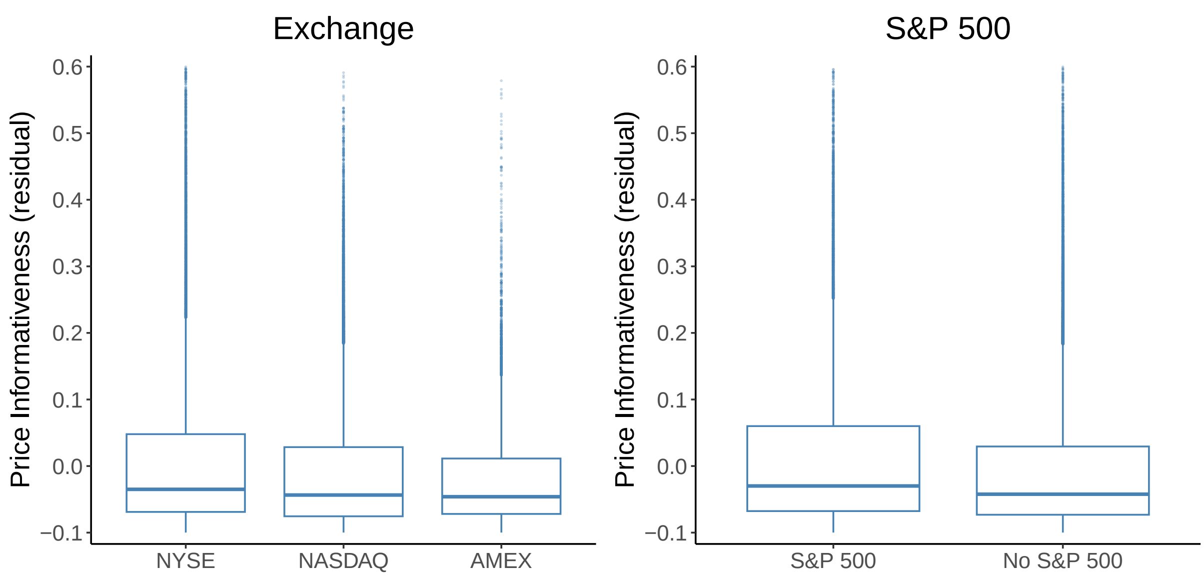


Figure IA-11: Cross-sectional results, no public signals

Note: The left panel in Figure IA-11 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 IA-11 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 75th and 25th percentiles. The whiskers extend up to 1.5 times the interquartile range.

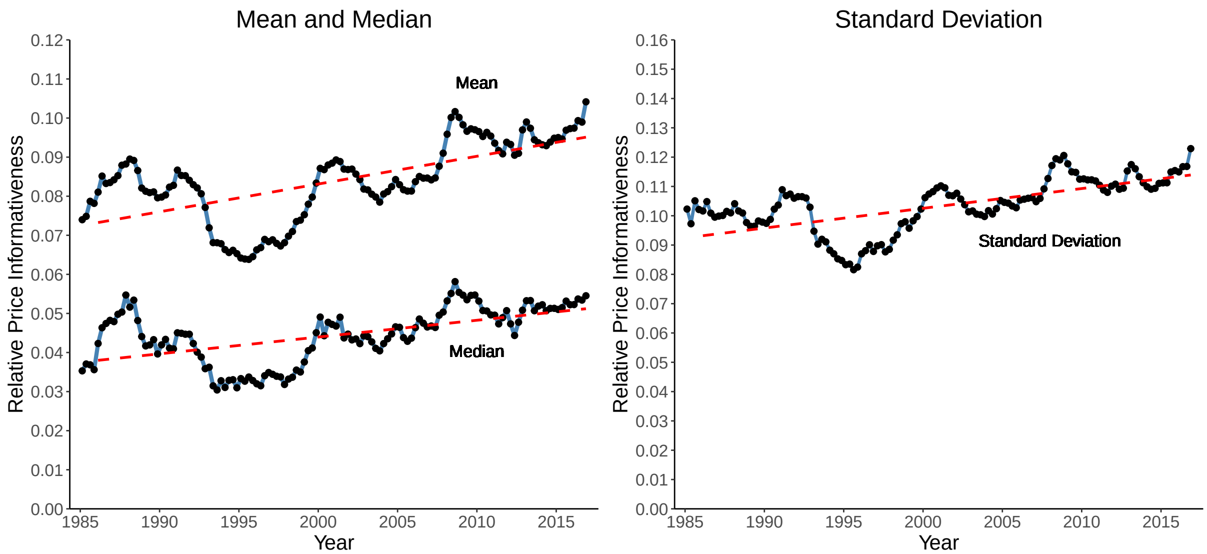


Figure IA-10: Price informativeness over time, no public signals

Note: The left panel in Figure IA-10 shows the time-series evolution of the cross-sectional mean and median relative price informativeness. The right panel in Figure IA-10 shows the time-series evolution of the cross-sectional standard deviation of price informativeness. The red dashed lines show linear trends starting in 1986.

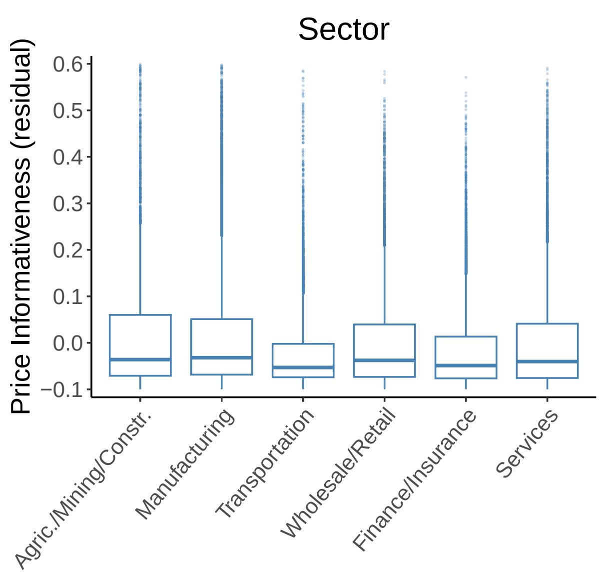


Figure IA-12: Cross-sectional results, no public signals

Note: Figure IA-12 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 75th and 25th percentiles. The whiskers extend up to 1.5 times the interquartile range.

E Additional Results

E.1 Absolute Price Informativeness

For simplicity, we provide our results in the absence of public signals and when investors have information about the payoff one period ahead.

Proposition 4. (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, then*

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

where $\Delta p_t = p_t - p_{t-1}$ denotes the date t change in log-price, and $\Delta x_t = x_t - x_{t-1}$ 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, $\operatorname{plim}\left(\widehat{\tau_\pi}\right) = \operatorname{plim}\left(\frac{\left(\widehat{\beta_1}\right)^2}{\widehat{\sigma_e^2}}\right) = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_{\Delta n} = \tau_\pi$. $\square$

F Structural Interpretation of Empirical Findings

It is possible to interpret the empirical findings presented in Section 4 through the lens of the general framework developed in Section 2 and the structural models developed in Section 3. If one were merely interested in knowing the precision of the signal contained in asset prices about future payoffs, our empirical results directly conclude that such signal is more precise for large, high turnover, and high institutional ownership stocks, and has become more precise on average over the last few decades.

However, one may be interested in translating these empirical patterns of informativeness to particular elements of a model. For simplicity, we consider the case without public signals and when investors have information about the payoff one period ahead. In this case, we can express relative price informativeness as

$$\tau_{\pi}^{R} = \frac{1}{1 + \left(\frac{\phi_n}{\phi_1}\right)^2 \frac{\tau_u}{\tau_{\Delta n}}}. \tag{35}$$

From Equation (35), we can conclude that stocks with high informativeness are those with a high value of ϕ_1 (sensitivity of the asset price to the future payoff) relative to ϕ_n (sensitivity of the asset price to its non-payoff relevant component) and/or a high value of τ_u^{-1} (variance of the innovation to the payoff) relative to $\tau_{\Delta n}^{-1}$ (variance of the non-payoff relevant component, i.e., noise). Therefore, our empirical results imply that the (either of the) ratios $\frac{\phi_1}{\phi_n}$ and $\frac{\tau_{\Delta n}^{-1}}{\tau_u^{-1}}$ must be higher for large, high turnover, and high institutional ownership stocks, and that (either of) such ratios must have increased on average over the last few decades. Equation (35) clearly highlights that price informativeness captures the signal-to-noise ratio in asset prices, but not the sources of noise or information independently.

The models developed in Section 3 allow us to go one step further by relating our empirical findings on informativeness to deeper primitives. In all three models, the ratio $\frac{\phi_1}{\phi_n}$ corresponds to $\frac{\overline{\alpha_s}}{\overline{\alpha_n}}$, which denotes the ratio of the aggregate demand sensitivities to information and noise, respectively. Consequently, across all three models, higher price informativeness can be interpreted as either a higher $\frac{\overline{\alpha_s}}{\overline{\alpha_n}}$ and/or a higher $\frac{\tau_{\Delta n}}{\tau_u}$. While τ_u is a primitive in all three models, $\overline{\alpha_s}$, $\overline{\alpha_n}$, and, in some cases, $\tau_{\Delta n}$, are equilibrium objects, as we explain below.

The first model considered in Section 3, in which noise arises from investors' sentiment, provides the clearest connection between relative price informativeness and model primitives in the context of a fully structural model. In this model, the aggregate demand sensitivity to information relative to noise is exactly given by the ratio of the precision of investors' private signals (τ_s) about the future payoff relative to the precision of the innovation (τ_u), that is, $\frac{\overline{\alpha_s}}{\overline{\alpha_n}} = \frac{\tau_s}{\tau_u}$. In this model, the noise embedded in the price is only coming from the investors' sentiment and $\tau_{\Delta n}$ is also a primitive of the model. Therefore, price informativeness can be expressed as the following combination of primitives:

$$\tau_{\pi}^{R} = \frac{1}{1 + \left(\frac{\tau_u}{\tau_s}\right)^2 \frac{\tau_u}{\tau_{\Delta n}}}. \tag{36}$$

Equation (36) implies that price informativeness is increasing in the precision of investors' private signals (τ_s), decreasing in the volatility of the payoff innovation (τ_u^{-1}), and decreasing in the volatility of aggregate noise ($\tau_{\Delta n}^{-1}$). Through the lens of this model, one interpretation of our empirical results is that investors have more precise private information about stocks with higher market capitalization and high turnover. It is conceivable that investors acquire more private information about stocks with higher market capitalization and high turnover because they can benefit from such information at a larger scale. However, this conclusion is not obvious, since one may conjecture that larger firms attract the attention of more unsophisticated traders, which would make the prices of those stocks noisy and uninformative, or that high turnover stocks feature a large number of noise traders, thus engendering low price informativeness. Similarly, our time series empirical findings are consistent with an increase in the average precision of private information relative to noise over the last few decades. Similar arguments can be given for the other characteristics, e.g., value, institutional ownership, idiosyncratic volatility.

In the second model considered in Section 3, which features a representative agent, the relative aggregate demand sensitivity to noise and information is also given $\frac{\overline{\alpha_s}}{\overline{\alpha_n}} = \frac{\tau_s}{\tau_u}$, but in this case the precision of the noise embedded in the price $\tau_{\Delta n}$ is endogenous, which makes the connection between informativeness and primitives less direct. As in the model with sentiment, we show that price informativeness is increasing in the precision of investors' private signals (τ_s) and decreasing in the volatility of the innovation (τ_u^{-1}). All else equal, it is also the case that price informativeness is decreasing in the volatility of aggregate noise ($\tau_{\Delta n}^{-1}$). Therefore, the interpretation of the results is almost identical to the interpretation of the model with sentiment as noise. We should note that if we had allowed for time varying risk aversion, the movements in discount rates could be interpreted as changes in the (endogenous) volatility of $\tau_{\Delta n}$.

Finally, the model with informed, uninformed, and noise traders delivers similar implications to the model with sentiment. In this last model, price informativeness is increasing in the fraction of informed investors and decreasing in the volatility of noise trading. Through the lens of this model, our empirical results can be interpreted as concluding that large, high turnover, and high institutional ownership stocks feature a higher share of informed investors relative to noise traders, and that the share of informed investors has increased over the last few decades in relation to the volatility of noise trading.

G Alternative Modeling Frameworks

Our identification results extend 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. Third, we also provide the respective CARA-Normal models to microfound these exact linear formulations.

G.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

$$x_{t+1} = \mu_x + \rho x_t + u_t, \tag{37}$$

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 $\mathbb{V}\text{ar}\left[u_t\right] = \sigma_u^2 = \tau_u^{-1}$, and are identically and independently distributed over time.[17] We assume that the equilibrium price is given by

$$p_t = \overline{\phi} + \phi_0 x_t + \phi_1 x_{t+1} + \phi_n n_t, \tag{38}$$

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 $\mathbb{V}\text{ar}\left[\varepsilon_t^n\right] = \sigma_n^2 = \tau_n^{-1}$. 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 level*, which we denote by $\tilde{\pi}_t$, is given by

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

and absolute and relative price informativeness are respectively given by

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

Proposition 5. (Identifying price informativeness: log-linear case)

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

[17] 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.

recovered by

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

The OLS estimation of Regression R1-LL 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-LL. Let R^2_x, 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, $\tau^R_{\tilde{\pi}}$, can be recovered by

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

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

Proof. a) By comparing Regression R1-LL with the structural Equation (38), 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, $\sigma_e^2 = \mathbb{V}\text{ar}[e_t] = (\phi_n)^2 \mathbb{V}\text{ar}[\varepsilon_t^n] = (\phi_n)^2 \tau_n^{-1}$. 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 (37) and (38), 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, $\text{plim}\left(\widehat{\tau_{\tilde{\pi}}}\right) = \text{plim}\left(\frac{\left(\widehat{\beta_1}\right)^2}{\widehat{\sigma_e^2}}\right) = \left(\frac{\phi_1}{\phi_n}\right)^2 \tau_n = \tau_\pi$.

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

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

After substituting Equation (37) in Regression R1-LL, the following relation holds

$$p_t = \overline{\phi} + \phi_1 \mu_x + \phi_n \mu_n + (\phi_0 + \rho \phi_1) x_t + \phi_1 u_t + \phi_n \varepsilon_t^n. \tag{39}$$

By comparing Regression R2-LL with the structural Equation (39), 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 (39), the following variance decomposition must hold

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

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

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

Therefore, relative price informativeness can be written as

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

□

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_1^i, with flow utility given by $U_i\left(w_1^i\right)$, 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 $x_t = \ln\left(X_t\right)$ and which trades at a (log) price $p_t = \ln\left(P_t\right)$. The process followed by x_t is given by

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

where $\Delta x_{t+1} = x_{t+1} - x_t$, 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 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_0^i, and optimally chooses a portfolio share in the risky asset, denoted by θ_t^i, to solve

$$\max_{\theta_t^i} \mathbb{E}_t^i \left[U_i \left(w_1^i \right) \right] \tag{40}$$

subject to a wealth accumulation constraint

$$w_1^i = \left(R^f + \theta_t^i \left(\frac{X_{t+1} + P_{t+1}}{P_t} - R^f \right) \right) w_0^i, \tag{41}$$

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

The optimality condition of an investor who maximizes Equation (40) subject to the wealth accumulation constraint in Equation (41) is given by

$$\mathbb{E} \left[U_i' \left(w_1^i \right) \left(\frac{X_{t+1} + P_{t+1}}{P_t} - R^f \right) \middle| \mathcal{I}_t^i \right] = 0. \tag{42}$$

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 $U' \left(w_1^i \right)$ around the current date t wealth level w_0^i. Formally, we approximate $U' \left(w_1^i \right)$ as follows

$$U' \left(w_1^i \right) \approx U' \left(w_0^i \right) + U'' \left(w_0^i \right) \Delta w_1^i,$$

which allows us to express Equation (42) as

$$U' \left(w_0^i \right) \mathbb{E}_i \left[\frac{X_{t+1} + P_{t+1}}{P_t} - R^f \right] + U'' \left(w_0^i \right) w_0^i \mathbb{E}_i \left[\left(\left(R^f - 1 \right) + \theta_t^i \left(\frac{X_{t+1} + P_{t+1}}{P_t} - R^f \right) \right) \left(\frac{X_{t+1} + P_{t+1}}{P_t} - R^f \right) \right] \approx 0.$$

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 (42) as

$$U' \left(w_0^i \right) \mathbb{E}_t^i \left[\frac{X_{t+1} + P_{t+1}}{P_t} - R^f \right] + U'' \left(w_0^i \right) w_0^i \theta_t^i \mathbb{V}\mathrm{ar}_t^i \left[\frac{X_{t+1} + P_{t+1}}{P_t} \right] \approx 0.$$

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

$$\theta_t^i \approx \frac{1}{\gamma^i} \frac{\mathbb{E}_t^i \left[\frac{X_{t+1} + P_{t+1}}{P_t} - R^f \right]}{\mathbb{V}\mathrm{ar}_t^i \left[\frac{X_{t+1} + P_{t+1}}{P_t} \right]}, \tag{43}$$

where $\gamma^i \equiv -\frac{w^i U''(w^i)}{U'(w^i)}$ 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

$$\frac{X_{t+1} + P_{t+1}}{P_t} = e^{\ln \left(\frac{\left(1 + \frac{P_{t+1}}{X_{t+1}} \right) \frac{X_{t+1}}{X_t}}{\frac{P_t}{X_t}} \right)},$$

and

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

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

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

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

Therefore, starting from Equation (43), we have that the risky asset demand of an investor i can be approximated as

$$\theta_t^i\approx\frac{1}{\gamma^i}\frac{k_0+k_1\mathbb{E}_t^i\left[p_{t+1}-x_{t+1}\right]+\mathbb{E}_t^i\left[\Delta x_{t+1}\right]-\left(p_t-x_t\right)-r^f}{\mathbb{V}\mathrm{ar}\left[k_1\left(p_{t+1}-x_{t+1}\right)+\Delta x_{t+1}|\mathcal{I}_t^i\right]},\tag{44}$$

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 $\mathbb{V}\mathrm{ar}\left[k_1\left(p_{t+1}-x_{t+1}\right)+\Delta x_{t+1}\right]$ is a constant, which we denote by V. Under this conjecture, θ_t^i is a linear function of s_t^i, x_t, and $\overline{n}_t^i$, 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 $\int\theta_t^i w_0^i di=Q$ implies

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

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

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

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

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

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

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

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,

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

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

In this equilibrium, our guess in Equation (44) is verified and the equilibrium price is linear and can be expressed as in Equation (38).

G.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.

G.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

$$\Delta X_{t+1} = \mu_{\Delta X} + \rho \Delta X_t + u_t, \tag{45}$$

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 $\mathbb{V}\text{ar}\left[u_t\right] = \sigma_u^2 = \tau_u^{-1}$, and are identically and independently distributed over time. We assume that the equilibrium price difference is given by

$$\Delta P_t = \overline{\phi} + \phi_0 \Delta X_t + \phi_1 \Delta X_{t+1} + \phi_n \Delta n_t, \tag{46}$$

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 $\mathbb{V}\text{ar}\left[\varepsilon_t^{\Delta n}\right] = \sigma_n^2 = \tau_{\Delta n}^{-1}$. 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

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

and absolute and relative price informativeness are given by

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

Proposition 6. (Identifying price informativeness: difference-stationary linear case)

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

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

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,

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

Then, relative price informativeness, τ_{Π}^{R}, can be recovered by

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

The OLS estimation of Regressions R1-Linear-Diff and R2-Linear-Diff 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 (46), 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 (45) and (46), 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-Diff 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-Diff 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 (45) in Regression R1-Linear-Diff, the following relation holds:

$$\Delta P_t = \overline{\phi} + \phi_1 \mu_{\Delta X} + \phi_n \mu_{\Delta n} + \left(\phi_0 + \rho \phi_1\right) \Delta X_t + \phi_1 u_t + \phi_n \varepsilon_t^{\Delta n}. \tag{47}$$

By comparing Regression R2-Linear-Diff with the structural Equation (47), 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 (47), 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:

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

Therefore, relative price informativeness can be written as

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

$\square$

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 its last period wealth. An investor born at time t has preferences given by

$$U_i\left(w_{t+1}\right) = -e^{-\gamma_i w_{t+1}},$$

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 a price P_t in period t.[18] 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 of 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 Dávila and Parlatore (2021) 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 $\Delta n_t = \mu_{\Delta n} + \varepsilon_t^{\Delta n}$ with $\varepsilon_t^{\Delta n} \sim N\left(0, \tau_{\Delta n}^{-1}\right)$. 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.[19]

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 $i \neq j$, 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

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

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

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

[18] To simplify notation, we denote the risk-free rate by R, instead of R^f as we did in the body of the paper.
[19] Dávila and Parlatore (2021) show that the equilibrium structure is preserved if this assumption is relaxed.

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

$$Q_t^i = \alpha_X^i X_t + \alpha_s^i s_t^i + \alpha_n^i \overline{n}_t^i - \alpha_P^i P_t + \psi^i, \tag{48}$$

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

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

where

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

with price informativeness given by

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

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

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

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

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

Then, the first-order condition is given by

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

Matching coefficients we have

$$\begin{aligned}
\alpha_s^i &= \frac{\left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa^i}\mathbb{V}\text{ar}\left[u_t|\mathcal{I}_t^i\right]\tau_s \qquad (49)\\
\alpha_n^i &= \frac{\left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa^i}\mathbb{V}\text{ar}\left[u_t|\mathcal{I}_t^i\right]\tau_\eta \\
\alpha_X^i &= \frac{\left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa^i}\left(1 - \mathbb{V}\text{ar}\left[u_t|\mathcal{I}_t^i\right]\tau_\Pi \frac{\overline{\alpha_X}}{\overline{\alpha_s}}\right) \\
\alpha_P^i &= \frac{1}{\kappa^i}\left(1 - \left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)\mathbb{V}\text{ar}\left[u_t|\mathcal{I}_t^i\right]\tau_\Pi \frac{\overline{\alpha_p}}{\overline{\alpha_s}}\right) \\
\psi^i &= -\frac{1}{\kappa^i}\left(\left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)\mathbb{V}\text{ar}\left[u_t|\mathcal{I}_t^i\right]\tau_\Pi\left(\frac{\overline{\alpha_n}}{\overline{\alpha_s}}\mu_{\Delta n} + \frac{\overline{\psi}}{\overline{\alpha_s}}\right) - R^{-1}\left(\frac{\overline{\alpha_n}}{\overline{\alpha_P}}\mu_{\Delta n} + \frac{\overline{\psi}}{\overline{\alpha_P}}\right)\right),
\end{aligned}$$

where

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

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

Then, an equilibrium in linear strategies exists if the system above has a solution. In this equilibrium, our guess in Equation (48) is verified and the equilibrium price is linear and can be expressed as in Equation (51).

Note that the if the investors are ex-ante identical, the demand sensitivities are the same for all i. Then, there exists a unique solution to the system in Equations (49) 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

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

G.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

$$X_{t+1} = \mu_X + \rho X_t + u_t, \tag{50}$$

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 $\mathbb{V}\mathrm{ar}[u_t] = \sigma_u^2 = \tau_u^{-1}$, and are identically and independently distributed over time. We assume that the equilibrium price is given by

$$P_t = \overline{\phi} + \phi_0 X_t + \phi_1 X_{t+1} + \phi_n n_t, \tag{51}$$

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 $\mathbb{E}[\varepsilon_t^n] = 0$ and $\mathbb{V}\mathrm{ar}[\varepsilon_t^n] = \sigma_n^2 = \tau_n^{-1}$. 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

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

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 7. (Identifying price informativeness: difference-stationary linear case)

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 $\sigma_e^2 = \mathbb{V}ar[e_t]$ denotes the variance of the error. Then, absolute price informativeness, $\tau_{\hat{\Pi}}$, can be recovered by

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

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

$$\tau_{\hat{\Pi}}^R = \frac{R^2_{X,X'} - R^2_X}{1 - R^2_X}.$$

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

Proof. a) By comparing Regression R1-Linear with the structural Equation (38), 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 (50) and (51), 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 (50) in Equation (51), the following relation holds

$$P_t = \overline{\phi} + \phi_1 \mu_X + \phi_n \mu_n + \left(\phi_0 + \rho \phi_1\right) X_t + \phi_1 u_t + \phi_n \varepsilon_t^n. \tag{52}$$

By comparing Regression R2-Linear with the structural Equation (52), 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 (52), 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

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

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 its last period wealth. An investor born at time t has preferences given by

$$U_i\left(w_{t+1}\right) = -e^{-\gamma^i w_{t+1}},$$

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 a 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 of 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 Dávila and Parlatore (2021) 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 $\varepsilon_t^n \sim N\left(0, \tau_n^{-1}\right)$. Note that 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.[20]

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 $i \neq j$, and $u_t \perp \varepsilon_{st}^i$ 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:

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

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

The optimality condition for an investor i in period t satisfies

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

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

$$\Delta Q_t^i = \alpha_X^i X_t + \alpha_s^i s_t^i + \alpha_n^i \overline{n}_t^i - \alpha_P^i P_t + \psi^i, \tag{53}$$

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

[20] Dávila and Parlatore (2021) show that the equilibrium structure is preserved if this assumption is relaxed.

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 $\overline{\alpha_X} = \int \alpha_X^i di$, $\overline{\alpha_s} = \int \alpha_s^i di$, $\overline{\alpha_P} = \int \alpha_P^i di$, and $\overline{\psi} = \int \psi^i di - Q$.

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

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

where

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

with price informativeness given by

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

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

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

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

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

Then, the first-order condition is given by

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

Matching coefficients we have

$$
\begin{aligned}
\alpha_s^i &= \frac{\left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa^i} \mathbb{V}\text{ar}\left[u_t | \mathcal{I}_t^i\right] \tau_s \\
\alpha_n^i &= \frac{\left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa^i} \mathbb{V}\text{ar}\left[u_t | \mathcal{I}_t^i\right] \tau_\eta \\
\alpha_X^i &= \frac{\left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right)}{\kappa^i} \left(\rho - \mathbb{V}\text{ar}\left[u_t | \mathcal{I}_t^i\right] \tau_{\hat{\Pi}} \frac{\overline{\alpha_X}}{\overline{\alpha_s}}\right) \\
\alpha_P^i &= \frac{1}{\kappa^i}\left(1 - \left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right) \mathbb{V}\text{ar}\left[u_t | \mathcal{I}_t^i\right] \tau_{\hat{\Pi}} \frac{\overline{\alpha_p}}{\overline{\alpha_s}}\right) \\
\psi^i &= -\frac{1}{\kappa^i}\left(\left(1 + R^{-1}\frac{\overline{\alpha_X}}{\overline{\alpha_P}}\right) \mathbb{V}\text{ar}\left[u_t | \mathcal{I}_t^i\right] \tau_{\hat{\Pi}} \left(\frac{\overline{\alpha_n}}{\overline{\alpha_s}}\mu_n + \frac{\overline{\psi}}{\overline{\alpha_s}}\right) - R^{-1}\left(\frac{\overline{\alpha_n}}{\overline{\alpha_P}}\mu_n + \frac{\overline{\psi}}{\overline{\alpha_P}}\right)\right),
\end{aligned}
\tag{54}
$$

where

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

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

Then, an equilibrium in linear strategies exists if the system above has a solution. In this equilibrium, our guess in Equation (53) is verified and the equilibrium price is linear and can be expressed as in Equation (51).

Note that when investors are ex-ante identical, the demand sensitivities are the same for all i. Then, there exists a unique solution to the system in Equations (54), 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

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