Measuring the quantity and value of the average investor's private information

Philip Bond (University of Washington)*

December 2021

Abstract

I use market-clearing identities to derive sufficient statistics formulae for the quantity and value of the average investor's private information. The inputs required to operationalize these formulae are closely related to the outputs of the price-dividend predictability literature. Using estimates from this literature suggests that the average investor possesses substantial information; but that little of this information is impounded in prices; and that an uninformed investor would experience only a very small increase in returns if given access to this information.

*I thank Alan Morrison, Yiyao Wang, and seminar audiences at Emory, EPFL, ITAM, OxFit, SUFE, UNC, and the University of Washington for helpful comments.

1 Introduction

The standard view of financial markets is that they aggregate the dispersed private information of participating investors (Grossman 1976). The objective of this paper is simple to state, and is to give quantitative answers to the following questions: How much private information do investors have? How much of this information gets impounded into prices? And how valuable is this information, as measured by its effect on portfolio returns?

In this paper I develop a methodology for answering these questions. The methodology yields formulae based on a small number of sufficient statistics. These formulae emerge from a relatively general framework in which investors trade both because of differing expectations about returns, and because of shocks to desired holdings that resemble discount rate shocks.[1] The formulae arise from the exploitation of the market clearing condition, coupled with a focus on the first two moments of all relevant distributions.

I focus on information about aggregates. In this case, the sufficient statistics that enter the formulae I derive for the quantity and value of investor information are closely related to the objects estimated by the literature that studies the predictability of market returns and dividends using the price-dividend ratio. I operationalize the formulae for the quantity and value of investor information using the estimates of Binsbergen and Koijen (2010).[2] The formulae I derive make clear the relevance of conditioning on the history of prices and dividends when forming expectations, and a distinguishing feature of Binsbergen and Koijen (2010) is to present a tractable method of doing exactly this.

The formulae I derive yield the following findings, given inputs from Binsbergen and Koijen's empirical analysis. First, the average investor has substantial information: quantitively, this information reduces an investor's perceived variance of future dividends by approximately 10%. Second, the information possessed by the average investor is large relative to the information contained in prices: quantitively, information in prices reduces perceived variance of future dividends by only approximately 1.5%. Third, the value of the average investor's information is nonetheless small: an investor who is affected by discount rate shocks in the same way as the average investor, but who lacks the average investor's information about future dividends, would realize a return that is just a handful of basis points lower than that of the average investor.

The economic underpinning of the formula for the average investor's information is as follows. The price-dividend predictability literature has found that high ratios predict low

[1] See Diamond and Verrecchia (1981) and its subsequent dynamic extensions, e.g., Watanabe (2008), Biais et al (2010).

[2] Closely related to these authors' approach, see the survey of Koijen and Van Nieuwerburgh (2011), and Cochrane (2011).

returns. Since subsequent dividends are observed, an econometrician can isolate fluctuations in the price-dividend ratio due to fluctuations in future dividends that are forecastable by investors. The only effect of positive innovations to future dividends on the econometrician's information is to raise current prices. Hence positive innovations to future dividends reduce the econometrician's expectation of future returns. But for market clearing to hold, positive innovations to future dividends *cannot* reduce the average investor's expectation of future returns. By quantifying this observation, I quantify the amount of information that the average investor must possess.

Related literature:

This paper is related to the large literature on price-dividend predictability. That literature adopts the (typically implicit) view that investors are symmetrically informed, and possess information unobservable to the econometrician. A separate and primarily theoretical literature has studied the process by which prices come to contain information, and emphasizes the idea that investors observe independent and noisy signals of economic fundamentals, which are aggregated into the price (classical references are Hayek 1945, Grossman 1976, Hellwig 1980). I link these literatures by showing how estimates from the price-dividend predictability literature can be used to infer how much information dispersed investors have.

Related, a significant literature quantifies the information content of prices (e.g., Bai, Savov, Philippon 2015). In this paper I use estimates of the information content of prices to infer the information possessed by individual investors.

This paper is complementary to Kadan and Manela (2019), who quantify the value to an investor of a given signal, for example, learning the content of a macroeconomic employment report before other investors. Instead, in this paper I use observed correlations between prices and subsequent returns and dividends to infer how much information the average investor possesses. I also estimate the value of information in a very different way to Kadan and Manela, specifically, via a second application of market-clearing implications.

This paper is likewise complementary to Egan, MacKay, and Yang (2021), who use a revealed-preference approach to infer the expectations of investors who buy index funds. Theirs is a partial equilibrium approach in the sense that it takes fees as given. In contrast, the approach in this paper depends critically on market-clearing arguments. Relative to Egan, MacKay, and Yang, this paper has the advantage of shedding light on the information of the average of *all* investors in the market; but the disadvantage of saying nothing about heterogeneity among investors.

The Probability of Informed Trade (PIN; Easley, Kiefer, and O'Hara 1996) quantifies the fraction of trade stemming from informed traders, and as such measures the extensive margin of information. In contrast, the measure in this paper captures both intensive and extensive margins. The measure in the current paper is also constructed entirely from pricing and dividend data, and does not require the use of order flow information, and as such is robust to changes in trading patterns (such as the proliferation of trading venues and the increasing prevalence of high-frequency traders). Also related, Kyle's lambda (Kyle 1985) is frequently estimated, and used as a proxy for the prevalence of informed trade; but it is challenging to relate the estimated value to a cardinal measure of the amount of informed trade.

The framework used to derive sufficient statistics formulae is related to Watanabe (2008) and Biais et al (2010). These papers note that uninformed investors will (rationally) experience below-market returns, because they increase their holdings when future returns are low. Glode (2011) and Savov (2014) use related observations to rationalize investment in active mutual funds with negative alphas, along with providing some evidence. The same economic force operates in this paper. Nonetheless, my estimates suggest that the "underperformance" of an uninformed investor is quantitively small.

Kurlat (2019) derives and implements a sufficient statistics formula for the ratio of private to social value of information in what is essentially the origination part of financial markets. This paper instead examines the amount and private value of information in a secondary financial market. Bond and García (2021) theoretically characterize the social value of private information in a related setting.

2 Framework

I derive sufficient statistics formulae using a general framework in which investors trade both because of differing expectations about returns, and because of shocks to desired holdings that resemble discount rate shocks. The framework is closely related to the canonical models of Grossman and Stiglitz (1980), Hellwig (1980), and especially Diamond and Verrecchia (1981). Following these papers, all random variables below are normal. As such, all results should be interpreted as approximations based on the first two moments of distributions.

The framework features a single risky asset. In the empirical implementation I will consider the S&P 500 index. I conjecture that many of the insights of the paper can be extended to multi-asset models.

A unit continuum of investors, indexed by i, trade a risky asset and a risk-free asset. The gross return of the risk-free asset is R_t, and the price of the risky asset at date t is P_t. The risk-free asset is in zero supply, and the supply of the risky asset is normalized to 1. The risky asset pays dividends D_t at the start of each period t. The (absolute) excess return on the risky asset from date t to $t+1$ is

$$X_{t+1} \equiv P_{t+1} + D_{t+1} - R_t P_t.$$

Write $\mathcal{I}_{i,t}$ for investor i's information at date t. Let investor i's demand $q_{i,t}$ for the asset be a function of the expected return, $E\left[X_{t+1}|\mathcal{I}_{i,t}\right]$, and factors unrelated to returns, $Z_t + u_{i,t}$:

$$q_{i,t} = A_i E\left[X_{t+1}|\mathcal{I}_{i,t}\right] - B_i\left(Z_t + u_{i,t}\right). \tag{1}$$

In (1), A_i and B_i are potentially equilibrium objects. In particular, in the standard mean-variance framework, A_i depends on the combination of investor i's risk tolerance and $var\left[X_{t+1}|\mathcal{I}_{i,t}\right]$, where the latter is an equilibrium object.

The term Z_t is an aggregate shock to investors' desired asset holdings. As such, Z_t shifts prices independent of expectations of dividends. Following the literature, I will typically refer to Z_t as a discount rate shock. Similarly, $u_{i,t}$ is an investor specific shock to desired holdings.

Write $\mathcal{H}_t$ for the history of exogenous cash flows and aggregate discount rate shocks, i.e., $\mathcal{H}_t = \{D_t, Z_{t-1}, D_{t-1}, Z_{t-2}, D_{t-2}, \ldots\}$. Define the innovations

$$\begin{aligned} \epsilon_{D,t} &= D_{t+1} - E\left[D_{t+1}|\mathcal{H}_t\right] \\ \epsilon_{Z,t} &= Z_t - E\left[Z_t|\mathcal{H}_t\right]. \end{aligned}$$

Assume $\epsilon_{D,t}$ and $\epsilon_{Z,t}$ are stationary, uncorrelated, and normally distributed. As noted, the normality assumption means that all results should be interpreted as approximations based on the first two moments of distributions.

The assumption that $\epsilon_{D,t}$ and $\epsilon_{Z,t}$ are uncorrelated is important. While the analytical results below can be generalized to allow for correlation (notes available upon request), these generalizations are much harder to empirically implement. The assumption that dividend and discount rate innovations are uncorrelated fits well with how the literature has conceived of the origins of discount rate fluctuations (e.g., see review in Cochrane 2011). Even in a consumption-based asset pricing paper such as Bansal and Yaron (2004), discount rate fluctuations stem largely from fluctuations in cash flow volatility that are assumed to be uncorrelated with cash flow innovations.

Investor i observes at date t a private signal of date $t+1$ dividends,

$$y_{i,t} = \epsilon_{D,t} + \epsilon_{i,t},$$

where $\epsilon_{i,t} \sim N\left(0, \tau_i^{-1}\right)$. In particular, τ_i is the precision of investor i's private signal. Note that the shock $Z_t + u_{i,t}$ both directly affects investor i's asset demand, and also serves as a signal about Z_t.

By standard arguments, in equilibrium the price innovation $P_t - E\left[P_t|\mathcal{H}_t\right]$ is linear in the exogenous innovations to dividends and aggregate discount rates,

$$P_t - E\left[P_t|\mathcal{H}_t\right] = c_D \epsilon_{D,t} + c_Z \epsilon_{Z,t}.$$

Investor i's information $\mathcal{I}_{i,t}$ consists of $\mathcal{H}_t$, combined with prices $\{P_t, P_{t-1}, \ldots\}$, private shocks to asset demand, $\{Z_t + u_{i,t}, Z_{t-1} + u_{i,t-1}\}$, and private dividend signals $\{y_{i,t}, y_{i,t-1}, \ldots\}$. The inclusion of the lagged realization of the aggregate discount rate, i.e., Z_{t-1}, in $\mathcal{I}_{i,t}$ reflects the fact that this realization can be inferred from the history of dividend and price realizations. But an individual investor i does not know the contemporaneous aggregate discount rate shock; instead, investors know only their own trading preferences, $Z_t + u_{i,t}$.

It is useful to keep track of two further information sets. First, $\mathcal{I}_{0,i,t}$ is investor i's information excluding the sequence of private signals about dividends $y_{i,t}$. For investors for whom private signals are worthless, $\tau_i = 0$, the information sets $\mathcal{I}_{0,i,t}$ and $\mathcal{I}_{i,t}$ coincide. I refer to an investor with information $\mathcal{I}_{0,i,t}$ as being uninformed; though such an investor has a signal $Z_t + u_{i,t}$ about the discount rate innovation $\epsilon_{Z,t}$. Second, $\mathcal{J}_t$ is the econometrician's information set, which consists of $\mathcal{H}_t$ augmented with the history of prices $\{P_t, P_{t-1}, \ldots\}$.

As written, the only public signals are the realizations of prices and dividends. It is straightforward to extend the framework to incorporate additional public signals, such as public macroeconomic announcements. Under such an extension, the required inputs would be the outputs of predictability regressions that incorporate the same set of public announcements.

3 Conditional distributions

I start by collecting results on the conditional distributions of $\begin{pmatrix} \epsilon_{D,t} \\ \epsilon_{Z,t} \end{pmatrix}$ under different conditioning information. In particular, the moments of $\begin{pmatrix} \epsilon_{D,t} \\ \epsilon_{Z,t} \end{pmatrix}$ conditional on $\mathcal{I}_{i,t}$, $\mathcal{I}_{0,i,t}$, and $\mathcal{J}_t$ have simple linear relations.

An input for the results in this section is the observation that, because $\mathcal{I}_{i,t}$, $\mathcal{I}_{0,i,t}$, and $\mathcal{J}_t$ all include the date t price P_t, and because the price is determined by the innovations $\epsilon_{D,t}$ and $\epsilon_{Z,t}$, the variance-covariance matrix of $\begin{pmatrix} \epsilon_{D,t} \\ \epsilon_{Z,t} \end{pmatrix}$ conditional on each of these information sets is singular. Formally, this follows from (using $\mathcal{J}_t$ as an example)

$$\begin{aligned} var\left[\epsilon_{Z,t}|\mathcal{J}_t\right] &= \left(\frac{c_D}{c_Z}\right)^2 var\left[\epsilon_{D,t}|\mathcal{J}_t\right] \\ cov\left[\epsilon_{Z,t},\epsilon_{D,t}|\mathcal{J}_t\right] &= -\frac{c_D}{c_Z} var\left[\epsilon_{D,t}|\mathcal{J}_t\right], \end{aligned} \tag{2}$$

implying

$$var\left[\epsilon_{Z,t}|\mathcal{J}_t\right] var\left[\epsilon_{D,t}|\mathcal{J}_t\right] - cov\left[\epsilon_{Z,t},\epsilon_{D,t}|\mathcal{J}_t\right]^2 = 0. \tag{3}$$

The results in this section follow from this observation, combined with manipulation of standard updating rules for normally distributed random variables.

Relative to the econometrician's information set $\mathcal{J}_t$, $\mathcal{I}_{0,i,t}$ contains $Z_t + u_{i,t}$. Define

$$\Upsilon = \frac{var\left[\epsilon_{Z,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{Z,t}|\mathcal{J}_t\right] + var\left[u_{i,t}\right]},$$

which measures an uninformed investor's informational advantage relative to the econometrician. It ranges from 0 (if $var\left[u_{i,t}\right] = \infty$, no information advantage) to 1 (if $var\left[u_{i,t}\right] = 0$, maximal information advantage since an uninformed investor knows $\epsilon_{Z,t}$ and hence can infer $\epsilon_{D,t}$ from the price P_t).

Lemma 1

$$var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] = \left(1-\tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right) var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{0,i,t}\right] \tag{4}$$

$$var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{0,i,t}\right] = \left(1-\Upsilon\right) var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{J}_{t}\right]. \tag{5}$$

A key ingredient for equilibrium relations is how the sensitivity of the conditional expectation of $\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}$ to its true value depends on the information set. The relation is most easily expressed in terms of forecast errors:

Lemma 2

$$\begin{aligned} \frac{\partial}{\partial\epsilon_{D,t}}\left(E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right]-\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right) &= \left(1-\tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right)\frac{\partial}{\partial\epsilon_{D,t}}\left(E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{0,i,t}\right]-\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right. \\ \frac{\partial}{\partial\epsilon_{Z,t}}\left(E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right]-\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right) &= \left(1-\tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right)\frac{\partial}{\partial\epsilon_{Z,t}}\left(E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{0,i,t}\right]-\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right. \end{aligned}$$

$$\frac{\partial}{\partial \epsilon_{D,t}}\left(E\left[\left(\begin{array}{c}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{array}\right)|\mathcal{I}_{0,i,t}\right]-\left(\begin{array}{c}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{array}\right)\right) = (1-\Upsilon)\frac{\partial}{\partial \epsilon_{D,t}}\left(E\left[\left(\begin{array}{c}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{array}\right)|\mathcal{J}_{t}\right]-\left(\begin{array}{c}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{array}\right)\right)$$
$$\frac{\partial}{\partial \epsilon_{Z,t}}\left(E\left[\left(\begin{array}{c}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{array}\right)|\mathcal{I}_{0,i,t}\right]-\left(\begin{array}{c}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{array}\right)\right) = (1-\Upsilon)\frac{\partial}{\partial \epsilon_{Z,t}}\left(E\left[\left(\begin{array}{c}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{array}\right)|\mathcal{J}_{t}\right]-\left(\begin{array}{c}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{array}\right)\right).$$

Finally, the following is helpful for parameterizing Υ, an uninformed investor's informational advantage relative to the econometrician:

Lemma 3

$$\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}\right]}+\frac{var\left[\epsilon_{Z,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{Z,t}\right]}=1.$$

In words: The econometrician's information set includes the price P_t, which is a function of dividend innovations $\epsilon_{D,t}$ and discount rate innovations $\epsilon_{Z,t}$. Lemma 3 says that the more information that the price contains about $\epsilon_{D,t}$, the less it contains about $\epsilon_{Z,t}$.

The immediate implication of Lemma 3 is that an uninformed investor's informational advantage Υ relative to the econometrician is given by

$$\Upsilon=\frac{1-\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}\right]}}{1-\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}\right]}+\frac{var\left[u_{i,t}\right]}{var\left[\epsilon_{Z,t}\right]}}. \tag{6}$$

4 Private information of the average investor

4.1 The average investor

The general framework allows for a great deal of heterogeneity of investors; in particular, it places no restrictions on the distribution of characteristics (A_i, B_i, τ_i) over the investor population.

Nonetheless, the equilibrium price coincides with the equilibrium price in an economy in which agents are ex ante identical, i.e., share a common $\left(\bar{A}, \bar{B}, \bar{\tau}\right)$. This representative-agent characterization of the economy is simply a modest generalization of various results in the existing literature.

The market-clearing condition is

$$\int_i q_{i,t} di = 1,$$

which implies

$$\int_i \frac{\partial q_{i,t}}{\partial \epsilon_{D,t}} di = 0 \tag{7}$$

$$\int_i \frac{\partial q_{i,t}}{\partial \epsilon_{Z,t}} di = 0. \tag{8}$$

A trivial decomposition of $\frac{\partial q_{i,t}}{\partial \epsilon_{D,t}}$ and $\frac{\partial q_{i,t}}{\partial \epsilon_{Z,t}}$ is

$$\begin{aligned}
\frac{\partial q_{i,t}}{\partial \epsilon_{D,t}} &= A_i \frac{\partial X_{t+1}}{\partial \epsilon_{D,t}} - A_i \frac{\partial}{\partial \epsilon_{D,t}} \left(X_{t+1} - E\left[X_{t+1} | \mathcal{I}_{i,t} \right] \right) \\
\frac{\partial q_{i,t}}{\partial \epsilon_{Z,t}} &= \left(A_i \frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}} - B_i \right) - A_i \frac{\partial}{\partial \epsilon_{Z,t}} \left(X_{t+1} - E\left[X_{t+1} | \mathcal{I}_{i,t} \right] \right).
\end{aligned}$$

Here, the first term in each expression is the change in asset demand of investor who perfectly forecasts the return X_{t+1}, and the second term represents the underreaction stemming from imperfect information.

Combined with Lemma 2, and using the fact that the information sets $\mathcal{I}_{i,t}$ and $\mathcal{I}_{0,i,t}$ both contain the contemporaneous price P_t, these decompositions yield

$$\frac{\partial q_{i,t}}{\partial \epsilon_{D,t}} = A_i \frac{\partial X_{t+1}}{\partial \epsilon_{D,t}} - A_i \left(1 - \tau_i var\left[\epsilon_{D,t} | \mathcal{I}_{i,t} \right] \right) \frac{\partial}{\partial \epsilon_{D,t}} \left(X_{t+1} - E\left[X_{t+1} | \mathcal{I}_{0,i,t} \right] \right) \tag{9}$$

$$\frac{\partial q_{i,t}}{\partial \epsilon_{Z,t}} = \left(A_i \frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}} - B_i \right) - A_i \left(1 - \tau_i var\left[\epsilon_{D,t} | \mathcal{I}_{i,t} \right] \right) \frac{\partial}{\partial \epsilon_{Z,t}} \left(X_{t+1} - E\left[X_{t+1} | \mathcal{I}_{0,i,t} \right] \right). \tag{10}$$

Substituting into the market-clearing conditions (7) and (8) delivers the following representative agent result. Here, the notation $var\left[\cdot | \mathcal{I}_{i,t}(\bar{\tau}) \right]$ means the conditional variance as perceived by an investor who observes private signals of precision $\bar{\tau}$.

Lemma 4 *The equilibrium price coefficients c_D and c_Z coincide with those in an economy in which all investors are ex ante identical, i.e., $(A_i, B_i, \tau_i) = \left(\bar{A}, \bar{B}, \bar{\tau} \right)$ for all i, where*

$$\begin{aligned}
\bar{\tau} var\left[\epsilon_{D,t} | \mathcal{I}_{i,t}(\bar{\tau}) \right] &= \frac{\int A_i \tau_i var\left[\epsilon_{D,t} | \mathcal{I}_{i,t} \right]}{\int A_i} \\
\frac{\bar{B}}{\bar{A}} &= \frac{\int B_i}{\int A_i}.
\end{aligned}$$

I will refer to the $\bar{\tau}$ defined by Lemma 4 as the precision of the average investor's private information. As the lemma establishes, it is determined as a weighted average of the cross-section of signal precisions, where the weights are the coefficients A_i that determine the sensitivity of investor i's trade to expectations about the excess return. For use throughout, define

$$\mathcal{T} \equiv \bar{\tau} var\left[\epsilon_{D,t} | \mathcal{I}_{i,t}\left(\bar{\tau}\right)\right],$$

which measures the information advantage of the average investor relative to an uninformed investor (see Lemma 1, or (12) below). Observe that $\mathcal{T}$ ranges from 0, corresponding to no information advantage, to 1, corresponding to $\bar{\tau} = \infty$, i.e., the average investor perfectly observes the dividend innovation $\epsilon_{D,t}$.

4.2 Measuring the information of the average investor

The main observation of this section is that the precision of the average investor's private information can be expressed in terms of sufficient statistics that can be estimated using only aggregate data.

The key step is the application of market-clearing. Condition (7) for the average investor is

$$\bar{A}\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}} - \bar{A}\left(1 - \mathcal{T}\right)\frac{\partial}{\partial \epsilon_{D,t}}\left(X_{t+1} - E\left[X_{t+1} | \mathcal{I}_{0,i,t}\right]\right) = 0, \tag{11}$$

which rewrites to

$$1 - \mathcal{T} = \frac{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}}{\frac{\partial}{\partial \epsilon_{D,t}}\left(X_{t+1} - E\left[X_{t+1} | \mathcal{I}_{0,i,t}\right]\right)}.$$

By Lemma 2, and using the fact that the information sets $\mathcal{I}_{0,i,t}$ and $\mathcal{J}_t$ both contain the contemporaneous price P_t,

$$\frac{\partial}{\partial \epsilon_{D,t}}\left(X_{t+1} - E\left[X_{t+1} | \mathcal{I}_{0,i,t}\right]\right) = \left(1 - \Upsilon\right)\frac{\partial}{\partial \epsilon_{D,t}}\left(X_{t+1} - E\left[X_{t+1} | \mathcal{J}_t\right]\right),$$

implying

$$\left(1 - \mathcal{T}\right)\left(1 - \Upsilon\right) = \frac{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}}{\frac{\partial}{\partial \epsilon_{D,t}}\left(X_{t+1} - E\left[X_{t+1} | \mathcal{J}_t\right]\right)}.$$

By construction, the dividend innovation $\epsilon_{D,t}$ affects the econometrician's information set $\mathcal{J}_t$ only via the price P_t. Hence:

Proposition 1 *The average investor's informational advantage is given by*

$$\left(1 - \mathcal{T}\right)\left(1 - \Upsilon\right) = \frac{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}} - \frac{\partial E\left[X_{t+1} | \mathcal{J}_t\right]}{\partial P_t}\frac{\partial P_t}{\partial \epsilon_{D,t}}}.$$

Proposition 1 says that the average investor's informational advantage, as given by the combination of $\mathcal{T}$ (the precision of the average investor's signal about dividend innovations) and Υ (the informational advantage of investors, who observe their own discount rate shocks $Z_t + u_{i,t}$, relative to the econometrician) can be inferred from the objects $\frac{\partial P_t}{\partial \epsilon_{D,t}}$, $\frac{\partial E[X_{t+1}|\mathcal{J}_t]}{\partial P_t}$, and $\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}$.

In words, $\frac{\partial P_t}{\partial \epsilon_{D,t}}$ is the relation between today's price and the innovations to next period's dividend. It is related to ability of prices to forecast future dividends. Likewise, $\frac{\partial E[X_{t+1}|\mathcal{J}_t]}{\partial P_t}$ is related to the ability of today's price to forecast next period's return. Finally, $\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}$ is a combination of $\frac{\partial P_t}{\partial \epsilon_{D,t}}$, the direct effect of the dividend innovation $\epsilon_{D,t}$ on D_{t+1}, which is simply 1, and the effect of the dividend innovation $\epsilon_{D,t}$ on P_{t+1}, which is determined largely by the persistence of dividend innovations.

Section 5 details how to relate these objects to the outputs of a typical predictability analysis.

The economic idea behind Proposition 1 is as follows. Suppose that higher prices today lead an econometrician to forecast lower returns, as the price-dividend ratio literature suggests. But by definition, the average investor's holding doesn't change, since the average investor must continue to hold the market supply. For this to happen, it must be the case that the average investor's expectation differs from the econometrician's. In particular, the average investor must observe private signals that are on average positive when prices are high. Proposition 1 quantifies this statement.

To interpret Proposition 1: From Lemma 1,

$$1 - \mathcal{T} = \frac{var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}(\bar{\tau})\right]}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}(\bar{\tau})\right]}, \quad (12)$$

$$1 - \Upsilon = \frac{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}(\bar{\tau})\right]}{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}, \quad (13)$$

i.e., the percentage reductions in the conditional variance of $\epsilon_{D,t}$ associated with observing a signal with the precision of the average investor's signal, and associated with an uninformed investor's informational advantage relative to the econometrician.

4.3 Information in the price relative to the average investor's information

The information of the average investor is given by Proposition 1. How does this quantity of information compare to the information in the price?

The average investor forecasts dividends using the combination of his/her private signal; his/her private information about the discount rate; and the date t price. Together, these three pieces of information affected an investor's perceived variance of dividends according to

$$\frac{var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}(\bar{\tau})\right]}{var\left[\epsilon_{D,t}|D_t,\mathcal{J}_{t-1}\right]} = \frac{var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}(\bar{\tau})\right]}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}(\bar{\tau})\right]}\frac{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}(\bar{\tau})\right]}{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}|D_t,\mathcal{J}_{t-1}\right]}. \tag{14}$$

The combination of the first two terms in this decomposition is given by Proposition 1. The final term, $\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}|D_t,\mathcal{J}_{t-1}\right]}$, measures the information in the price, and can be estimated.

4.4 The value of private information

What is the value of the average investor's private information? I calculate by how much giving an uninformed investor access to the average investor's information would increase the uninformed investor's expected return. Observe that this exercise is well-defined regardless of whether or not uninformed investors are actually present in the market.

The benefit of focusing on return differentials rather than, for example, willingness-to-pay measures of utility differences, is that it is possible to give a sufficient statistics formula for the return differential. In contrast, I have been unable to find a sufficient statistics formula for willingness-to-pay in which the components can be estimated. In particular, a utility-based measure would require estimates of investor risk aversion.

To isolate the value of information, I evaluate this expected return differential for an investor who resembles the average in other dimensions. Concretely, I consider an investor with characteristics

$$\left(\bar{A},\bar{B}\right) = \left(\int_i A_i di, \int_i B_i di\right), \tag{15}$$

and compare the expected return from investment strategies made under the information of the average investor,

$$q_{\bar{\tau},t} \equiv \bar{A}E\left[X_{t+1}|\mathcal{I}_{i,t}\left(\bar{\tau}\right)\right] - \bar{B}\left(Z_t + u_{i,t}\right),$$

and under the information of the uninformed investor,

$$q_{0,t} \equiv \bar{A}E\left[X_{t+1}|\mathcal{I}_{0,i,t}\right] - \bar{B}\left(Z_t + u_{i,t}\right).$$

Getting access to the average investor's information would raise an uninformed investor's expected return by a fraction

$$V \equiv \frac{E\left[q_{\bar{\tau},t}X_{t+1}\right]}{E\left[q_{0,t}X_{t+1}\right]}. \tag{16}$$

As a first step in evaluating V, note that an investor with characteristics $\left(\bar{A},\bar{B},\bar{\tau}\right)$ is a representative agent for the economy, in the sense of Lemma 4. So as a direct consequence of Lemma 4, the investment strategy of such an investor is independent of the aggregate shocks $\epsilon_{D,t}$ and $\epsilon_{Z,t}$. This is a version of the "average investor theorem" of Sharpe (1991).

Corollary 1 $\frac{\partial q_{\bar{\tau},t}}{\partial \epsilon_{D,t}} = \frac{\partial q_{\bar{\tau},t}}{\partial \epsilon_{Z,t}} = 0.$

Corollary 1 and some straightforward manipulation and substitution (see proof of Proposition 2) yields

$$V = \frac{1}{1 + \frac{cov\left[\frac{q_{0,t}}{\bar{A}}, X_{t+1}\right]}{E[X_{t+1}]^2}}. \tag{17}$$

Expression (17) relates the value of information to the covariance between an uninformed investor's asset position, and returns. Decomposing this covariance into the parts attributable to dividend and discount rate innovations gives

$$cov\left[\frac{q_{0,t}}{\bar{A}}, X_{t+1}\right] = \frac{\frac{1}{\bar{A}}\frac{\partial q_{0,t}}{\partial \epsilon_{D,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}}\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right] + \frac{\frac{1}{\bar{A}}\frac{\partial q_{0,t}}{\partial \epsilon_{Z,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}}\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right]. \tag{18}$$

The key challenge in evaluating (18) is that while $\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right]$ can be estimated—loosely speaking, as a residual, i.e., the variance of date t expected excess returns that isn't attributable to cash flow innovations—this leaves the term $\frac{\frac{1}{\bar{A}}\frac{\partial q_{0,t}}{\partial \epsilon_{Z,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}}$. In particular, this term depends on the ratio $\bar{B}/\bar{A}$, which is unobservable.

The key step in characterizing the value of information is to use the market-clearing condition (8) to infer $\frac{\frac{1}{\bar{A}}\frac{\partial q_{0,t}}{\partial \epsilon_{Z,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}}$. Doing so yields a sufficient statistics expression for the value of information V:

Proposition 2 *The value of information V is given by (17), where*

$$cov\left[\frac{q_{0,t}}{\bar{A}}, X_{t+1}\right] = -\frac{\mathcal{T}}{1-\mathcal{T}}\left(1 - \frac{\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}}\frac{\frac{\partial P_t}{\partial \epsilon_{D,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}}\right)\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right], \tag{19}$$

and

$$\frac{\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}} = \pm\sqrt{\frac{\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right]}{\left(\frac{\partial P_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right]}}. \tag{20}$$

The covariance term (19) can be estimated. The first can be bounded using Proposition 1's expression for the average investor's information $(1-\mathcal{T})(1-\Upsilon)$. The final term is the variance of returns attributable to shocks to date $t+1$ dividends. The ratio $\frac{\frac{\partial P_t}{\partial \epsilon_{D,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}}$ can be straightforwardly estimated in the data. Finally, the ratio $\frac{\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}}$ can estimated using (20), i.e., from the ratio of the variances, each of which can be measured as a residual.

Note that estimating the ratio $\frac{\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}}$ via (20) fails to determine the sign of this ratio. In the numerical implementation, I compute (19) for both possibilities.

The estimated value of the covariance (19) is negative, meaning that uninformed investors end up negatively timing the market, in the sense of increasing their asset holdings when future returns are low. To see why negative covariance arises, consider a date t shock to either discount rates or dividends that increases the return from date t to $t+1$. In response to this shock, an uninformed investor's expectations about returns rises less than informed investor's (see Lemma 2). By market clearing, the average investor's desired asset holding cannot respond (Corollary 1). It follows that an uninformed investor's asset holding must fall.

As noted, the negative covariance between uninformed asset holdings and subsequent returns is closely related to results in Watanabe (2008) and Biais et al (2010). It occurs even though the uninformed investor is fully Bayesian in forming expectations about future returns.

As a final comment on Proposition 1: Recall that the value of information V is based on an uninformed investor who responds to return expectations and discount rate shocks $Z_t + u_{it}$ in the same way as the average investor in the economy, i.e., has characteristics $\bar{A}$ and $\bar{B}$ given by (15). However, in many underpinnings of asset demand (1), the coefficient A_i on return expectations is a joint function of the endogenous perceived return variance, $var\left[X_{t+1}|\mathcal{I}_{i,t}\right]$, and exogenous risk aversion. Moreover, the same is potentially true of the coefficient B_i on discount rate shocks. As such, the measure V doesn't account for the increase in the average size of an uninformed investor's average position that may accompany getting access to better information and thereby reducing perceived return variance.

Two points are worth making here. First, incorporating this effect into the measure V would require substantially stronger assumptions about investors' asset demands than I have made so far. That is: How exactly does an investor's asset demand depend on perceived return variance? Second: Although estimates of the reduction in the perceived variance of dividend innovations $\epsilon_{D,t}$ turn out to be relatively large (Section 6), estimates of the reduction in the perceived variance of returns X_{t+1} are much smaller, because estimates indicate that most return variance is driven by date $t+1$ discount rate shocks (Section 5). As such, incorporating the reduction-in-perceived-variance effect into V is likely to have a relatively modest impact.

5 Predictability empirics

The key quantities required to operationalize Propositions 1 and 2 are tightly related to the outputs of empirical analysis that studies the predictability of returns and dividends, emphasizing the role of the price-dividend ratio.

The literature is sizeable. In this paper, I use of the estimates of Binsbergen and Koijen (2010), henceforth BK, to illustrate the methodology developed in Section 4. As the preceding analysis indicates, the interpretation of today's prices requires incorporating the history of prices and dividends in order to infer the history of discount rate shocks. An important advantage of BK's estimates for the evaluation of the amount and value of private information is that, as they write, "Our latent variables approach aggregates the whole history of price-dividend ratios and dividend growth rates to estimate expected returns and expected growth rates."

5.1 Estimated VAR

BK estimate an empirical model with exogenous shocks to dividend growth, expected return, and realized dividends. As emphasized by BK and Cochrane (2008), such an empirical model is indistinguishable from an alternative one with exogenous shocks to dividend growth and the price-dividend ratio. I will work with this latter specification because it is stated in terms of observables, and as such is closer to quantities needed as inputs for Propositions 1 and 2.

Specifically, write $r_{t+1} = \log \frac{P_{t+1}+D_{t+1}}{P_t}$ and $\Delta d_{t+1} = \log \frac{D_{t+1}}{D_t}$ for the asset return and dividend growth rate between dates t and $t+1$. Let μ_t and g_t denote the econometrician's date t expectations about returns and dividend growth:

$$\begin{aligned} \mu_t &= E\left[r_{t+1}|\mathcal{J}_t\right] \\ g_t &= E\left[\Delta d_{t+1}|\mathcal{J}_t\right]. \end{aligned}$$

BK assume that μ_t and g_t follow AR1 processes:

$$\begin{aligned} \mu_{t+1} &= \bar{\mu} + \phi_\mu\left(\mu_t - \bar{\mu}\right) + \nu_{\mu,t+1} \\ g_{t+1} &= \bar{\Delta d} + \phi_g\left(g_t - \bar{\Delta d}\right) + \nu_{g,t+1}. \end{aligned}$$

Writing $pd_t = \log \frac{P_t}{D_t}$ for the log price-dividend ratio, along with $\bar{pd}$ for its steady state value, and $\rho = \frac{\exp\left(\bar{pd}\right)}{1+\exp\left(\bar{pd}\right)}$, the now-standard present value approximation (Campbell and Shiller

1988; for completeness, see Appendix D) is

$$pd_t - \bar{pd} = \frac{g_t - \bar{\Delta d}}{1 - \rho\phi_g} - \frac{\mu_t - \bar{\mu}}{1 - \rho\phi_\mu}.$$

Let $\nu_{d,t+1}$ and $\nu_{pd,t+1}$ denote the unforecastable (to the econometrician) innovations to Δd_{t+1} and pd_{t+1}, i.e.,

$$\begin{aligned} \Delta d_{t+1} &= g_t + \nu_{d,t+1} \\ pd_{t+1} &= E\left[pd_{t+1}|\mathcal{J}_t\right] + \nu_{pd,t+1}. \end{aligned}$$

Since

$$\begin{aligned} pd_{t+1} - \bar{pd} &= \frac{\phi_g\left(g_t - \bar{\Delta d}\right)}{1 - \rho\phi_g} - \frac{\phi_\mu\left(\mu_t - \bar{\mu}\right)}{1 - \rho\phi_\mu} + \frac{\nu_{g,t+1}}{1 - \rho\phi_g} - \frac{\nu_{\mu,t+1}}{1 - \rho\phi_\mu} \\ &= \frac{\phi_g - \phi_\mu}{1 - \rho\phi_g}\left(g_t - \bar{\Delta d}\right) + \phi_\mu\left(pd_t - \bar{pd}\right) + \frac{\nu_{g,t+1}}{1 - \rho\phi_g} - \frac{\nu_{\mu,t+1}}{1 - \rho\phi_\mu}, \end{aligned}$$

it follows that

$$\nu_{pd,t+1} = \frac{\nu_{g,t+1}}{1 - \rho\phi_g} - \frac{\nu_{\mu,t+1}}{1 - \rho\phi_\mu}. \tag{21}$$

Since the econometrician observes only dividends and prices, the innovations $\nu_{g,t+1}$ and $\nu_{\mu,t+1}$ must be functions of $\nu_{pd,t+1}$ and $\nu_{d,t+1}$:[3]

$$\begin{aligned} \nu_{g,t+1} &= a_{pd}\nu_{pd,t+1} + a_d\nu_{d,t+1} \\ \nu_{\mu,t+1} &= b_{pd}\nu_{pd,t+1} + b_d\nu_{d,t+1}. \end{aligned}$$

From (21), b_{pd} and b_d satisfy

$$\begin{aligned} \frac{a_{pd}}{1 - \rho\phi_g} - \frac{b_{pd}}{1 - \rho\phi_\mu} &= 1 \\ \frac{a_d}{1 - \rho\phi_g} - \frac{b_d}{1 - \rho\phi_\mu} &= 0. \end{aligned}$$

So the estimated system is

$$g_{t+1} - \bar{\Delta d} = \phi_g\left(g_t - \bar{\Delta d}\right) + a_{pd}\nu_{pd,t+1} + a_d\nu_{d,t+1} \tag{22}$$

[3]Linearity here follows from Cochrane (2008, p. 11). Roughly: $\{\nu_g, \nu_\mu, \nu_d\}$ is assumed to be stationary. Hence $\{\nu_{pd}, \nu_d\}$ is stationary; and ν_g and ν_μ must be linear functions of ν_{pd}, ν_d.

$$\Delta d_{t+1} = g_t + \nu_{d,t+1} \tag{23}$$

$$pd_{t+1} - \bar{pd} = \frac{\phi_g - \phi_\mu}{1 - \rho\phi_g}\left(g_t - \bar{\Delta d}\right) + \phi_\mu\left(pd_t - \bar{pd}\right) + \nu_{pd,t+1}. \tag{24}$$

The system has ten parameters,

$$\left\{\bar{\Delta d}, \bar{pd}, \rho, \phi_g, \phi_\mu, a_{pd}, a_d, \sigma^2_{pd}, \sigma^2_d, \sigma_{pd,d}\right\}.$$

Appendix B details how to recover estimates of these ten values from the estimates reported in BK. The first five parameters $\left\{\bar{\Delta d}, \bar{pd}, \rho, \phi_g, \phi_\mu\right\}$ coincide with the BK values. Using the estimates reported in the first column of BK's Table II yields the values reported in Table 1.

Parameter	Estimated value
$\bar{\Delta d}$	0.062
$\bar{pd}$	3.571
ρ	0.969
ϕ_g	0.354
ϕ_μ	0.932
a_{pd}	0.0482
a_d	0.3952
σ_{pd}	0.1596
σ_d	0.0576
$\frac{\sigma_{pd,d}}{\sigma_{pd}\sigma_d}$	−0.3118
$\bar{\mu}$	0.090
b_{pd}	−0.0898
R^2 of Δd_{t+1} regressed on $\mathcal{J}_t$	13.9%

Table 1: Parameter estimates recovered by BK

5.2 From predictability estimates to required inputs

The evaluation of Propositions 1 and 2 requires the estimates of the quantities listed in Table 2, which also summarizes the corresponding economic quantities. Recall that $\epsilon_{D,t}$ is the innovation to date $t+1$ dividends, though investors observe noisy signals at date t.

In particular, $\frac{\partial E[X_{t+1}|\mathcal{J}_t]}{\partial P_t}$ is closely related to the estimated parameter b_{pd}, i.e., the relation between the expected return ($\nu_{\mu,t}$) and the price-dividend ratio ($\nu_{pd,t}$), holding dividends fixed ($\nu_{d,t} = 0$). Similarly, $\frac{\partial P_t}{\partial \epsilon_{D,t}}$ is closely related to a_{pd}, i.e., the relation between expectations about the dividend growth rate ($\nu_{g,t}$) and today's price ($\nu_{pd,t}$), holding current dividends fixed ($\nu_{d,t} = 0$).

Term	Description	Estimated value (see below)
$\frac{\partial E[X_{t+1}\|\mathcal{J}_t]}{\partial P_t}$	Expected return and price	-0.0463
$\frac{\partial P_t}{\partial \epsilon_{D,t}}$	Price and next-period dividend	10.961
$\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}$	Realized return and contemporaneous dividend	4.5182
$\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right]$	Variance of returns due to $\epsilon_{D,t}$	$P_{t-1}^2 \times 0.0120^2$
$\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right]$	Variance of returns due to $\epsilon_{Z,t}$	$P_{t-1}^2 \times 0.0119^2$
$\left(\frac{\partial P_t}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right]$	Variance of price due to $\epsilon_{Z,t}$	$P_{t-1}^2 \times 0.16^2$
$E\left[X_{t+1}\right]$	Equity premium	$P_{t-1} \times 0.0789$

Table 2: Quantities to estimate

BK's estimates are based on nominal annual returns and nominal annual dividend growth rates. Accordingly, the appropriate risk free rate R_t is a nominal annual rate. I use a risk free rate of 2% in the calculations below.

5.3 The term $\frac{\partial E[X_{t+1}|\mathcal{J}_t]}{\partial P_t}$

The estimated coefficient is

$$b_{pd} = \frac{\partial E\left[\ln\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right)|\mathcal{J}_t\right]}{\partial \ln P_t} = -0.0898.$$

In words: a 100% increase in prices is associated with a decline in the econometrician's expected return of 9*pp* (percentage points).

Manipulation yields (both here and below see appendix for details)

$$\frac{\partial}{\partial P_t} E\left[X_{t+1}|\mathcal{J}_t\right] \approx \frac{E\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]+R}{1+\frac{var\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]}{\left(E\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]+R\right)^2}}\left(1+b_{pd}\right)-R. \tag{25}$$

To evaluate (25), replace $E\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]$ and $var\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]$ with their steady state values. To do so, note that the steady state value of $E\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]+R$ is approximately the steady state (gross) expected return, i.e., $1+\bar{\mu}$. The steady state value of $var\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]$ is simply the square of return volatility, i.e., approx 0.15^2.

Hence

$$\frac{\partial}{\partial P_t} E\left[X_{t+1}|\mathcal{J}_t\right] \approx \frac{1.09}{1+\left(\frac{0.15}{1.09}\right)^2}\left(1-0.0898\right)-R=-0.0463. \tag{26}$$

That is: an increase in today's price of \$1 reduces the (\$) expected return by approximately \$0.05.

5.4 The term $\frac{\partial P_t}{\partial \epsilon_{D,t}}$

The estimated coefficient is

$$a_{pd}=\frac{\partial E\left[\ln D_{t+1}|\mathcal{J}_t\right]}{\partial \ln P_t}=0.0482.$$

In words: a 100% increase in prices is associated with an increase in the econometrician's expected dividend growth rate of 5*pp*.

I first relate a_{pd} to $\left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t, \mathcal{J}_{t-1}}$, and then subsequently relate $\left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t, \mathcal{J}_{t-1}}$ to the desired term $\frac{\partial P_t}{\partial \epsilon_{D,t}}$. The first step corresponds to switching from "does today's price predict future dividends" to "do future dividends predict today's price"? The second step is a simply a shift from percentage changes to level changes. The first step is related to Dávila and Parlatore (2021), who argue that the residual variance of current prices after controlling for future dividends is the correct measure of price informativeness, as opposed to the more commonly measured residual variance of future dividends after controlling for the current price.

For the first step:

$$\left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t, \mathcal{J}_{t-1}}=\frac{cov\left[pd_t, \Delta d_{t+1}|\Delta d_t, \mathcal{J}_{t-1}\right]}{var\left[\Delta d_{t+1}|\Delta d_t, \mathcal{J}_{t-1}\right]}.$$

Evaluating,

$$\begin{aligned} cov\left[pd_t, \Delta d_{t+1} | \Delta d_t, \mathcal{J}_{t-1}\right] &= cov\left[pd_t, g_t | \Delta d_t, \mathcal{J}_{t-1}\right] \\ &= cov\left[\nu_{pd,t}, a_{pd}\nu_{pd,t} | \nu_{d,t}\right] \\ &= a_{pd} var\left[\nu_{pd,t} | \nu_{d,t}\right] \end{aligned} \tag{27}$$

and

$$\begin{aligned} var\left[\Delta d_{t+1} | \Delta d_t, \mathcal{J}_{t-1}\right] &= var\left[\nu_{d,t+1}\right] + var\left[g_t | \Delta d_t, \mathcal{J}_{t-1}\right] \\ &= var\left[\nu_{d,t+1}\right] + var\left[a_{pd}\nu_{pd,t} | \nu_{d,t}\right] \\ &= var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\nu_{pd,t} | \nu_{d,t}\right], \end{aligned} \tag{28}$$

and hence

$$\left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t, \mathcal{J}_{t-1}} = \frac{1}{a_{pd}} \frac{a_{pd}^2 var\left[\nu_{pd,t} | \nu_{d,t}\right]}{var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\nu_{pd,t} | \nu_{d,t}\right]}. \tag{29}$$

The second term in (29) is the fraction of the residual variance of Δd_{t+1} after controlling for Δd_t and $\mathcal{J}_{t-1}$ that is explained by further controlling for pd_t. If this ratio is 1, i.e., pd_t and Δd_{t+1} are perfectly correlated conditional on Δd_t and $\mathcal{J}_{t-1}$, then $\left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t, \mathcal{J}_{t-1}}$ is simply the reciprocal of a_{pd}.

Evaluating

$$\begin{aligned} var\left[\nu_{pd,t} | \nu_{d,t}\right] &= var\left[\nu_{pd,t}\right] - \left(\frac{cov\left[\nu_{pd,t}, \nu_{d,t}\right]}{var\left[\nu_{d,t}\right]}\right)^2 var\left[\nu_{d,t}\right] \\ &= var\left[\nu_{pd,t}\right]\left(1 - corr\left[\nu_{pd,t}, \nu_{d,t}\right]^2\right) = 0.0230 \end{aligned} \tag{30}$$

and so

$$\frac{a_{pd}^2 var\left[\nu_{pd,t} | \nu_{d,t}\right]}{var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\nu_{pd,t} | \nu_{d,t}\right]} = 0.0158. \tag{31}$$

The low value of this last term reflects the standard result that today's price contains very limited predictive power for tomorrow's dividend. Returning to (29),

$$\left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t, \mathcal{J}_{t-1}} = \frac{0.0158}{0.0482} = 0.328,$$

i.e., a 100% increase in date $t+1$ dividends suggests that date t prices were 33% higher.

For the second step, along with numerical implementation:

$$\frac{\partial P_t}{\partial \epsilon_{D,t}} \approx \frac{\exp\left(\bar{pd}\right)}{\exp\left(\bar{g}\right)}\left(1 - cov\left[pd_t, \Delta d_{t+1}\right]\right) \left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t, \mathcal{J}_{t-1}} \tag{32}$$

$$= e^{3.571-0.062} \times 1.0001 \times 0.328 = 10.961. \tag{33}$$

5.5 The term $\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}$

A first step is to decompose the effect of the date $t+1$ dividend innovation on the excess return X_{t+1} into its anticipated and unanticipated components:

$$\begin{aligned}
\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}} &= \frac{\partial E\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{\partial \epsilon_{D,t}} \frac{\partial X_{t+1}}{\partial E\left[\epsilon_{D,t}|\mathcal{J}_t\right]} + \frac{\partial\left(\epsilon_{D,t} - E\left[\epsilon_{D,t}|\mathcal{J}_t\right]\right)}{\partial \epsilon_{D,t}} \frac{\partial X_{t+1}}{\partial\left(\epsilon_{D,t} - E\left[\epsilon_{D,t}|\mathcal{J}_t\right]\right)} \\
&= \frac{\partial E\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{\partial \epsilon_{D,t}} \frac{\partial E\left[X_{t+1}|\mathcal{J}_t\right]}{\partial E\left[\epsilon_{D,t}|\mathcal{J}_t\right]} + \frac{\partial\left(\epsilon_{D,t} - E\left[\epsilon_{D,t}|\mathcal{J}_t\right]\right)}{\partial \epsilon_{D,t}} \frac{\partial\left(P_{t+1} + D_{t+1}\right)}{\partial\left(\epsilon_{D,t} - E\left[\epsilon_{D,t}|\mathcal{J}_t\right]\right)} \\
&= \frac{\partial P_t}{\partial \epsilon_{D,t}} \frac{\partial E\left[X_{t+1}|\mathcal{J}_t\right]}{\partial P_t} + \frac{\partial\left(\epsilon_{D,t} - E\left[\epsilon_{D,t}|\mathcal{J}_t\right]\right)}{\partial \epsilon_{D,t}} \left.\frac{\partial\left(P_{t+1} + D_{t+1}\right)}{\partial D_{t+1}}\right|_{\mathcal{J}_t},
\end{aligned} \tag{34}$$

where the second inequality follows from the fact that the innovation $\epsilon_{D,t}$ affects the econometrician's date t expectation only via the price P_t.

Both elements of the first term of (34), corresponding to effect of the anticipated component of $\epsilon_{D,t}$, are calculated above.

In the second term, corresponding to the effect of the unanticipated component, $\frac{\partial\left(\epsilon_{D,t} - E\left[\epsilon_{D,t}|\mathcal{J}_t\right]\right)}{\partial \epsilon_{D,t}}$ is $1 - R^2\left(\Delta d_{t+1} \text{ on } \mathcal{J}_t\right)$. This is reported in BK. Finally,

$$\left.\frac{\partial\left(P_{t+1} + D_{t+1}\right)}{\partial D_{t+1}}\right|_{\mathcal{J}_t} \approx 1 + e^{\bar{pd}}\left(\frac{cov\left[\nu_{pd,t+1}, \nu_{d,t+1}\right]}{var\left[\nu_{d,t+1}\right]} + 1\right) \tag{35}$$

$$= 1 + e^{3.571}\left(-0.3118 \times \frac{0.1596}{0.0576} + 1\right) = 5.8370. \tag{36}$$

So from (34) (and making use of (26) and (33))

$$\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}} \approx 10.961 \times (-0.0463) + (1 - 0.139) \times 5.8370 = 4.5182.$$

Perhaps as one would expect, most of the effect of the dividend's innovation $\epsilon_{D,t}$ stems from the unanticipated component. In turn, the unanticipated component both directly increases the date $t+1$ dividend D_{t+1}, and also increases the date $t+1$ price P_{t+1}.

5.6 The term $\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right]$

I make use of the decomposition (34). By the law of total variance,[4]

$$var\left[\frac{\partial P_t}{\partial \epsilon_{D,t}}\epsilon_{D,t}\right] \approx var\left[P_t | D_t, \mathcal{J}_{t-1}\right] - var\left[P_t | D_{t+1}, D_t, \mathcal{J}_{t-1}\right]. \quad (37)$$

That is, the variance of P_t given $D_t, \mathcal{J}_{t-1}$ stems from $\epsilon_{D,t}$ and $\epsilon_{Z,t}$. The term $var\left[P_t|D_{t+1}, D_t, \mathcal{J}_t\right]$ isolates the effect stemming from $\epsilon_{Z,t}$. Evaluating,

$$var\left[\frac{\partial P_t}{\partial \epsilon_{D,t}}\epsilon_{D,t}\right] \approx P_{t-1}^2 e^{2\bar{\Delta d}} \frac{a_{pd}^2 var\left[\upsilon_{pd,t}|\nu_{d,t}\right]^2}{var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\upsilon_{pd,t}|\nu_{d,t}\right]}. \quad (38)$$

Next, note that

$$var\left[\frac{\partial\left(\epsilon_{D,t} - E\left[\epsilon_{D,t}|\mathcal{J}_t\right]\right)}{\partial \epsilon_{D,t}}\epsilon_{D,t}\right] = var\left[\epsilon_{D,t}|\mathcal{J}_t\right],$$

i.e., simply the variance of the component of $\epsilon_{D,t}$ that is unanticipated by the econometrician given date t information. Evaluating,

$$var\left[\epsilon_{D,t}|\mathcal{J}_t\right] = D_t^2 var\left[e^{\Delta d_{t+1}}|\mathcal{J}_t\right] \approx P_{t-1}^2 e^{2\Delta d_t - 2pd_{t-1}}\left(e^{\bar{\Delta d}}\right)^2 var\left[\nu_{d,t+1}\right].$$

Putting everything together,

$$\begin{aligned}\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right] &\approx P_{t-1}^2 e^{2\bar{\Delta d}} \\ &\times \left(\left(\frac{\partial E\left[X_{t+1}|\mathcal{J}_t\right]}{\partial P_t}\right)^2 \frac{a_{pd}^2 var\left[\upsilon_{pd,t}|\nu_{d,t}\right]^2}{var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\upsilon_{pd,t}|\nu_{d,t}\right]}\right. \\ &+ \left.\left(\left.\frac{\partial\left(P_{t+1} + D_{t+1}\right)}{\partial D_{t+1}}\right|_{\mathcal{J}_t}\right)^2 e^{2\left(\bar{\Delta d} - \bar{pd}\right)} var\left[\nu_{d,t+1}\right]\right).\end{aligned}$$

The two terms on the RHS correspond to return variation stemming from anticipated and unanticipated dividend innovations.

From (26), (30), (31),

$$e^{2\bar{\Delta d}}\left(\frac{\partial E\left[X_{t+1}|\mathcal{J}_t\right]}{\partial P_t}\right)^2 \frac{a_{pd}^2 var\left[\upsilon_{pd,t}|\nu_{d,t}\right]^2}{var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\upsilon_{pd,t}|\nu_{d,t}\right]} = e^{2\times.062}\times.0463^2\times.0230\times.0158 = 8.82\times 10^{-7}.$$

[4]Equation (37) holds exactly under joint normality.

From (36),

$$e^{2\bar{\Delta d}}\left(\left.\frac{\partial\left(P_{t+1}+D_{t+1}\right)}{\partial D_{t+1}}\right|_{\mathcal{J}_t}\right)^2 e^{2\bar{\Delta d}-2\bar{pd}} var\left[\nu_{d,t+1}\right] = e^{2\times .062}\times 5.8370^2\times e^{2\times(.062-3.571)}\times 0.0576^2 = 1.45\times 10^{-4}.$$

Anticipated date $t+1$ dividend innovations contribute a negligible amount of return variance relative to unanticipated date $t+1$ dividend innovations.

Hence

$$\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right] \approx P_{t-1}^2 \times 1.45\times 10^{-4} = P_{t-1}^2 \times 0.0120^2.$$

5.7 The term $\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right]$

Note that

$$\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right] = var\left[E\left[X_{t+1}|D_{t+1},\mathcal{J}_t\right]|D_{t+1},D_t,\mathcal{J}_{t-1}\right].$$

Evaluating,

$$\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right] \approx P_{t-1}^2 e^{2\bar{\Delta d}}\left(e^{\bar{\Delta d}}\left(\frac{\phi_g-\phi_\mu}{1-\rho\phi_g}a_{pd}+\phi_\mu\right)-R\right)^2 var\left[\nu_t^{pd}|\nu_t^d\right]. \tag{39}$$

From (30),

$$\begin{aligned}\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right] &\approx P_{t-1}^2 e^{2\times .062}\times\left(e^{.062}\left(\frac{.354-.932}{1-.969\times .354}\times .0482+.932\right)-1.02\right)^2\times 0.0230\\ &= P_{t-1}^2\times 1.41\times 10^{-4} = P_{t-1}^2\times .0119^2.\end{aligned}$$

Date t discount rate innovations make approximately the same contribution to the variance of the return X_{t+1} as do unanticipated date $t+1$ dividend shocks.

Remark (Aside): Combining the variance to returns X_{t+1} stemming from dividend shocks $\epsilon_{D,t}$ and discount rate shocks $\epsilon_{Z,t}$ gives

$$\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right] + \left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right] \approx P_{t-1}^2\times 2.86\times 10^{-4} = P_{t-1}^2\times .0169^2. \tag{40}$$

Expressed in percentage terms, date $t+1$ innovations and date t discount rate innovations generate a standard deviation of returns of approximately 1.7%. Why is the this value so much lower than the total standard deviation of returns, which is on the order of 15% − 20%? The reason is that (40) omits the return variation stemming from date $t+1$ discount rate innovations, i.e., $\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t+1}}\right)^2 var\left[\epsilon_{Z,t+1}\right]$. Using the decomposition

$P_{t+1} = P_{t-1} e^{\Delta d_{t+1} + \Delta d_t - pd_{t-1}} e^{pd_{t+1}},$

$$\left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t+1}}\right)^2 var\left[\epsilon_{Z,t+1}\right] \approx P_{t-1}^2 e^{4\bar{\Delta d}} var\left[\nu_{t+1}^{pd} | \nu_{t+1}^{d}\right] = P_{t-1}^2 \times 0.1717^2.$$

In other words, most of the variation in returns from date t to $t+1$ stems from date $t+1$ discount rate innovations; and the estimated value of this variation is consistent with total return variation lying in the $15 - 20\%$ range.

In contrast: Because discount rates are highly persistent (the estimated value of the autoregression coefficient ϕ_μ is 0.93), date t innovations affect prices at both dates t and $t+1$, with modest effects on the return from date t to $t+1$.

5.8 The term $\left(\frac{\partial P_t}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right]$

By the law of total variance,

$$\begin{aligned}
\left(\frac{\partial P_t}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right] &= D_t^2 var\left[e^{pd_t} | \Delta d_{t+1}, \Delta d_t, \mathcal{J}_{t-1}\right] \\
&\approx D_t^2 var\left[e^{pd_t} | \Delta d_t, \mathcal{J}_{t-1}\right] \\
&- D_t^2 var\left[E\left[e^{pd_t} | \Delta d_{t+1}, \Delta d_t, \mathcal{J}_{t-1}\right] | \Delta d_t, \mathcal{J}_{t-1}\right].
\end{aligned}$$

Evaluating, making use of (28) and (29), and then (30) in the numerical evaluation:

$$\begin{aligned}
\left(\frac{\partial P_t}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right] &\approx P_{t-1}^2 e^{2\bar{\Delta d}} \frac{var\left[\nu_{d,t+1}\right] var\left[\nu_{pd,t} | \nu_{d,t}\right]}{var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\nu_{pd,t} | \nu_{d,t}\right]}. \qquad (41) \\
&= P_{t-1}^2 e^{2 \times 0.062} \frac{.0576^2 \times .0230}{.0576^2 + .0482^2 \times .0230} = P_{t-1}^2 \times .0256 = P_{t-1}^2 \times .16^2.
\end{aligned}$$

5.9 The term $E\left[X_{t+1}\right]$

Finally, I consider the term $E\left[X_{t+1}\right]$. This is simply the equity premium,. While many estimates are available, for consistency I use one based on the same BK estimates as the other terms:

$$\begin{aligned}
E\left[X_{t+1}\right] &\approx P_{t-1} e^{\bar{\Delta d}} \left(e^{\bar{\mu}} - R\right) \qquad (42) \\
&= P_{t-1} e^{.062} \left(e^{.090} - 1.02\right) = P_{t-1} \times 0.0789. \qquad (43)
\end{aligned}$$

6 The quantity and value of the average investor's private information

Finally, I use the estimated values to operationalize Propositions 1 and 2, along with the decomposition (14).

6.1 The average investor's information

The average investor's information is, by Proposition 1,

$$(1-\mathcal{T})(1-\Upsilon) \approx \frac{4.5182}{4.5182-(-.0463 \times 10.961)} = 0.899. \tag{44}$$

That is: The average investor's information reduces the conditional variance of the dividend forecast by just over 10%.

6.2 Information in the price relative to the average investor's information

To operationalize the decomposition of informational sources (14), note that

$$var\left[\epsilon_{D,t}|\mathcal{J}_t\right] = var\left[D_t e^{\Delta d_{t+1}}|\mathcal{J}_t\right] \approx \left(D_t e^{\bar{\Delta d}}\right)^2 var\left[\Delta d_{t+1}|\mathcal{J}_t\right] = \left(D_t e^{\bar{\Delta d}}\right)^2 var\left[\nu_{d,t+1}\right],$$

and similarly, and using (28),

$$var\left[\epsilon_{D,t}|\Delta d_t, \mathcal{J}_{t-1}\right] = \left(D_t e^{\bar{\Delta d}}\right)^2 \left(var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\nu_{pd,t}|\nu_{d,t}\right]\right).$$

Hence

$$\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}|D_t, \mathcal{J}_{t-1}\right]} = \frac{var\left[\nu_{d,t+1}\right]}{var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\nu_{pd,t}|\nu_{d,t}\right]}.$$

Hence, by (31),

$$\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}|D_t, \mathcal{J}_{t-1}\right]} \approx 1 - 0.0158 = 0.9842.$$

In other words, the average investor learns much more from his/her own information than from the price.

How can the average investor have substantial information, and yet the price have little information? Loosely speaking, the aggregate discount rate shock dominates price fluctuations, reducing the information content of the price for dividend innovations.

6.3 The value of private information

The value of the average investor's information is given by Proposition 2. A key input in Proposition 2 is $\mathcal{T}$, relating to the precision of the average investor's signal about dividends. Proposition 1 does not separably identify $\mathcal{T}$ and Υ. However, the value of the average investor's information is bounded by the case of the largest possible value of $\mathcal{T}$, coupled with the lowest possible value of Υ, i.e., $\mathcal{T} = 0.101$ and $\Upsilon = 0$. Evaluating at these values:

$$\begin{aligned} cov\left[\frac{q_{0,t}}{\bar{A}}, X_{t+1}\right] &\approx -\frac{.101}{1-.101}\left(1 \pm \frac{0.0119}{0.16}\frac{10.961}{4.5182}\right) \times P_{t-1}^2 \times 0.0120^2 \\ &= \left\{-P_{t-1}^2 \times 1.91 \times 10^{-5}, -P_{t-1}^2 \times 1.33 \times 10^{-5}\right\}, \end{aligned} \quad (45)$$

and so

$$V \approx \left\{\frac{1}{1-\frac{1.91\times 10^{-5}}{.0789^2}}, \frac{1}{1-\frac{1.33\times 10^{-5}}{.0789^2}}\right\} = \{1.003, 1.002\}.$$

Why is the value of private information so small? While the standard deviation of returns (expressed as a percentage) is in the 15% range, the standard deviation of expected returns is an order of magnitude lower. In particular, the estimates of Section 5 suggest that the standard deviation of expected returns is on the order of 1%, stemming overwhelming stemming from date t discount rate shocks. So if the average investor's information reduces perceived variance by 10%, this corresponds to a reduction of $.1 \times .01^2 = 10^{-5}$. This back-of-the-envelope calculation is consistent with the output of the calculation in (45).

6.4 Bounding the ratio of idiosyncratic to aggregate discount rate shocks

Proposition 1 can also be used to bound the relative importance of idiosyncratic to aggregate discount rate shocks.

From (44), an uninformed investor's informational advantage Υ relative to the econometrician is bounded above by $1 - .899 = 0.101$. From (6), it follows that the ratio of the variances of idiosyncratic to aggregate discount rate shocks satisfies

$$\frac{var\left[u_{i,t}\right]}{var\left[\epsilon_{Z,t}\right]} \geq \frac{.899}{1-.899}\left(1 - \frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}\right]}\right).$$

The ratio $\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]}{var\left[\epsilon_{D,t}\right]}$ is the R^2 of Δd_{t+1} on the econometrician's information set. BK report a value of 13.9%. Hence, and expressed in more familiar standard deviation terms,

$$\sqrt{\frac{var\left[u_{i,t}\right]}{var\left[\epsilon_{Z,t}\right]}} \geq 2.77.$$

That is, the standard deviation of idiosyncratic discount rate shocks must be at least 2.77 times that of aggregate discount rate shocks. If this were not true, the information that individual investors would be able to extract from public prices would exceed the total quantity of their information estimated in (44).

7 Concluding remarks

In this paper I derive a methodology for estimating the quantity and value of the average investor's private information. The methodology is based on relatively general assumptions about asset demand, coupled with a focus on the first two moments of all relevant distributions. The methodology makes use of market clearing conditions to yield sufficient statistics formulae for the quantity and value of private information.

I operationalize the methodology using estimates from the price-dividend predictability literature, specifically, those of Binsbergen and Koijen (2010). Using BK's estimates suggests that the average investor possesses substantial private information about future dividends, but little of this information is impounded into prices, and the value of this information is relatively small. The combination of these results reflects the large role that discount rate innovations play in price fluctuations.

BK's estimates reflect, naturally, a number of specification decisions, including the time horizon (one year) and the set of public conditioning variables (prices and dividends). The sufficient statistics formulae hold independently of these estimation choices, though of course the estimates of the quantity and value of private information are sensitive to these choices.

References

Biais, Bossaerts, and Spatt 2010
Bai, Savov, Philippon 2015
Binsbergen and Koijen 2010
Bond and García 2021
Campbell and Shiller 1988
Cochrane 2008

Cochrane 2011
Dávila and Parlatore 2021
Diamond and Verrecchia 1981
Easley, Kiefer, and O'Hara 1996
Egan, MacKay, and Yang 2021
Glode 2011
Grossman 1976
Grossman and Stiglitiz 1980
Hayek 1945
Hellwig 1980
Kurlat 2019
Kadan and Manela 2019
Koijen and Van Nieuwerburgh 2011
Kyle 1985
Savov 2014
Sharpe 1991
Watanabe 2008

A Appendix: Proofs

Proof of Lemma 1: By the standard formula for the conditional variance given joint normality,

$$
\begin{aligned}
var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] &= var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{0,i,t}\right] \\
&- \frac{\begin{pmatrix}var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]\\ cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{0,i,t}\right]\end{pmatrix}\begin{pmatrix}var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right] & cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{0,i,t}\right]\end{pmatrix}}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]+\tau_i^{-1}} \qquad (46) \\
&= \frac{\tau_i^{-1}}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]+\tau_i^{-1}} var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{0,i,t}\right] \\
&+ \frac{1}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]+\tau_i^{-1}}\begin{pmatrix}0 & 0\\ 0 & var\left[\epsilon_{Z,t}|\mathcal{I}_{0,i,t}\right]var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]-cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{0,i,t}\right]^2\end{pmatrix} \\
&= \frac{\tau_i^{-1}}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]+\tau_i^{-1}} var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{0,i,t}\right], \qquad (47)
\end{aligned}
$$

where the final equality follows from (3). In particular, (47) implies

$$\frac{\tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]} = \frac{1}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]+\tau_i^{-1}},$$

and hence

$$1-\tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right] = 1-\frac{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]+\tau_i^{-1}} = \frac{\tau_i^{-1}}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]+\tau_i^{-1}} = \frac{var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]}{var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]}. \quad (48)$$

Substituting (48) into (47) yields (4).

Similarly,

$$\begin{aligned}
var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{0,i,t}\right] &= var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{J}_t\right] - \frac{\begin{pmatrix}cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{J}_t\right]\\ var\left[\epsilon_{Z,t}|\mathcal{J}_t\right]\end{pmatrix}\begin{pmatrix}cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{J}_t\right] & var\left[\epsilon_{Z,t}|\mathcal{J}_t\right]\end{pmatrix}}{var\left[\epsilon_{Z,t}|\mathcal{J}_t\right]+var\left[u_{i,t}\right]}\\
&= \frac{var\left[u_{i,t}\right]}{var\left[\epsilon_{Z,t}|\mathcal{J}_t\right]+var\left[u_{i,t}\right]} var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{J}_t\right]\\
&- \begin{pmatrix}\frac{var\left[\epsilon_{D,t}|\mathcal{J}_t\right]var\left[\epsilon_{Z,t}|\mathcal{J}_t\right]-cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{J}_t\right]^2}{var\left[\epsilon_{Z,t}|\mathcal{J}_t\right]+var\left[u_{i,t}\right]} & 0\\ 0 & 0\end{pmatrix},
\end{aligned}$$

which by (3) delivers (5) and completes the proof.

Proof of Lemma 2: Let W denote the row vector of variables in $\mathcal{I}_{i,t}$. Since all random variables are normally distributed,

$$\begin{aligned}
\frac{\partial}{\partial \epsilon_{D,t}} E\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right] &= cov\left[\epsilon_{D,t},W\right]var\left[W\right]^{-1}\left(\frac{\partial W}{\partial \epsilon_{D,t}}\right)^{\intercal} = \frac{1}{var\left[\epsilon_{D,t}\right]}cov\left[\epsilon_{D,t},W\right]var\left[W\right]^{-1}cov\left[\epsilon_{D,t},W\right]^{\intercal}\\
\frac{\partial}{\partial \epsilon_{Z,t}} E\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right] &= cov\left[\epsilon_{D,t},W\right]var\left[W\right]^{-1}\left(\frac{\partial W}{\partial \epsilon_{Z,t}}\right)^{\intercal} = \frac{1}{var\left[\epsilon_{Z,t}\right]}cov\left[\epsilon_{D,t},W\right]var\left[W\right]^{-1}cov\left[\epsilon_{Z,t},W\right]^{\intercal}
\end{aligned}$$

Moreover,

$$var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] = var\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right] - cov\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix},W\right]var\left[W\right]^{-1}cov\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix},W\right]^{\intercal},$$

and so

$$var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right] = var\left[\epsilon_{D,t}\right] - cov\left[\epsilon_{D,t},W\right]var\left[W\right]^{-1}cov\left[\epsilon_{D,t},W\right]^{\intercal}$$

$$cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right] = cov\left[\epsilon_{D,t},\epsilon_{Z,t}\right] - cov\left[\epsilon_{D,t},W\right] var\left[W\right]^{-1} cov\left[\epsilon_{Z,t},W\right]^{\intercal}.$$

Hence

$$\begin{aligned}
\frac{\partial}{\partial\epsilon_{D,t}}E\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right] &= \frac{1}{var\left[\epsilon_{D,t}\right]}\left(var\left[\epsilon_{D,t}\right] - var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right)\\
\frac{\partial}{\partial\epsilon_{Z,t}}E\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right] &= \frac{1}{var\left[\epsilon_{Z,t}\right]}\left(cov\left[\epsilon_{D,t},\epsilon_{Z,t}\right] - cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]\right),
\end{aligned}$$

and similarly,

$$\begin{aligned}
\frac{\partial}{\partial\epsilon_{D,t}}E\left[\epsilon_{Z,t}|\mathcal{I}_{i,t}\right] &= \frac{1}{var\left[\epsilon_{D,t}\right]}\left(cov\left[\epsilon_{D,t},\epsilon_{Z,t}\right] - cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]\right)\\
\frac{\partial}{\partial\epsilon_{Z,t}}E\left[\epsilon_{Z,t}|\mathcal{I}_{i,t}\right] &= \frac{1}{var\left[\epsilon_{Z,t}\right]}\left(var\left[\epsilon_{Z,t}\right] - var\left[\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]\right).
\end{aligned}$$

Since $cov\left[\epsilon_{D,t},\epsilon_{Z,t}\right] = 0$,

$$\frac{\partial}{\partial\epsilon_{D,t}}E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] - \begin{pmatrix}1\\0\end{pmatrix} = -\frac{1}{var\left[\epsilon_{D,t}\right]}\begin{pmatrix}var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\\ cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]\end{pmatrix} \tag{49}$$

$$\frac{\partial}{\partial\epsilon_{Z,t}}E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] - \begin{pmatrix}0\\1\end{pmatrix} = -\frac{1}{var\left[\epsilon_{Z,t}\right]}\begin{pmatrix}cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]\\ var\left[\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]\end{pmatrix}. \tag{50}$$

By Lemma 1, these expressions rewrite as

$$\begin{aligned}
\frac{\partial}{\partial\epsilon_{D,t}}\left(E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] - \begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right) &= -\frac{1}{var\left[\epsilon_{D,t}\right]}\begin{pmatrix}var\left[\epsilon_{D,t}|\mathcal{I}_{0,i,t}\right]\\ cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{0,i,t}\right]\end{pmatrix}\left(1-\tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right)\\
\frac{\partial}{\partial\epsilon_{Z,t}}\left(E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] - \begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right) &= -\frac{1}{var\left[\epsilon_{Z,t}\right]}\begin{pmatrix}cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{0,i,t}\right]\\ var\left[\epsilon_{Z,t}|\mathcal{I}_{0,i,t}\right]\end{pmatrix}\left(1-\tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right),
\end{aligned}$$

which in turn imply the first two expressions in Lemma 2.

Similarly, the analogues of (49) and (50) for the information sets $\mathcal{I}_{0,i,t}$ and $\mathcal{J}_t$ combine with Lemma 1 to deliver the third and fourth expressions in Lemma 2, completing the proof.

Proof of Lemma 3: By conditional normality,

$$\begin{aligned}
var\left[\epsilon_{D,t}|\mathcal{J}_t\right] &= var\left[\epsilon_{D,t}\right] - \frac{c_D^2 var\left[\epsilon_{D,t}\right]^2}{var\left[c_D\epsilon_{D,t}+c_Z\epsilon_{Z,t}\right]} = \frac{c_Z^2 var\left[\epsilon_{Z,t}\right] var\left[\epsilon_{D,t}\right]}{c_D^2 var\left[\epsilon_{D,t}\right] + c_Z^2 var\left[\epsilon_{Z,t}\right]}\\
var\left[\epsilon_{Z,t}|\mathcal{J}_t\right] &= var\left[\epsilon_{Z,t}\right] - \frac{c_Z^2 var\left[\epsilon_{Z,t}\right]^2}{var\left[c_D\epsilon_{D,t}+c_Z\epsilon_{Z,t}\right]} = \frac{c_D^2 var\left[\epsilon_{D,t}\right] var\left[\epsilon_{Z,t}\right]}{c_D^2 var\left[\epsilon_{D,t}\right] + c_Z^2 var\left[\epsilon_{Z,t}\right]},
\end{aligned}$$

delivering the result.

Proof of Lemma 4: Using (9) and (10), the market-clearing conditions (7) and (8) rewrite as

$$\begin{aligned}
\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}} - \frac{\int A_i \left(1 - \tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right)}{\int A_i} \frac{\partial}{\partial \epsilon_{D,t}} \left(X_{t+1} - E\left[X_{t+1}|\mathcal{I}_{0,i,t}\right]\right) &= 0 \\
\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}} - \frac{\int B_i}{\int A_i} - \frac{\int A_i \left(1 - \tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right)}{\int A_i} \frac{\partial}{\partial \epsilon_{Z,t}} \left(X_{t+1} - E\left[X_{t+1}|\mathcal{I}_{0,i,t}\right]\right) &= 0.
\end{aligned}$$

The result then follows immediately.

Proof of Proposition 2: Substitution of Corollary 1 into the definition (16) of V gives

$$V = \frac{E\left[q_{\bar{\tau},t}\right] E\left[X_{t+1}\right]}{E\left[q_{0,t}\right] E\left[X_{t+1}\right] + cov\left[q_{0,t}, X_{t+1}\right]}.$$

Substituting in asset demand,

$$V = \frac{\bar{A} E\left[X_{t+1}\right]^2}{\bar{A} E\left[X_{t+1}\right]^2 + cov\left[q_{0,t}, X_{t+1}\right]} = \frac{1}{1 + \frac{cov\left[\frac{q_{0,t}}{\bar{A}}, X_{t+1}\right]}{E\left[X_{t+1}\right]^2}},$$

establising (17).

From the market-clearing condition (8) ,

$$\int A_i \frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}} - \int B_i - \int A_i \left(1 - \tau_i var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\right) \frac{\partial}{\partial \epsilon_{Z,t}} \left(X_{t+1} - E\left[X_{t+1}|\mathcal{I}_{0,i,t}\right]\right) = 0.$$

Dividing by $\bar{A} = \int A_i di$ and substituting in for $\mathcal{T}$ and $\bar{B}$ yields

$$\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}} - \frac{\bar{B}}{\bar{A}} - (1 - \mathcal{T}) \frac{\partial}{\partial \epsilon_{Z,t}} \left(X_{t+1} - E\left[X_{t+1}|\mathcal{I}_{0,i,t}\right]\right) = 0,$$

and hence

$$\frac{1}{\bar{A}} \frac{\partial q_{0,t}}{\partial \epsilon_{Z,t}} + \mathcal{T} \frac{\partial}{\partial \epsilon_{Z,t}} \left(X_{t+1} - E\left[X_{t+1}|\mathcal{I}_{0,i,t}\right]\right) = 0. \tag{51}$$

Further below, I establish that forecast error sensitivities are related according to

$$\frac{\frac{\partial}{\partial \epsilon_{Z,t}} \left(X_t - E\left[X_t|\mathcal{I}_{i,t}\right]\right)}{\frac{\partial}{\partial \epsilon_{D,t}} \left(X_t - E\left[X_t|\mathcal{I}_{i,t}\right]\right)} = -\frac{var\left[\epsilon_{D,t}\right]}{var\left[\epsilon_{Z,t}\right]} \frac{\frac{\partial P_t}{\partial \epsilon_{D,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}}. \tag{52}$$

Equality (52) follows entirely from updating rules, making use of the fact that all relevant date t information sets include the price P_t. Substitution into (51) yields

$$\frac{1}{\bar{A}}\frac{\partial q_{0,t}}{\partial \epsilon_{Z,t}} = \mathcal{T}\frac{var\left[\epsilon_{D,t}\right]}{var\left[\epsilon_{Z,t}\right]}\frac{\frac{\partial P_t}{\partial \epsilon_{D,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}}\frac{\partial}{\partial \epsilon_{D,t}}\left(X_t - E\left[X_t|\mathcal{I}_{0,i,t}\right]\right),$$

and hence, using (11),

$$\frac{1}{\bar{A}}\frac{\partial q_{0,t}}{\partial \epsilon_{Z,t}} = -\frac{var\left[\epsilon_{D,t}\right]}{var\left[\epsilon_{Z,t}\right]}\frac{\frac{\partial P_t}{\partial \epsilon_{D,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}}\frac{\partial E\left[X_t|\mathcal{I}_{0,i,t}\right]}{\partial \epsilon_{D,t}}.$$

Also from (11),

$$\frac{\frac{1}{\bar{A}}\frac{\partial q_{0,t}}{\partial \epsilon_{D,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}} = -\frac{\mathcal{T}}{1-\mathcal{T}}.$$

Substitution of these last two equalities into (18) gives

$$cov\left[\frac{q_{0,t}}{\bar{A}}, X_{t+1}\right] = -\frac{\mathcal{T}}{1-\mathcal{T}}\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right] - \frac{\frac{\partial E\left[X_{t+1}|\mathcal{I}_{0,i,t}\right]}{\partial \epsilon_{D,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}}\frac{\frac{\partial P_t}{\partial \epsilon_{D,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}}\frac{\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}}{\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}}\left(\frac{\partial X_{t+1}}{\partial \epsilon_{D,t}}\right)^2 var\left[\epsilon_{D,t}\right],$$

which simplifies to (19).

It remains to establish (52). By the analogue of (3) for the information set $\mathcal{I}_{i,t}$, expressions (49) and (50) from the proof of Lemma 2 can be written as

$$\begin{aligned}
\frac{\partial}{\partial \epsilon_{D,t}}E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] - \begin{pmatrix}1\\0\end{pmatrix} &= -\frac{1}{var\left[\epsilon_{D,t}\right]}\begin{pmatrix}var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]\\ cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]\end{pmatrix}\\
\frac{\partial}{\partial \epsilon_{Z,t}}E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] - \begin{pmatrix}0\\1\end{pmatrix} &= -\frac{1}{var\left[\epsilon_{Z,t}\right]}\begin{pmatrix}cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]\\ \frac{cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]^2}{var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]}\end{pmatrix},
\end{aligned}$$

and hence

$$\frac{\partial}{\partial \epsilon_{Z,t}}\left(E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] - \begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right) = \frac{var\left[\epsilon_{D,t}\right]}{var\left[\epsilon_{Z,t}\right]}\frac{cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]}{var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]}\frac{\partial}{\partial \epsilon_{D,t}}\left(E\left[\begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}|\mathcal{I}_{i,t}\right] - \begin{pmatrix}\epsilon_{D,t}\\ \epsilon_{Z,t}\end{pmatrix}\right)$$

Moreover, the analogue of (2) for the information set $\mathcal{I}_{i,t}$ implies

$$\frac{cov\left[\epsilon_{D,t},\epsilon_{Z,t}|\mathcal{I}_{i,t}\right]}{var\left[\epsilon_{D,t}|\mathcal{I}_{i,t}\right]} = -\frac{c_D}{c_Z} = -\frac{\frac{\partial P_t}{\partial \epsilon_{D,t}}}{\frac{\partial P_t}{\partial \epsilon_{Z,t}}},$$

delivering (52) and completing the proof.

B Appendix: Parameter values in (22)-(24)

B.1 From observable moments to parameter values

The parameters to estimate are: $\left\{\bar{\Delta d}, \bar{pd}, \rho, \phi_g, \phi_\mu, a_{pd}, a_d, \sigma_{pd}^2, \sigma_d^2, \sigma_{pd,d}\right\}$.

Of these, $\bar{pd}$ and $\bar{\Delta d}$ are estimated using the sample means of pd_t and Δd_t, and ρ is in turn a function of $\bar{pd}$.

The remaining seven parameters $\left\{\bar{\Delta d}, \bar{pd}, \rho, \phi_g, \phi_\mu, a_{pd}, a_d, \sigma_{pd}^2, \sigma_d^2, \sigma_{pd,d}\right\}$ are estimated from the observed variance and covariance of pd_t and Δd_t, including lags.

As preliminaries: Let a_0 denote the coefficient on g_t in the pd_t transition equation,

$$a_0 = \frac{\phi_g - \phi_\mu}{1 - \rho\phi_g}.$$

The following variance and covariances, which are not directly observable, enter many expressions below:

$$\begin{aligned}
var\left[g_t\right] &= \frac{var\left[a_{pd}\nu_{pd,t} + a_d\nu_{d,t}\right]}{1-\phi_g^2} = \frac{a_{pd}^2\sigma_{pd}^2 + a_d^2\sigma_d^2 + 2a_{pd}a_d\sigma_{pd,d}}{1-\phi_g^2} \qquad (53)\\
cov\left[g_t, \Delta d_t\right] &= \phi_g var\left[g_t\right] + cov\left[a_{pd}\nu_{pd,t} + a_d\nu_{d,t}, \nu_{d,t}\right] = \phi_g var\left[g_t\right] + a_{pd}\sigma_{pd,d} + a_d\sigma_d^2\\
cov\left[g_t, pd_t\right] &= \frac{a_0\phi_g var\left[g_t\right] + cov\left[a_{pd}\nu_{pd,t} + a_d\nu_{d,t}, \nu_{pd,t}\right]}{1-\phi_g\phi_\mu} = \frac{a_0\phi_g var\left[g_t\right] + a_{pd}\sigma_{pd}^2 + a_d\sigma_{pd,d}}{1-\phi_g\phi_\mu}.
\end{aligned}$$

The observable moments are

$$\begin{aligned}
var\left[pd_t\right] &= \frac{a_0^2 var\left[g_t\right] + 2a_0\phi_\mu cov\left[g_t, pd_t\right] + \sigma_{pd}^2}{1-\phi_\mu^2} \qquad (54)\\
var\left[\Delta d_t\right] &= var\left[g_t\right] + \sigma_d^2 \qquad (55)\\
cov\left[\Delta d_t, pd_t\right] &= a_0 var\left[g_t\right] + \phi_\mu cov\left[g_t, pd_t\right] + \sigma_{pd,d}, \qquad (56)
\end{aligned}$$

and

$$\begin{aligned}
cov\left[\Delta d_{t+1}, \Delta d_t\right] &= cov\left[g_t + \nu_{d,t+1}, \Delta d_t\right] = cov\left[g_t, \Delta d_t\right] \qquad (57)\\
cov\left[\Delta d_{t+1}, pd_t\right] &= cov\left[g_t + \nu_{d,t+1}, pd_t\right] = cov\left[g_t, pd_t\right] \qquad (58)\\
cov\left[pd_{t+1}, \Delta d_t\right] &= cov\left[a_0 g_t + \phi_\mu pd_t + \nu_{pd,t+1}, \Delta d_t\right]\\
&= a_0 cov\left[g_t, \Delta d_t\right] + \phi_\mu cov\left[pd_t, \Delta d_t\right]\\
&= a_0 cov\left[\Delta d_{t+1}, \Delta d_t\right] + \phi_\mu cov\left[pd_t, \Delta d_t\right]\\
cov\left[pd_{t+1}, pd_t\right] &= cov\left[a_0 g_t + \phi_\mu pd_t + \nu_{pd,t+1}, pd_t\right]
\end{aligned}$$

$$
\begin{aligned}
&= a_0 cov\left[g_t, pd_t\right] + \phi_\mu var\left[pd_t\right] \\
&= a_0 cov\left[\Delta d_{t+1}, pd_t\right] + \phi_\mu var\left[pd_t\right]
\end{aligned}
$$

and

$$
\begin{aligned}
cov\left[\Delta d_{t+2}, \Delta d_t\right] &= \phi_g cov\left[g_t, \Delta d_t\right] = \phi_g cov\left[\Delta d_{t+1}, \Delta d_t\right] \\
cov[\Delta d_{t+2}, pd_t] &= \phi_g cov\left[g_t, pd_t\right] = \phi_g cov\left[\Delta d_{t+1}, pd_t\right] \\
cov\left[pd_{t+2}, \Delta d_t\right] &= cov\left[a_0 g_{t+1} + \phi_\mu pd_{t+1}, \Delta d_t\right] \\
&= a_0 \phi_g cov\left[g_t, \Delta d_t\right] + \phi_\mu cov\left[pd_{t+1}, \Delta d_t\right] \\
&= a_0 \phi_g cov\left[\Delta d_{t+1}, \Delta d_t\right] + \phi_\mu cov\left[pd_{t+1}, \Delta d_t\right] \\
cov\left[pd_{t+2}, pd_t\right] &= cov\left[a_0 g_{t+1} + \phi_\mu pd_{t+1}, pd_t\right] \\
&= a_0 \phi_g cov\left[g_t, pd_t\right] + \phi_\mu cov\left[pd_{t+1}, pd_t\right] \\
&= a_0 \phi_g cov\left[\Delta d_{t+1}, pd_t\right] + \phi_\mu cov\left[pd_{t+1}, pd_t\right].
\end{aligned}
$$

(One can continue to compute further lag covariance, but doing so does not yield any additional information.)

The parameter ϕ_g is given by

$$
\phi_g = \frac{cov\left[\Delta d_{t+2}, \Delta d_t\right]}{cov\left[\Delta d_{t+1}, \Delta d_t\right]}.
$$

Given ϕ_g, the parameter ϕ_μ can be inferred from the combination of $cov\left[pd_{t+2}, pd_t\right]$, $cov\left[\Delta d_{t+1}, pd_t\right]$ and $cov\left[pd_{t+1}, pd_t\right]$.

Given ϕ_g and ϕ_μ, the remaining non-redundant moment conditions are (54)-(58).

To solve for $\left\{a_{pd}, a_d, \sigma_{pd}^2, \sigma_d^2, \sigma_{pd,d}\right\}$, first substitute (58) into (54) and (56) and rearrange to yield expressions for $\sigma_{pd}^2, \sigma_d^2, \sigma_{pd,d}$ in terms of observable moments and $var\left[g_t\right]$.

$$
\begin{aligned}
\sigma_{pd}^2 &= \left(1 - \phi_\mu^2\right) var\left[pd_t\right] - a_0^2 var\left[g_t\right] - 2a_0\phi_\mu cov\left[\Delta d_{t+1}, pd_t\right] && (59) \\
\sigma_d^2 &= var\left[\Delta d_t\right] - var\left[g_t\right] && (60) \\
\sigma_{pd,d} &= cov\left[\Delta d_t, pd_t\right] - a_0 var\left[g_t\right] - \phi_\mu cov\left[\Delta d_{t+1}, pd_t\right]. && (61)
\end{aligned}
$$

Next, substitute in for $cov\left[g_t, \Delta d_t\right]$ and $cov\left[g_t, pd_t\right]$ in (57) and (58) and rearrange to yield

$$
\begin{aligned}
a_{pd}\sigma_{pd,d} + a_d\sigma_d^2 &= cov\left[\Delta d_{t+1}, \Delta d_t\right] - \phi_g var\left[g_t\right] \\
a_{pd}\sigma_{pd}^2 + a_d\sigma_{pd,d} &= \left(1 - \phi_g\phi_\mu\right) cov\left[\Delta d_{t+1}, pd_t\right] - a_0\phi_g var\left[g_t\right],
\end{aligned}
$$

and hence

$$\begin{aligned}\left(\sigma_{pd,d}^2-\sigma_{pd}^2\sigma_d^2\right)a_{pd} &= \sigma_{pd,d}\left(cov\left[\Delta d_{t+1},\Delta d_t\right]-\phi_g var\left[g_t\right]\right)\\ &- \sigma_d^2\left(\left(1-\phi_g\phi_\mu\right)cov\left[\Delta d_{t+1},pd_t\right]-a_0\phi_g var\left[g_t\right]\right) \qquad (62)\\ \left(\sigma_{pd,d}^2-\sigma_{pd}^2\sigma_d^2\right)a_d &= \sigma_{pd,d}\left(\left(1-\phi_g\phi_\mu\right)cov\left[\Delta d_{t+1},pd_t\right]-a_0\phi_g var\left[g_t\right]\right)\\ &- \sigma_{pd}^2\left(cov\left[\Delta d_{t+1},\Delta d_t\right]-\phi_g var\left[g_t\right]\right). \qquad (63)\end{aligned}$$

Together, equations (59)-(63) give $\left\{a_{pd},a_d,\sigma_{pd}^2,\sigma_d^2,\sigma_{pd,d}\right\}$ in terms of observable moments and $var\left[g_t\right]$. The term $var\left[g_t\right]$ itself can be solved for using (53).

B.2 Recovering observable moments from the reported estimates in BK

BK estimate the system

$$\begin{aligned}g_{t+1}-\bar{\Delta d}^{BK} &= \phi_g^{BK}\left(g_t-\bar{\Delta d}^{BK}\right)+\nu_{g,t+1}^{BK}\\ \Delta d_{t+1} &= g_t+\nu_{d,t+1}^{BK}\\ pd_{t+1}-\bar{pd}^{BK} &= \frac{\phi_g^{BK}-\phi_\mu^{BK}}{1-\rho^{BK}\phi_g^{BK}}\left(g_t-\bar{\Delta d}^{BK}\right)+\phi_\mu^{BK}\left(pd_t-\bar{pd}^{BK}\right)\\ &- \frac{1}{1-\rho^{BK}\phi_\mu^{BK}}\nu_{\mu,t+1}^{BK}+\frac{1}{1-\rho^{BK}\phi_g^{BK}}\nu_{g,t+1}^{BK}\end{aligned}$$

under the restriction that $cov\left[\nu_{g,t+1}^{BK},\nu_{d,t+1}^{BK}\right]=0$.

First note that the relation between $\left\{\bar{\Delta d}^{BK},\bar{pd}^{BK},\rho^{BK},\phi_g^{BK},\phi_\mu^{BK}\right\}$ and observable moments is exactly the same as the relation between $\left\{\bar{\Delta d},\bar{pd},\rho,\phi_g,\phi_\mu\right\}$ and observable moments. So is immediate that

$$\left\{\bar{\Delta d},\bar{pd},\rho,\phi_g,\phi_\mu\right\}=\left\{\bar{\Delta d}^{BK},\bar{pd}^{BK},\rho^{BK},\phi_g^{BK},\phi_\mu^{BK}\right\}. \qquad (64)$$

As such, it is only necessary to recover the five moments that are use to infer $\left\{a_{pd},a_d,\sigma_{pd}^2,\sigma_d^2,\sigma_{pd,d}\right\}$, namely $var\left[pd_t\right]$, $var\left[\Delta d_t\right]$, $cov\left[\Delta d_t,pd_t\right]$, $cov\left[\Delta d_{t+1},\Delta d_t\right]$ and $cov\left[\Delta d_{t+1},pd_t\right]$. Explicit evaluation implies that these five moments are given by the following expressions (in light of (64), I drop the BK subscripts on $\left\{\bar{\Delta d},\bar{pd},\rho,\phi_g,\phi_\mu\right\}$):

$$var\left[g_t\right] = \frac{\left(\sigma_g^{BK}\right)^2}{1-\phi_g^2}$$

$$
\begin{aligned}
cov\left[g_t, \Delta d_t\right] &= \phi_g var\left[g_t\right] \\
cov\left[g_t, pd_t\right] &= \frac{a_0\phi_g var\left[g_t\right] - \frac{\sigma_{\mu,g}^{BK}}{1-\rho\phi_\mu} + \frac{\left(\sigma_g^{BK}\right)^2}{1-\rho\phi_g}}{1-\phi_g\phi_\mu} \\
var\left[pd_t\right] &= \frac{a_0^2 var\left[g_t\right] + 2a_0\phi_\mu cov\left[g_t, pd_t\right]}{1-\phi_\mu^2} \\
&+ \frac{\left(\frac{1}{1-\rho\phi_\mu}\right)^2\left(\sigma_\mu^{BK}\right)^2 + \left(\frac{1}{1-\rho\phi_g}\right)^2\left(\sigma_g^{BK}\right)^2 - 2\frac{1}{1-\rho\phi_\mu}\frac{1}{1-\rho\phi_g}\sigma_{\mu,g}^{BK}}{1-\phi_\mu^2} \\
var\left[\Delta d_t\right] &= var\left[g_t\right] + \left(\sigma_d^{BK}\right)^2 \\
cov\left[\Delta d_t, pd_t\right] &= a_0 var\left[g_t\right] + \phi_\mu cov\left[g_t, pd_t\right] - \frac{1}{1-\rho\phi_\mu}\sigma_{\mu,d}^{BK} \\
cov\left[\Delta d_{t+1}, \Delta d_t\right] &= cov\left[g_t, \Delta d_t\right] \\
cov\left[\Delta d_{t+1}, pd_t\right] &= cov\left[g_t, pd_t\right].
\end{aligned}
$$

C Appendix: Details for calculations in subsections 5.3 to 5.9

C.1 Details for term (25)

The estimated coefficient b_{pd} is

$$
b_{pd} = \frac{\partial E\left[\ln\left(\frac{P_{t+1}+D_{t+1}}{P_t}\right)|\mathcal{J}_t\right]}{\partial \ln P_t}\frac{\partial E\left[\ln\left(\frac{X_{t+1}+RP_t}{P_t}\right)|\mathcal{J}_t\right]}{\partial \ln P_t} = P_t\frac{\partial E\left[\ln\left(X_{t+1}+RP_t\right)|\mathcal{J}_t\right]}{\partial P_t} - 1.
$$

For any random variable Y, $E\left[\ln Y\right] \approx \ln E\left[Y\right] - \frac{1}{2}\frac{var[Y]}{E[Y]^2}$.[5] Hence

$$
b_{pd} \approx P_t\left(\frac{1}{E\left[X_{t+1}+RP_t|\mathcal{J}_t\right]} + \frac{var\left[X_{t+1}+RP_t|\mathcal{J}_t\right]}{E\left[X_{t+1}+RP_t|\mathcal{J}_t\right]^3}\right)\frac{\partial}{\partial P_t}E\left[X_{t+1}+RP_t|\mathcal{J}_t\right] - 1
$$

[5] To obtain this approximation: The second-order Taylor expansion is

$$
\ln Y \approx \ln E\left[Y\right] + \frac{Y-E\left[Y\right]}{E\left[Y\right]} - \frac{1}{2}\frac{\left(Y-E\left[Y\right]\right)^2}{E\left[Y\right]^2}.
$$

Hence

$$
E\left[\ln Y\right] \approx \ln E\left[Y\right] - \frac{1}{2}\frac{var\left[Y\right]}{E\left[Y\right]^2}.
$$

$$= \frac{1}{E\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]+R}\left(1+\frac{var\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]}{\left(E\left[\frac{X_{t+1}}{P_t}|\mathcal{J}_t\right]+R\right)^2}\right)\left(\frac{\partial}{\partial P_t}E\left[X_{t+1}|\mathcal{J}_t\right]+R\right)-1,$$

yielding (25).

C.2 Details for (32)

$$\frac{\partial P_t}{\partial \epsilon_{D,t}} = \left.\frac{\partial P_t}{\partial D_{t+1}}\right|_{\Delta d_t,\mathcal{J}_{t-1}} = \frac{P_t}{D_{t+1}}\left.\frac{\partial \log P_t}{\partial \log D_{t+1}}\right|_{\Delta d_t,\mathcal{J}_{t-1}} = \frac{P_t}{D_{t+1}}\left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t,\mathcal{J}_{t-1}}.$$

To evaluate, replace $\frac{P_t}{D_{t+1}}$ with its expected value,

$$\begin{aligned}
E\left[\frac{P_t}{D_{t+1}}\right] = E\left[\frac{\frac{P_t}{D_t}}{\frac{D_{t+1}}{D_t}}\right] &\approx \frac{E\left[\frac{P_t}{D_t}\right]}{E\left[\frac{D_{t+1}}{D_t}\right]} + cov\left[\frac{P_t}{D_t}, \frac{1}{\frac{D_{t+1}}{D_t}}\right] \\
&\approx \frac{\exp\left(E\left[pd_t\right]\right)}{\exp\left(\Delta d_{t+1}\right)} + cov\left[\exp\left(pd_t\right), \exp\left(-\Delta d_{t+1}\right)\right] \\
&\approx \frac{\exp\left(\bar{pd}\right)}{\exp\left(\bar{g}\right)}\left(1 - cov\left[pd_t, \Delta d_{t+1}\right]\right).
\end{aligned}$$

C.3 Details for (35)

$$\begin{aligned}
\left.\frac{\partial\left(P_{t+1}+D_{t+1}\right)}{\partial D_{t+1}}\right|_{\mathcal{J}_t} &= 1 + \frac{P_{t+1}}{D_{t+1}}\left.\frac{\partial \log P_{t+1}}{\partial \log D_{t+1}}\right|_{\mathcal{J}_t} \\
&= 1 + \frac{P_{t+1}}{D_{t+1}}\left.\frac{\partial \log \frac{P_{t+1}}{D_{t+1}}}{\partial \log \frac{D_{t+1}}{D_t}} + 1\right|_{\mathcal{J}_t} \\
&= 1 + e^{pd_{t+1}}\left(\left.\frac{\partial pd_{t+1}}{\partial \Delta d_{t+1}}\right|_{\mathcal{J}_t} + 1\right) \\
&= 1 + e^{pd_{t+1}}\left(\frac{cov\left[\nu_{pd,t+1}, \nu_{d,t+1}\right]}{var\left[\nu_{d,t+1}\right]} + 1\right) \\
&\approx 1 + e^{\bar{pd}}\left(\frac{cov\left[\nu_{pd,t+1}, \nu_{d,t+1}\right]}{var\left[\nu_{d,t+1}\right]} + 1\right).
\end{aligned}$$

C.4 Details for (38)

Evaluating (37):

$$
\begin{aligned}
var\left[\frac{\partial P_t}{\partial \epsilon_{D,t}}\epsilon_{D,t}\right] &= D_t^2\left(var\left[e^{pd_t}|\Delta d_t,\mathcal{J}_{t-1}\right] - var\left[e^{pd_t}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right]\right) \\
&= \left(P_{t-1}e^{\Delta d_t - pd_{t-1}}\right)^2\left(var\left[e^{pd_t}|\Delta d_t,\mathcal{J}_{t-1}\right] - var\left[e^{pd_t}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right]\right) \\
&\approx P_{t-1}^2 e^{2\bar{\Delta}d - 2\bar{pd}}\left(e^{\bar{pd}}\right)^2\left(var\left[pd_t|\Delta d_t,\mathcal{J}_{t-1}\right] - var\left[pd_t|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right]\right) \\
&= P_{t-1}^2 e^{2\bar{\Delta}d}\left(var\left[\nu_{pd,t}|\nu_{d,t}\right] - var\left[\nu_{pd,t}|\nu_{d,t}, a_{pd}\nu_{pd,t} + a_d\nu_{d,t} + \nu_{d,t+1}\right]\right).
\end{aligned}
$$

Evaluating,

$$
\begin{aligned}
var\left[\nu_{pd,t}|\nu_{d,t}, a_{pd}\nu_{pd,t} + a_d\nu_{d,t} + \nu_{d,t+1}\right] &= var\left[\nu_{pd,t}|\nu_{d,t}, a_{pd}\nu_{pd,t} + \nu_{d,t+1}\right] \\
&= var\left[\nu_{pd,t}|\nu_{d,t}\right] - \frac{a_{pd}^2 var\left[\upsilon_{pd,t}|\nu_{d,t}\right]^2}{var\left[\nu_{d,t+1}\right] + a_{pd}^2 var\left[\upsilon_{pd,t}|\nu_{d,t}\right]},
\end{aligned}
$$

yielding (38).

C.5 Details for (39)

Expanding

$$
\begin{aligned}
& \frac{1}{D_t}E\left[X_{t+1}|D_{t+1},\mathcal{J}_t\right] \\
= \;& E\left[e^{pd_{t+1}+\Delta d_{t+1}} + e^{\Delta d_{t+1}} - Re^{pd_t}|D_{t+1},\mathcal{J}_t\right] \\
\approx \;& e^{E[pd_{t+1}+\Delta d_{t+1}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}]}\left(E\left[pd_{t+1}+\Delta d_{t+1}|\Delta d_{t+1},\mathcal{J}_t\right] - E\left[pd_{t+1}+\Delta d_{t+1}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right]\right) \\
+ \;& e^{E[\Delta d_{t+1}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}]}\left(E\left[\Delta d_{t+1}|\Delta d_{t+1},\mathcal{J}_t\right] - E\left[\Delta d_{t+1}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right]\right) \\
- \;& Re^{E[pd_t|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}]}\left(E\left[pd_t|\Delta d_{t+1},\mathcal{J}_t\right] - E\left[pd_t|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right]\right) \\
+ \;& e^{E[pd_{t+1}+\Delta d_{t+1}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}]} + e^{E[\Delta d_{t+1}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}]} - Re^{E[pd_t|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}]}
\end{aligned}
$$

Note that

$$
\begin{aligned}
& E\left[pd_{t+1}|\Delta d_{t+1},\mathcal{J}_t\right] - E\left[pd_{t+1}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right] \\
= \;& E\left[\frac{\phi_g-\phi_\mu}{1-\rho\phi_g}g_t + \phi_\mu pd_t|\Delta d_{t+1},\mathcal{J}_t\right] - E\left[\frac{\phi_g-\phi_\mu}{1-\rho\phi_g}g_t + \phi_\mu pd_t|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right] \\
= \;& \left(\frac{\phi_g-\phi_\mu}{1-\rho\phi_g}a_{pd} + \phi_\mu\right)\left(\nu_t^{pd} - E\left[\nu_t^{pd}|\nu_t^d\right]\right)
\end{aligned}
$$

and

$$\begin{aligned}
& E\left[pd_t|\Delta d_{t+1},\mathcal{J}_t\right]-E\left[pd_t|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}\right] \\
= \; & \nu_t^{pd}-E\left[\nu_t^{pd}|\nu_t^d\right].
\end{aligned}$$

Hence

$$\begin{aligned}
& \frac{1}{D_t}E\left[X_{t+1}|D_{t+1},\mathcal{J}_t\right] \\
\approx \; & \left(e^{E[pd_{t+1}+\Delta d_{t+1}|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}]}\left(\frac{\phi_g-\phi_\mu}{1-\rho\phi_g}a_{pd}+\phi_\mu\right)-Re^{E[pd_t|\Delta d_{t+1},\Delta d_t,\mathcal{J}_{t-1}]}\right)\left(\nu_t^{pd}-E\left[\nu_t^{pd}|\nu_t^d\right]\right),
\end{aligned}$$

and so

$$\begin{aligned}
& \left(\frac{\partial X_{t+1}}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right] \\
\approx \; & \left(P_{t-1}e^{\Delta d_t-pd_{t-1}}\right)^2\left(e^{\bar{pd}}\left(e^{\Delta d}\left(\frac{\phi_g-\phi_\mu}{1-\rho\phi_g}a_{pd}+\phi_\mu\right)-R\right)\right)^2 var\left[\nu_t^{pd}|\nu_t^d\right] \\
\approx \; & P_{t-1}^2 e^{2\bar{\Delta d}}\left(e^{\bar{\Delta d}}\left(\frac{\phi_g-\phi_\mu}{1-\rho\phi_g}a_{pd}+\phi_\mu\right)-R\right)^2 var\left[\nu_t^{pd}|\nu_t^d\right].
\end{aligned}$$

C.6 Details for (41)

Using (28) and (29) in the second equality:

$$\begin{aligned}
\left(\frac{\partial P_t}{\partial \epsilon_{Z,t}}\right)^2 var\left[\epsilon_{Z,t}\right] \; & \approx \; D_t^2 e^{2\bar{pd}}\left(var\left[\nu_{pd,t}|\nu_{d,t}\right]-\left(\left.\frac{\partial pd_t}{\partial \Delta d_{t+1}}\right|_{\Delta d_t,\mathcal{J}_{t-1}}\right)^2 var\left[\Delta d_{t+1}|\Delta d_t,\mathcal{J}_{t-1}\right]\right) \\
& = \; D_t^2 e^{2\bar{pd}}\left(var\left[\nu_{pd,t}|\nu_{d,t}\right]-\frac{a_{pd}^2 var\left[\nu_{pd,t}|\nu_{d,t}\right]^2}{var\left[\nu_{d,t+1}\right]+a_{pd}^2 var\left[\nu_{pd,t}|\nu_{d,t}\right]}\right) \\
& = \; P_{t-1}^2 e^{2\Delta d_t-2pd_{t-1}}e^{2\bar{pd}}\frac{var\left[\nu_{pd,t}|\nu_{d,t}\right]var\left[\nu_{d,t+1}\right]}{var\left[\nu_{d,t+1}\right]+a_{pd}^2 var\left[\nu_{pd,t}|\nu_{d,t}\right]} \\
& \approx \; P_{t-1}^2 e^{2\bar{\Delta d}}\frac{var\left[\nu_{d,t+1}\right]var\left[\nu_{pd,t}|\nu_{d,t}\right]}{var\left[\nu_{d,t+1}\right]+a_{pd}^2 var\left[\nu_{pd,t}|\nu_{d,t}\right]}.
\end{aligned}$$

C.7 Details for (42)

$$E\left[X_{t+1}\right] \; = \; E\left[E\left[D_t\left(e^{r_{t+1}}-R\right)\frac{P_t}{D_t}|\mathcal{J}_t\right]\right]$$

$$
\begin{aligned}
&= E\left[E\left[P_{t-1}e^{\Delta d_t - pd_{t-1}}\left(e^{r_{t+1}}-R\right)e^{pd_t}|\mathcal{J}_t\right]\right] \\
&\approx P_{t-1}e^{\bar{\Delta d}}\left(e^{\bar{\mu}}-R\right).
\end{aligned}
$$

D Appendix: Standard present value approximation

$$
\begin{aligned}
r_{t+1} &= \log\frac{P_{t+1}+D_{t+1}}{P_t} \\
&= \log\frac{P_{t+1}+D_{t+1}}{D_{t+1}}\frac{D_{t+1}}{D_t}\frac{D_t}{P_t} \\
&= \log\left(1+\exp\left(\log\frac{P_{t+1}}{D_{t+1}}\right)\right)+\log\frac{D_{t+1}}{D_t}-\log\frac{P_t}{D_t} \\
&= \log\left(1+\exp\left(pd_{t+1}\right)\right)+\Delta d_{t+1}-pd_t \\
&\approx \log\left(1+\exp\left(\bar{pd}\right)\right)+\frac{\exp\left(\bar{pd}\right)}{1+\exp\left(\bar{pd}\right)}\left(pd_{t+1}-\bar{pd}\right) \\
&+ \Delta d_{t+1}-pd_t.
\end{aligned}
$$

Recalling $\rho=\frac{\exp\left(\bar{pd}\right)}{1+\exp\left(\bar{pd}\right)}$ and defining $K_{\bar{pd}}=\log\left(1+\exp\left(\bar{pd}\right)\right)-\rho\bar{pd}$

$$
pd_t \approx \rho pd_{t+1}+\Delta d_{t+1}-r_{t+1}+K_{\bar{pd}}.
$$

Iterating forwards

$$
pd_t \approx \sum_{s=0}^{\infty}\rho^{s-1}\left(\Delta d_{t+s}-r_{t+s}\right)+\frac{K_{\bar{pd}}}{1-\rho}.
$$

Taking expectations of both sides, the AR1 assumption implies

$$
pd_t \approx \frac{g_t-\bar{\Delta d}}{1-\rho\phi_g}-\frac{\mu_t-\bar{\mu}}{1-\rho\phi_\mu}+\frac{\bar{\Delta d}-\bar{\mu}}{1-\rho}+\frac{K_{\bar{pd}}}{1-\rho},
$$

and hence

$$
pd_t-\bar{pd} \approx \frac{g_t-\bar{\Delta d}}{1-\rho\phi_g}-\frac{\mu_t-\bar{\mu}}{1-\rho\phi_\mu}.
$$