Collegio Carlo Alberto

Rational Learning and the Term Structures of Value and Growth Risk Premia

Michael Hasler
Mariana Khapko
Roberto Marfè

No. 622
December 2020

Carlo Alberto Notebooks

www.carloalberto.org/research/working-papers

Rational Learning and the Term Structures of Value and Growth Risk Premia*

Michael Hasler[†] Mariana Khapko[‡] Roberto Marfè[§]

May 26, 2020

Abstract

This paper studies the impact of information processing and rational learning about economic fundamentals on the level and timing of risk premium in the cross-section of firms. Learning helps explain the level of the value premium, and why the term structure of risk premium is increasing for value firms and decreasing for growth firms. Moreover, learning yields an upward-sloping term structure of interest rates and a downward-sloping term structure of market risk premium, whereas the full information economy predicts the opposite shapes. Therefore, rational learning helps understand the level and timing of expected returns observed in the cross-section of risky and risk-free assets.

Keywords: Asset Pricing, Rational Learning, Term Structures, Value and Growth Firms

JEL Classification. D51, D53, D83, G12

*The paper was previously circulated under the title: “Rational Learning and Term Structures”. We would like to thank Daniel Andrei, Adem Atmaz, Patrick Augustin, Andrea Buffa, Alexandre Corhay, Julien Cujean, Olivier Dessaint, Alexandre Jeanneret, Ralph Koijen, Holger Kraft, Jean-Marie Meier, Johannes Muhle-Karbe, Chayawat Ornthanalai, Marcel Rindisbacher, Alejandro Rivera, Scott Robertson, Christian Schlag, Mike Simutin, Julian Thimme, Steven Xiao, and seminar and conference participants at the Collegio Carlo Alberto, Goethe University Frankfurt, the University of Toronto, and the University of Texas at Dallas for their insightful comments. Research support from Long-Term Investors@UniTo (LTI@UniTO), the University of Texas at Dallas, and the University of Toronto is gratefully acknowledged.

[†]Naveen Jindal School of Management, University of Texas at Dallas, 800 W Campbell Rd, Richardson, TX 75080, USA; michael.hasler5@gmail.com; <https://sites.google.com/michaelhasler>

[‡]Department of Management, University of Toronto Scarborough, 1095 Military Trail, Toronto, ON, M1C 1A4, Canada; Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, ON, M5S 3E6, Canada; mariana.khapko@utoronto.ca; www.marianakhapko.com

[§]Collegio Carlo Alberto, Piazza Vincenzo Arbarello 8, 10122 Torino, Italy; roberto.marfe@carloalberto.org; <http://robertomarfe.altervista.org/>

1 Introduction

To build forecasts about future economic conditions, investors need to gather, analyze, and filter a large amount of available information. By efficiently and rationally processing this flow of information over time, investors update their forecasts and use them to dynamically invest in financial markets. Therefore, information processing impacts asset prices through investors' trades.

We study the role of information processing and learning in an asset-pricing model with time-varying economic fundamentals. Investors, unable to directly observe the fundamentals, update their beliefs about future growth prospects at both the firm and aggregate levels by processing available sources of information. We show that information processing and rational learning help explain the level and timing of risk premium observed in the cross-section of firms. More accurate information on a firm's fundamental translates into higher prices and lower risk, particularly so in the long term. This explains why firms with high valuations—growth firms—have a low risk premium and a downward-sloping term structure of risk premium. In contrast, less accurate information on firms' fundamentals implies that risk is higher and concentrated in the long term. As a result, these firms have low valuations—value firms—a high risk premium and an upward-sloping term structure of risk premium.

The aggregate output of the economy is driven by a time-varying expected growth rate and a contemporaneous shock. The expected growth rate is unobservable, and therefore needs to be estimated by the representative agent using observed changes in output. Following [Marfè \(2017\)](#), we assume that the expected growth rate depends on two factors. The first is the standard permanent component considered in the literature.¹ The second is a transitory component that captures short-term business cycle fluctuations. The permanent component implies that output growth risk tends to increase with the time horizon, whereas the transitory component has the exact opposite impact. Since the impact of the transitory

¹See [Detemple \(1986, 1991\)](#), [Veronesi \(1999, 2000\)](#), [Brennan and Xia \(2001\)](#), [Bansal and Yaron \(2004\)](#), and [Berrada \(2006, 2009\)](#) among others.

component dampens that of the permanent component, the model-implied term structure of output growth risk is flat, as in the data (Marfè, 2016; Dew-Becker, 2017). The market and firms' dividends depend on both the permanent and transitory components of output. In addition, each firm's dividend is also driven by an unobservable idiosyncratic expected growth component. The agent needs to estimate this idiosyncratic component by observing the firm's dividend as well as a news signal representing information provided by analysts.

We show that information processing and rational learning help explain the level and timing of risk premium observed in the cross-section of equities. Specifically, learning implies that the risk premium of growth firms is about 5% lower than that of value firms, thereby explaining the observed value premium. In addition, the term structure of risk premium is downward-sloping for growth firms and upward-sloping for value firms, consistent with recent empirical findings (Giglio, Kelly, and Kozak, 2020). Learning also yields an increasing term structure of interest rates and a decreasing term structure of market risk premium (van Binsbergen, Brandt, and Koijen, 2012), whereas the corresponding economy with full information predicts the opposite shapes.²

The economic mechanism is as follows. Each source of risk has a specific effect on the term structure of cashflow growth risk. The permanent component has an upward-sloping effect because its shocks accumulate over time, which yields more risk in the long term than in the short term. In contrast, the transitory component has a downward-sloping effect because its shocks mean revert towards zero, which yields less risk in the long run than in the short run. Although the two effects offset each other when the permanent and transitory components are fully observable, the agent assigns a larger weight to the (estimated) transitory component when both components are unobservable and need to be filtered out. As a result, the term structures of output growth risk and market dividend growth risk are downward sloping.

²While we are agnostic about the empirical issue of correctly measuring the unconditional slope of the term structure of market risk premium (Bansal, Miller, Song, and Yaron, 2019), our paper identifies rational learning as an economic channel capable to explain sizable risk premia on short-term claims. These have been documented by van Binsbergen et al. (2012) and cannot be explained by leading models, independently of the sign of the unconditional slope.

Since bonds are used to hedge against consumption growth risk, and consumption equals output in equilibrium, the demand for short-term bonds is larger than that for long-term bonds. As a result, bond yields are lower in the short term than in the long term; the term structure of interest rates is upward sloping. Furthermore, the agent requires a higher risk premium to bear a higher dividend growth risk in equilibrium, which explains why the model-implied term structure of market risk premium is downward sloping.

In the cross-section, the impact of the firm's idiosyncratic expected dividend growth component on the risk premium adds to the downward-sloping effect implied by the combination of the permanent and transitory components. When the news signal about the idiosyncratic component is perfectly informative, the agent observes the shocks driving the idiosyncratic component. Since these idiosyncratic shocks are not priced, the idiosyncratic component has no effect on the firm's risk premium when the signal is perfectly informative. Therefore, the firm has a fairly low risk premium, high valuations, and a downward-sloping term structure of risk premium as the market. When the precision of the news signal is low, however, the agent uses information provided by both dividend growth surprises and the news signal to filter out the idiosyncratic expected growth component. When the dividend growth surprise is positive (resp., negative), the agent increases (resp., decreases) her estimate of the idiosyncratic component. That is, the correlation between dividends and the estimated idiosyncratic component is perceived by the investor to be positive. Since dividend shocks are priced, the idiosyncratic component generates a higher risk premium and lower valuations for the firm when the signal precision is lower. Furthermore, the positive correlation between dividends and the estimated idiosyncratic component implies that a negative dividend shock today yields a decrease in expected future dividends, thereby generating more risk in the long term than in the short term. In equilibrium, this upward-sloping effect dominates and yields an increasing term structure of risk premium for the firm when the signal precision is low.

To summarize, information processing and learning provide an economic mechanism that

helps explain the observed value premium (Fama and French, 1992), downward-sloping term structure of risk premium for growth firms (Giglio et al., 2020), upward-sloping term structure of risk premium for value firms (Giglio et al., 2020), downward-sloping term structure of market risk premium (van Binsbergen et al., 2012), and upward-sloping term structure of interest rates (Gürkaynak, Sack, and Wright, 2010).

This paper is related to the literature on the value premium. In Zhang (2005), costly investment reversibility represents a form of operating inflexibility that rationalizes the value premium. Yogo (2006) shows that utility derived from both nondurable and durable consumption goods can explain the value premium. In Bansal, Dittmar, and Lundblad (2005) and Hansen, Heaton, and Li (2008), the value premium results from the value firms' higher dividend exposure to aggregate expected growth relative to growth firms. Santos and Veronesi (2006) rationalize the value premium via the existence of fluctuations in both wages and dividends. Lettau and Wachter (2007, 2011) show that the value premium can be explained by the duration of assets. In their model, value and growth stocks are short- and long-horizon equity, whose cashflows concentrate in the short- and long-term, respectively. Since expected dividends are negatively correlated to dividends, growth stocks represent a better hedge and therefore feature a lower risk premium than value stocks. In Kogan and Papanicolaou (2014), the value premium is explained by the firms' different sensitivities to investment-specific technology shocks. However, the recent empirical findings of Golubov and Konstantinidi (2019) empirically question whether the value premium can be explained by operating inflexibility, cashflow risk, asset duration (see also Chen, 2017), and exposure to investment-specific technology shocks. They conclude that the value premium remains a puzzle.

We contribute to this literature by providing an alternative economic channel for the value premium, namely information processing and learning. In our model, more accurate information on a firm's expected dividend growth rate yields lower risk, and therefore higher valuations and a lower risk premium. That is, the learning channel implies that growth

firms have a lower risk premium than value firms. The theoretical prediction of the model is consistent with empirical findings. Indeed, [Brennan and Subrahmanyam \(1995\)](#), [Botosan \(1997\)](#), [Brennan and Tamarowski \(2000\)](#), and [Doukas, Kim, and Pantzalis \(2005\)](#) show that high information disclosure or high analyst coverage is associated with high valuations (low dividend yields) and low expected returns.

Our paper is also related to the literature providing theoretical foundations for the empirical finding that short-term claims on market dividends feature particularly high risk premia ([van Binsbergen et al., 2012](#); [van Binsbergen, Hueskes, Koijen, and Vrugt, 2013](#); [van Binsbergen and Koijen, 2017](#); [Giglio et al., 2020](#)). [Belo, Collin-Dufresne, and Goldstein \(2015\)](#) shows that, when dividend dynamics are such that leverage ratios are stationary, the term structure of market risk premium is downward sloping. In [Croce, Lettau, and Ludvigson \(2015\)](#), investors' limited information and bounded rationality generate a high market risk premium together with a decreasing term structure of market risk premium. [Hasler and Marfè \(2016\)](#) show that the observed term structures of market risk premium can be explained by the existence of recoveries following rare economic disasters. [Marfè \(2017\)](#) rationalizes the decreasing term structure of market risk premium by investigating the impact of labor on dividend payouts. [Ai, Croce, Diercks, and Li \(2018\)](#) show that, when investment responds positively to contemporaneous productivity shocks and negatively to news about long-term productivity shocks, the term structure of market risk premium is downward sloping. [Hasler and Khapko \(2018\)](#) derive parameter conditions under which a simple model with time-varying expected economic growth and a representative agent with standard preferences can produce a decreasing term structure of market risk premium.

Our results contribute to this literature by showing that information processing and rational learning represent a natural economic channel that helps explain sizable risk premia on short-term market dividend claims together with the observed increasing term structure of interest rates, increasing term structure of risk premium for value firms, and decreasing term structure of risk premium for growth firms. The model generates the aforementioned

term structures while keeping the levels of asset risk premia, the value premium, and the risk-free in line with the data.

The remainder of the paper is organized as follows. Section 2 describes the economic fundamentals and the learning problem; Section 3 solves for equilibrium asset prices; Section 4 presents the results and Section 5 concludes. Derivations and supplementary material are provided in the Appendix.

2 Economic Fundamentals

This section describes the economy and discusses the implications of learning on the term structure of output growth risk. The only information available to the representative agent is the one generated by the aggregate output process (which is equal to the aggregate consumption in equilibrium). Importantly, the latent factors driving the expected growth dynamics are not observable. This introduces learning into the decision problem of the agent and has implications for the agent's perception of risk at different horizons.

2.1 Output Dynamics

The aggregate output, C , has the following dynamics

$$d \log C_t = dy_t + dz_t, \tag{1}$$

where y is an integrated process with time-varying expected growth, x . The dynamics of y and x are

$$dy_t = (\mu + x_t)dt + \sigma_y dB_{y,t}, \tag{2}$$

$$dx_t = -\lambda_x x_t dt + \sigma_x dB_{x,t}, \tag{3}$$

and the mean-reverting process z satisfies

$$dz_t = -\lambda_z z_t dt + \sigma_z dB_{z,t}. \quad (4)$$

The Brownian motions B_y, B_x , and B_z are mutually independent, and the parameters $\mu, \lambda_x, \lambda_z, \sigma_y, \sigma_x, \sigma_z$ are known constants. The full filtration generated by observing all three Brownian shocks is denoted by $\mathbf{F}$. Importantly, the system (1)–(4) is only partially observable by the agent who does not have access to the full information contained in the filtration $\mathbf{F}$. The flow of information available to the agent is the one generated by continuously observing the level of output C . The fundamentals x and z driving the expected output growth rate are unobservable. Therefore, they need to be estimated through Bayesian learning, as detailed in Section 2.2.

The aim of considering the aggregate output dynamics as modeled in (1), is to introduce flexibility in modeling the timing of output growth risk (Marfè, 2017). The first output component, y , is an integrated process which depends on the time integral of x . Shocks in x accumulate and therefore permanently affect future output levels. For this reason, we call x the permanent component. The second output component depends on the current value of the process z . Since the process z is mean-reverting, shocks in z dissipate as time passes and therefore have a transitory impact on output. For this reason, we call z the transitory component. Without the transitory component, the aggregate output follows the standard dynamics considered in the literature on incomplete information and learning (e.g. Gennotte, 1986; Detemple, 1986) as well as in the long-run risk literature pioneered by Bansal and Yaron (2004). In this case, as we will show, the term structure of output growth risk is upward sloping because of the accumulation of x shocks in y . Adding the transitory component z generates risk in the short term that dissipates in the longer term. Consequently, the transitory component induces a downward-sloping effect on the term structure of output growth risk. The existence of both x and z therefore provides flexibility in the modeling of

the timing of output growth risk.

Equations (1), (2), and (4) imply the following dynamics for the logarithm of output

$$d \log C_t = (\mu + x_t - \lambda_z z_t) dt + \sqrt{v} dB_t, \quad (5)$$

where $v \equiv \sigma_y^2 + \sigma_z^2$ is the instantaneous variance and $dB_t \equiv (\sigma_y dB_{y,t} + \sigma_z dB_{z,t})/\sqrt{v}$ is the increment of a Brownian motion.

2.2 Bayesian Learning

The expected growth rate of output varies over time due to shocks that come from two sources: the permanent component x defined in (3) and the transitory component z defined in (4). The agent has access to information generated by the observation of the realized aggregate output path in (5), and does not have access to the full information contained in the filtration $\mathbf{F}$. Therefore, all her actions must be adapted to her observation filtration $\mathbf{F}^o = \{\mathcal{F}_t^o\}_{t \geq 0}$. In other words, the agent needs to filter out through Bayesian updating the unobservable components x and z by observing the history of output. Proposition 1 provides the dynamics of the filtered state variables.

Proposition 1. *With respect to the agent's observation filtration, the dynamics of output C_t , and the filtered state variables $\hat{x}_t \equiv \mathbb{E}[x_t | \mathcal{F}_t^o]$ and $\hat{z}_t \equiv \mathbb{E}[z_t | \mathcal{F}_t^o]$ satisfy*

$$d \log C_t = (\mu + \hat{x}_t - \lambda_z \hat{z}_t) dt + \sqrt{v} d\hat{B}_t, \quad (6)$$

$$d\hat{x}_t = -\lambda_x \hat{x}_t dt + \hat{\sigma}_{x,t} d\hat{B}_t, \quad (7)$$

$$d\hat{z}_t = -\lambda_z \hat{z}_t dt + \hat{\sigma}_{z,t} d\hat{B}_t. \quad (8)$$

where $\hat{B}_t$ is an $\mathcal{F}_t^o$ -Brownian motion, $\hat{\sigma}_{x,t} = \frac{\gamma_{x,t} - \lambda_z \gamma_{xz,t}}{\sqrt{v}}$, and $\hat{\sigma}_{z,t} = \frac{\sigma_z^2 + \gamma_{xz,t} - \lambda_z \gamma_{z,t}}{\sqrt{v}}$. The

posterior variance-covariance matrix Γ_t is defined as follows:

$$\Gamma_t \equiv \begin{pmatrix} \gamma_{x,t} & \gamma_{xz,t} \\ \gamma_{xz,t} & \gamma_{z,t} \end{pmatrix} = \begin{pmatrix} \text{Var}[x_t | \mathcal{F}_t^o] & \text{Cov}[x_t, z_t | \mathcal{F}_t^o] \\ \text{Cov}[x_t, z_t | \mathcal{F}_t^o] & \text{Var}[z_t | \mathcal{F}_t^o] \end{pmatrix} \quad (9)$$

and its elements satisfy

$$\frac{d\gamma_{x,t}}{dt} = \sigma_x^2 - 2\lambda_x \gamma_{x,t} - v^{-1} (\gamma_{x,t} - \lambda_z \gamma_{xz,t})^2, \quad (10)$$

$$\frac{d\gamma_{z,t}}{dt} = \sigma_z^2 - 2\lambda_z \gamma_{z,t} - v^{-1} (\sigma_z^2 - \lambda_z \gamma_{z,t} + \gamma_{xz,t})^2, \quad (11)$$

$$\frac{d\gamma_{xz,t}}{dt} = -(\lambda_x + \lambda_z) \gamma_{xz,t} - v^{-1} (\gamma_{x,t} - \lambda_z \gamma_{xz,t}) (\sigma_z^2 - \lambda_z \gamma_{z,t} + \gamma_{xz,t}). \quad (12)$$

Proof. See Appendix B.1.

Equation (6) gives the dynamics of log-output, $\log C_t$, projected on the observable filtration, while Equations (7) and (8) describe the agent's updating rule of the expectation of the latent state variables x_t , and z_t . We refer to $\hat{x}_t$ and $\hat{z}_t$ as the filter estimates. Equations (10), (11), and (12) provide the dynamics of the posterior variance-covariance matrix (9) and hence capture the evolution of uncertainty associated with the estimation of the unobserved components.

Note that the posterior variance-covariance matrix is a deterministic function of time. In line with the literature (e.g., Scheinkman and Xiong, 2003; Dumas, Kurshev, and Uppal, 2009), we replace Γ_t with its steady-state in the subsequent discussion (i.e., $\Gamma \equiv \lim_{t \rightarrow \infty} \Gamma_t$). That is, we assume that the agent has already observed a long enough history of output growth rates to reach the stationary variance estimate of the unobserved components. The corresponding steady-state volatilities of the two filter estimates are denoted by $\hat{\sigma}_x$ and $\hat{\sigma}_z$, which are characterized in Appendix B.1.

2.3 Timing of Output Growth Risk

The goal of this section is to study how output growth risk varies across different horizons. Understanding the risk characteristics of output growth is crucial to understanding the properties of expected asset returns, which are the main focus of this paper. To this end, we follow [Belo et al. \(2015\)](#) and [Marfè \(2016\)](#) and compute annualized measures of output growth volatility at horizon τ under full and partial information,

$$\sigma_C(t, \tau) = \sqrt{\frac{1}{\tau} \log \left(\frac{\mathbb{E}_t[C_{t+\tau}^2 | \mathcal{F}_t]}{\mathbb{E}_t[C_{t+\tau} | \mathcal{F}_t]^2} \right)}, \quad \widehat{\sigma}_C(t, \tau) = \sqrt{\frac{1}{\tau} \log \left(\frac{\mathbb{E}_t[C_{t+\tau}^2 | \mathcal{F}_t^o]}{\mathbb{E}_t[C_{t+\tau} | \mathcal{F}_t^o]^2} \right)}. \quad (13)$$

To provide a clear intuition, this section focuses on the simplified models with either a permanent shock only or a transitory shock only. First, consider the economy in which output is an integrated process with its drift driven by the permanent component x only. That is, $d \log C_t = dy_t$, where $dy_t = (\mu + x_t)dt + \sigma_y dB_{y,t}$. In this case the term structure of output growth risk is monotone increasing, both under full and partial information (see [Proposition 2](#)). This result obtains because fluctuations in the permanent component accumulate over time and contribute to the integrated path of output. Thus, the longer the horizon, the larger the accumulated variation of the permanent component, and so the larger the variance of output. Moreover, [Proposition 2](#) shows that a more volatile (i.e., larger σ_x) or more persistent (i.e., smaller λ_x) permanent component leads to a higher level of the term structure of growth risk.

Proposition 2. *In the model with permanent shocks only, the output growth volatilities under full and partial information, $\sigma_C(t, \tau)$ and $\widehat{\sigma}_C(t, \tau)$ as given in (13), (i) increase with the horizon τ , (ii) decrease with the mean-reversion speed λ_x , and (iii) increase with the volatility σ_x .*

Proof. See [Appendix B.2](#).

Consider now the economy in which the drift of output is only driven by the transitory

component z . That is, $d \log C_t = dy_t + dz_t$, where $dy_t = \mu dt + \sigma_y dB_{y,t}$. In this case the term structure of growth risk is monotone decreasing, both under full and partial information (see Proposition 3). The reason is that shocks to the transitory component affect output in a transitory way; as time passes, the impact of these shocks on output weakens. Therefore, risk is higher in the short term than in the long term. A more volatile (i.e., larger σ_z) or more persistent (i.e., smaller λ_z) transitory component yields a higher level of the term structure of growth risk.

Proposition 3. *In the model with transitory shocks only, the output growth volatilities under full and partial information, $\sigma_C(t, \tau)$ and $\widehat{\sigma}_C(t, \tau)$ as given in (13), (i) decrease with the horizon τ , (ii) decrease with the mean-reversion speed λ_z , and (iii) increase with the volatility σ_z .*

Proof. See Appendix B.3.

Importantly, in these two simple economies with either permanent shocks only or transitory shocks only, the term structures of growth risk under full and partial information are not equal. They are increasing (resp., decreasing) and share the same short-run and long-run limits. However, the output growth variance perceived by the agent under partial information is larger than that under full information in both economies (see Proposition 4).

Proposition 4. *In the economy with permanent shocks only, the short-end and long-end limits of the output growth variance are*

$$\begin{aligned} \lim_{\tau \rightarrow 0} \sigma_C^2(t, \tau) &= \lim_{\tau \rightarrow 0} \widehat{\sigma}_C^2(t, \tau) = \sigma_y^2, \\ \lim_{\tau \rightarrow \infty} \sigma_C^2(t, \tau) &= \lim_{\tau \rightarrow \infty} \widehat{\sigma}_C^2(t, \tau) = \sigma_y^2 + \frac{\sigma_x^2}{\lambda_x^2}. \end{aligned} \tag{14}$$

In the economy with transitory shocks only, these limits are given by

$$\begin{aligned} \lim_{\tau \rightarrow 0} \sigma_C^2(t, \tau) &= \lim_{\tau \rightarrow 0} \widehat{\sigma}_C^2(t, \tau) = \sigma_y^2 + \sigma_z^2, \\ \lim_{\tau \rightarrow \infty} \sigma_C^2(t, \tau) &= \lim_{\tau \rightarrow \infty} \widehat{\sigma}_C^2(t, \tau) = \sigma_y^2. \end{aligned} \tag{15}$$

In both economies, the risk perceived under partial information is higher than that under full information: $\hat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau) > 0$ for any finite horizon $\tau > 0$. Moreover, the difference between the partial and full information term structures is a hump-shaped function of the horizon:

$$\partial_\tau (\hat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau)) \begin{cases} > 0 & \tau < \bar{\tau}_l, \\ < 0 & \tau > \bar{\tau}_l, \end{cases} \quad (16)$$

where $\bar{\tau}_l \equiv -\frac{1}{2\lambda_l} \left(1 + 2\mathcal{L}(-1, -\frac{1}{2\sqrt{e}})\right) > 0$, $l \in \{x, z\}$ and $\mathcal{L}(k, \cdot)$ is the k -th solution of the Lambert W (or product logarithm) function.

Proof. See Appendix B.4.

Figure 1 illustrates the term structures of output growth risk in the model with either permanent shocks only (left panel) or transitory shocks only (right panel). Under partial information the agent observes the level of output only, which implies that uncertainty is generated by a unique Brownian motion. As a result, the filtered variables are instantaneously perfectly correlated. Positive (negative) shocks to the level y are perceived to come together with positive (negative) shocks to the expected output growth component x (resp., z). Such positive correlation increases $\hat{\sigma}_C^2(t, \tau)$ relative to $\sigma_C^2(t, \tau)$ because y and x (or, resp., z) are uncorrelated under full information. The difference between the partial and full information term structures is a hump-shaped function of the horizon. It increases up to a threshold $\bar{\tau}_l$, $l \in \{x, z\}$ and decreases afterwards (see Proposition 4). At the horizon $\bar{\tau}_l$, the divergence between the agent's perception of growth risk under full and partial information is maximal. We note that the threshold $\bar{\tau}_l$ is decreasing in λ_l , $l \in \{x, z\}$. Consequently, the difference between the perceived and the true growth risks increases over a longer horizon when the persistence of x (or, resp., z) increases, or in other words, when its mean-reversion speed decreases.

When the aggregate output is driven by both permanent and transitory shocks, the term structure of growth risk can be either increasing or decreasing depending on which of the

Figure 1: Term Structure of Output Growth Risk in Simple Economies with Permanent or Transitory Shocks Only.

Parameter values are $\mu = 0.025$, $\sigma_y = 0.01$, $\lambda_x = 0.5$, $\sigma_x = 0.01$, $\lambda_z = 0.1$, $\sigma_z = 0.03$.

two shocks dominates. At the short end, the output growth risk is driven by the volatility of the transitory shock, σ_z . On the contrary, it is the volatility of the permanent shock (scaled by the mean-reversion speed), σ_x , that influences growth risk at the long end. In Section 4.1 we demonstrate that, in an economy with both permanent and transitory shocks, partial information and learning can alter the slope of the term structure of growth risk perceived by the agent.

3 Asset Pricing

In this section, we study the impact of learning on the equilibrium term structure of risk premia for value firms, growth firms, and the market, as well as on the term structure of interest rates. We show that partial information and learning help explain the upward-sloping term structure of risk premia for value firms, the downward-sloping term structure of risk premia for growth firms and the market, together with the upward-sloping term structure of interest rates.

We consider a pure-exchange economy (Lucas, 1978) in which the output process follows

the dynamics in (1). The representative agent features recursive preferences in the spirit of Kreps and Porteus (1979), Epstein and Zin (1989), Weil (1989), and Duffie and Epstein (1992). These preferences allow for the separation between the elasticity of intertemporal substitution and the coefficient of relative risk aversion. Given a consumption process C , the utility at time t is defined as

$$U_t \equiv \left[(1 - \delta^{\Delta t}) C_t^{\frac{1-\gamma}{\theta}} + \delta^{\Delta t} \mathbb{E}_t [U_{t+\Delta t}^{1-\gamma}]^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}},$$

where δ is the time discount factor, γ is the coefficient of risk aversion, ψ is the elasticity of intertemporal substitution, $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$, and Δt is a time interval. In equilibrium, the representative agent's aggregate consumption is equal to the output.

As in Abel (1999), dividends paid by the market are modeled as levered consumption

$$d \log D_{m,t} = (\mu_m + \phi(x_t - \lambda_z z_t)) dt + \phi \sqrt{v} dB_t, \quad (17)$$

where $\phi \geq 1$ is the market's leverage parameter on both expected growth and growth volatility. To investigate the effect of learning on the cross-section of asset returns, we consider firms featuring an idiosyncratic component in expected dividend growth. Dividends paid by firm i satisfy

$$d \log D_{i,t} = (\bar{\mu} + \bar{\phi}(x_t - \lambda_z z_t + x_{i,t})) dt + \phi \sqrt{v} dB_t, \quad (18)$$

where $\bar{\phi} \geq 1$ is firm i 's leverage parameter on expected growth. Each firm i 's dividend is affected by the same contemporaneous shock as the aggregate market dividend. Moreover, each firm i 's dividend has the same exposure to the permanent and transitory components of expected consumption growth, x and z . The only difference across firms comes from the realized paths of the idiosyncratic expected growth components x_i , which are unobservable. This unobservable firm-specific expected growth component is assumed to evolve according

to

$$dx_{i,t} = -\bar{\lambda}x_{i,t}dt + \bar{\sigma}dB_{x_i,t}, \quad (19)$$

where B_{x_i} is an independent Brownian motion and the parameters $\bar{\lambda}$ and $\bar{\sigma}$ are the same across firms. As in [Dumas et al. \(2009\)](#) and [Xiong and Yan \(2010\)](#), the agent observes a news signal on B_{x_i} satisfying

$$ds_{i,t} = p_i dB_{x_i,t} + \sqrt{1 - p_i^2} dB_{s_i,t}, \quad (20)$$

where p_i denotes the precision of the signal and B_{s_i} is an independent Brownian motion. The news signal can be interpreted as information provided by analysts on the firm's future earnings. The signal precision p_i is the only source of cross-sectional heterogeneity. As detailed in [Section 4.3](#), information processing and rational learning endogenously produce, as an equilibrium outcome, firms that can be identified as either growth (high valuations) or value (low valuations).

Note that the dividends introduced in [\(17\)](#) and [\(18\)](#) do not modify the learning problem on x and z exposed earlier in [Section 2.2](#). Indeed, observing dividends and consumption does not bring more information on x and z compared to observing consumption only because dividend and consumption shocks are identical. Moreover, $\bar{\phi} \neq \phi$ is required to insure that the process x_i is unobservable. If $\bar{\phi}$ were equal to ϕ , the difference between market dividends in [\(17\)](#) and firm i dividends in [\(18\)](#) would yield the observation of the process x_i .

When $\bar{\phi} \neq \phi$, the agent has to filter out the unobservable process x_i by observing both firm cashflows and the information signal. [Proposition 5](#) characterizes the dynamics of market and firm dividends under the agent's observation filtration.

Proposition 5. *Under the observation filtration, the dynamics of the market dividend D_m ,*

firm i 's dividend D_i , and filter $\hat{x}_{i,t} \equiv \mathbb{E}[x_{i,t} | \mathcal{F}_t^o]$ satisfy

$$\begin{aligned} d \log D_{m,t} &= (\mu_m + \phi(\hat{x}_t - \lambda_z \hat{z}_t)) dt + \phi \sqrt{v} d\hat{B}_t \\ d \log D_{i,t} &= (\bar{\mu} + \bar{\phi}(\hat{x}_t - \lambda_z \hat{z}_t + \hat{x}_{i,t})) dt + \phi \sqrt{v} d\hat{B}_t \\ d\hat{x}_{i,t} &= -\bar{\lambda} \hat{x}_{i,t} dt + \hat{\sigma}_{x_i} d\hat{B}_t + p_i \bar{\sigma} ds_{i,t}, \end{aligned} \tag{21}$$

where $\hat{x}$ and $\hat{z}$ are as before (see Proposition 1), $\hat{\sigma}_{x_i} \equiv \frac{\bar{\phi} \gamma_{x_i}}{\sqrt{v} |\phi - \bar{\phi}|}$, and γ_{x_i} is the steady-state posterior variance of $x_{i,t}$ provided in Equation (28), Appendix B.5.

Remark: The steady-state posterior variance γ_{x_i} is a decreasing function of the signal precision p_i . When the signal is perfectly informative ($p_i = 1$), the steady-state posterior variance γ_{x_i} is equal to zero.

Proof. See Appendix B.5.

The dynamics in (21) show that the agent updates the filter estimate $\hat{x}_i$ using two pieces of information. The first is the dividend growth surprise $d\hat{B}$, and the second is the news signal shock ds_i . When the precision of the signal is equal to zero, the loading on the signal is unsurprisingly equal to zero too. Therefore, the filter is perfectly correlated with the dividend when $p_i = 0$. As the precision of the signal increases, the loading on the news signal rises and that on the dividend surprise decreases. As a result, the correlation between the filter and the dividend gradually decreases as the precision of the signal increases. When the signal is perfectly informative ($p_i = 1$), the dividend and the filter become uncorrelated.

We now turn our attention to the asset pricing results. Recursive preferences lead to a non-affine state-price density. Therefore, to solve for prices and preserve analytic tractability, we follow the methodology presented by Eraker and Shaliastovich (2008), which is based on the Campbell and Shiller (1988) log-linearization.³

³Campbell, Lo, and MacKinlay (1997) and Bansal, Kiku, and Yaron (2012) show that the log-linearization yields low approximation errors in a Gaussian setting. Hasler and Marf (2016) document that it is also accurate when state variables follow jump-diffusions. For notational ease, we will treat the approximated results as exact in what follows.

Proposition 6 characterizes the state-price density, the risk-free rate, and the price of risk in the economy.

Proposition 6. *The equilibrium state-price density has dynamics given by*

$$\frac{dM_t}{M_t} = -r_t dt - \Lambda d\widehat{B}_t,$$

where the risk-free rate satisfies $r_t = r_0 + r_{\widehat{x}}\widehat{x}_t + r_{\widehat{z}}\widehat{z}_t$, with

$$r_0 = -\frac{1-\gamma}{1-1/\psi} \log \delta + \frac{1/\psi - \gamma}{1-1/\psi} \log k_1 + \gamma\mu - \frac{1}{2} \Theta(\widehat{\sigma}_y, \widehat{\sigma}_x, \widehat{\sigma}_z), \quad r_{\widehat{x}} = \frac{1}{\psi}, \quad \text{and} \quad r_{\widehat{z}} = -\frac{\lambda_z}{\psi},$$

and $\Theta(\widehat{\sigma}_y, \widehat{\sigma}_x, \widehat{\sigma}_z)$ defined in Appendix B.6.

The market price of risk satisfies

$$\Lambda = \gamma\sqrt{v} + \frac{\gamma - 1/\psi}{1/k_1 - (1 - \lambda_x)} \widehat{\sigma}_x - \frac{\lambda_z(\gamma - 1/\psi)}{1/k_1 - (1 - \lambda_z)} \widehat{\sigma}_z,$$

where $\widehat{\sigma}_x$ and $\widehat{\sigma}_z$ are defined in Appendix B.1, and $\widehat{\sigma}_y \equiv \sqrt{v} - \widehat{\sigma}_z$.

Proof. See Appendix B.6.

The risk-free rate increases with $\widehat{x}_t$ and decreases with $\widehat{z}_t$, irrespective of the value of the parameters. The reason is that an increase in $\widehat{x}_t$ (resp., $\widehat{z}_t$) implies an increase (resp., decrease) in expected consumption growth (see Equation (6)). Because the investor smoothes consumption over time, an increase in expected future consumption yields an increase in current consumption and therefore a decrease in risk-free investments. As a result, the risk-free rate increases when $\widehat{x}_t$ increases and when $\widehat{z}_t$ decreases.

The market price of risk is the sum of three components. If the investor has CRRA preferences, only the first component of the market price of risk remains. This component corresponds to the reward for bearing consumption growth risk (Lucas, 1978). Both the second and third components are rewards for bearing expected consumption growth risk

associated to variations in $\hat{x}_t$ and $\hat{z}_t$, respectively. Their signs are opposite because the expected consumption growth rate increases with $\hat{x}_t$ but decreases with $\hat{z}_t$ (see Equation (6)). When the investor has a preference for early resolution of uncertainty, the second component is positive (Bansal and Yaron, 2004) and therefore the third component is negative. When the investor has a preference for late resolution of uncertainty, these two risk components flip sign.

Proposition 7 characterizes the zero-coupon bond price and yield.

Proposition 7. *The price of the zero-coupon bond with time to maturity τ is given by*

$$Q(t, \tau) \equiv e^{-YTM(t, \tau)\tau} = e^{q_0(\tau) + q_{\hat{x}}(\tau)\hat{x}_t + q_{\hat{z}}(\tau)\hat{z}_t},$$

where $YTM(t, \tau)$ is the yield to maturity, $q_0(\tau)$ is derived in Appendix B.7, and

$$\begin{aligned} q_{\hat{x}}(\tau) &= -\frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}), \\ q_{\hat{z}}(\tau) &= \frac{1}{\psi} (1 - e^{-\lambda_z \tau}). \end{aligned}$$

Proof. See Appendix B.7.

The price of the zero-coupon bond decreases with $\hat{x}_t$ and increases with $\hat{z}_t$. The reason is as follows. An increase in $\hat{x}_t$ signals higher future consumption. For consumption smoothing purposes, the investor increases current consumption and therefore decreases her investment in the risk-free bond. As a result, the bond price drops. The exact opposite mechanism occurs when $\hat{z}_t$ increases, which explains why the bond price decreases with $\hat{x}_t$ and increases with $\hat{z}_t$.

Proposition 8 characterizes dividend strip prices and risk premia.

Proposition 8. *The price of a (market or firm) dividend strip with time to maturity τ is*

$$S_j(t, \tau) = D_{j,t} e^{w_{0j}(\tau) + w_{\hat{x}j}(\tau)\hat{x}_t + w_{\hat{z}j}(\tau)\hat{z}_t + w_{\hat{x}ij}(\tau)\hat{x}_{i,t}}, \quad j \in \{m, i\}$$

with

$$\begin{aligned}
w_{\hat{x}m}(\tau) &= \frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}) (\phi \psi - 1), & w_{\hat{x}i}(\tau) &= \frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}) (\bar{\phi} \psi - 1), \\
w_{\hat{z}m}(\tau) &= -\frac{1}{\psi} (1 - e^{-\lambda_z \tau}) (\phi \psi - 1), & w_{\hat{z}i}(\tau) &= -\frac{1}{\psi} (1 - e^{-\lambda_z \tau}) (\bar{\phi} \psi - 1), \\
w_{\hat{x}_i m}(\tau) &= 0, & w_{\hat{x}_i i}(\tau) &= \frac{\bar{\phi} (1 - e^{-\bar{\lambda} \tau})}{\bar{\lambda}},
\end{aligned}$$

and $w_{0j}(\tau)$ as defined in Appendix B.8. The corresponding risk premium is equal to

$$RP_j(t, \tau) = (\phi \sqrt{v} + w_{\hat{x}j}(\tau) \hat{\sigma}_x + w_{\hat{z}j}(\tau) \hat{\sigma}_z + w_{\hat{x}_i j}(\tau) \hat{\sigma}_{x_i}) \Lambda, \quad j \in \{m, i\}.$$

Proof. See Appendix B.8.

The market and firm stock prices can also be approximated by an exponential affine function of the state variables. Proposition 9 characterizes the price and risk premium of both the market and the firm.

Proposition 9. *The (market or firm) stock price is given by*

$$P_{j,t} = D_{j,t} e^{A_{0j} + A_{\hat{x}j} \hat{x}_t + A_{\hat{z}j} \hat{z}_t + A_{\hat{x}_i j} \hat{x}_{i,t}}, \quad j \in \{m, i\}$$

with

$$\begin{aligned}
A_{\hat{x}m} &= \frac{\phi - 1/\psi}{1 - k_{1m}(1 - \lambda_x)}, & A_{\hat{x}i} &= \frac{\bar{\phi} - 1/\psi}{1 - k_{1i}(1 - \lambda_x)}, \\
A_{\hat{z}m} &= -\frac{\lambda_z(\phi - 1/\psi)}{1 - k_{1m}(1 - \lambda_z)}, & A_{\hat{z}i} &= -\frac{\lambda_z(\bar{\phi} - 1/\psi)}{1 - k_{1i}(1 - \lambda_z)}, \\
A_{\hat{x}_i m} &= 0, & A_{\hat{x}_i i} &= \frac{\bar{\phi}}{1 - k_{1i}(1 - \bar{\lambda})},
\end{aligned}$$

and A_{0j} , k_{1j} as defined in Appendix B.9. The corresponding risk premium is

$$RP_j(t) = (\phi\sqrt{v} + A_{\hat{x}j}\hat{\sigma}_x + A_{\hat{z}j}\hat{\sigma}_z + A_{\hat{x}ij}\hat{\sigma}_{x_i}) \Lambda, \quad j \in \{m, i\}.$$

Proof. See Appendix B.9.

Propositions 8 and 9 show that the first component of the risk premium corresponds to the dividend growth risk, while the second, third, and fourth components correspond to the risks associated with expected dividend growth. The second and third components have opposite signs because expected dividend growth increases with $\hat{x}_t$ and decreases with $\hat{z}_t$ (see Equations (6) and (17)). When the product $\phi\psi$ (or $\bar{\phi}\psi$) is larger than one, the substitution effect dominates. An increase in expected dividend growth due to either an increase in $\hat{x}_t$ or a decrease in $\hat{z}_t$ signals improved investment opportunities, which pushes the investor to invest more in risky assets. As a result, the price of the strip and the price of the stock respond positively to an increase in $\hat{x}_t$ and negatively to an increase in $\hat{z}_t$. This translates into a positive second risk premium component and a negative third component when $\phi\psi > 1$ (or $\bar{\phi}\psi > 1$). When the product $\phi\psi$ (or $\bar{\phi}\psi$) is smaller than one, the income effect dominates. An increase in expected dividend growth due to either an increase in $\hat{x}_t$ or a decrease in $\hat{z}_t$ signals higher future consumption. Because of her consumption smoothing motives, the investor increases current consumption and therefore decreases investments in risky assets. As a result, the price of the strip and the price of the stock respond negatively to an increase in $\hat{x}_t$ and positively to an increase in $\hat{z}_t$, thereby explaining a negative second risk premium component and a positive third component when $\phi\psi < 1$ (or $\bar{\phi}\psi < 1$). The fourth component of the risk premium is positive for firm i and zero for the market because the process $\hat{x}_i$ drives firm i 's expected dividend growth only.

4 Results

In this section we calibrate the model, show that it is in line with observed data, and study its asset-pricing implications on the level and timing of both equity returns and interest rates.

4.1 Calibration and Model Fit

We calibrate the output dynamics parameters as follows. First, we match the short term (1-year) empirical level of volatility, which is 3%. Second, we match the empirical observation that output growth volatility is flat across horizons. Evidence of this empirical observation is provided in [Hasler and Marfè \(2016\)](#) and [Marfè \(2017\)](#), who document that the variance ratios of output growth rates in the U.S. are approximatively flat around unity. Also, [Dew-Becker \(2017\)](#) documents that robust estimators of long-run growth volatility are very close to estimates of the one-year volatility. Therefore, the output parameters $\Theta = \{\sigma_y, \sigma_x, \lambda_x, \sigma_z, \lambda_z\}$ are obtained by minimizing the following objective

$$\Theta^* = \arg \min \left\{ [\sigma_C(t, 1) - 0.03]^2 + \alpha \int_0^{50} [VR_C(t, \tau) - 1]^2 d\tau \right\},$$

where α is a weighting constant and $VR_C(t, \tau) = \frac{\sigma_C^2(t, \tau)}{\sigma_C^2(t, 1)}$ is the τ -year variance ratio. This minimization procedure yields: $\sigma_y = 0.0002$, $\sigma_x = 0.040$, $\lambda_x = 1.346$, $\sigma_z = 0.033$, and $\lambda_z = 0.549$.

In addition, $\mu = \mu_m = 0.025$ such that the long-run growth rate of market dividends equals that of output. The parameter ϕ captures the excessive volatility of market dividends relative to output. We set $\phi = 7.5$ to match the one-year volatility of shareholders' remuneration in the U.S., which is about 20% (see [Belo et al., 2015](#)). Therefore, we use the label dividend with slight abuse of terminology, as we actually consider the more appropriate full shareholders' remuneration consisting of dividends plus net repurchases. The parameters driving the firm's dividend are $\bar{\phi} = 7$, $\bar{\mu} = 0.0025$, $\bar{\lambda} = 0.3$, and $\bar{\sigma} = 0.01$. The precision of

Table 1: Model-Implied and Empirical Market Moments.

The first and second columns use real S&P 500 prices and dividends as well as real interest rates. Monthly data from 1871 to 2018 are obtained from Robert Shiller’s website. The first column uses pre- and post-war data, while the second column uses post-war data only. The last six columns provide the model-implied counterparts.

	Data		Preference Parameters					
	1871-2018	1946-2018	$\gamma = 7.5, \psi = 1.5$		$\gamma = 10, \psi = 1.5$		$\gamma = 7.5, \psi = 1.25$	
			Information		Information		Information	
			Full	Partial	Full	Partial	Full	Partial
Risk-free rate (%)	2.7	1.3	2.5	2.1	2.5	1.9	2.9	2.4
Dividend yield (%)	4.4	3.4	2.1	2.1	3.5	3.7	2.4	2.4
Risk premium (%)	3.8	5.6	4.1	5.2	5.6	6.9	4.0	5.2
Sharpe ratio (%)	26.8	45.9	20.4	22.8	27.7	30.1	20.0	22.7

the news signal p_i belongs to the interval $[0, 1]$ (see Section 4.3). The relative risk aversion is $\gamma = 7.5$, the elasticity of intertemporal substitution is $\psi = 1.5$, and the time discount factor is $\delta = 0.99$. The firm and preference parameters are in line with the long-run risk literature. As shown in what follows, these parameters yield model-implied levels of the risk-free rate, dividend yield, risk premium, and Sharpe ratio that are consistent with the data.

Table 1 compares the average risk-free rate, market dividend yield, risk premium, and Sharpe ratio implied by our calibrated model and their empirical counterparts. Under full information and using the benchmark preference parameters ($\gamma = 7.5, \psi = 1.5$), the risk-free rate, risk premium, and Sharpe ratio are 2.5%, 4.1%, and 20.4%, respectively. The risk premium and risk-free rate are respectively lower and higher than their post-war counterparts. As a result, the Sharpe ratio is lower than its empirical counterpart.

Under partial information, learning yields a lower risk-free rate (2.1%) and a higher risk premium (5.2%) than under full information. This corresponds to a decrease in the risk-free rate and an increase in the risk premium of about 25%. Both the equity premium and the risk-free rate are fairly close to their post-war counterparts under partial information, which shows that learning helps to solve the equity premium and risk-free rate puzzles.

An increase in risk aversion ($\gamma = 10, \psi = 1.5$) modifies the results as follows. The risk-

Figure 2: Term Structure of Output Growth Risk Calibrated to U.S. Data.

free rate decreases slightly and the risk premium increases substantially, which implies that the Sharpe ratio increases substantially too. This result obtains under both full and partial information. In addition, the dividend yields are fairly close to their empirical counterparts.

A decrease in the elasticity of intertemporal substitution ($\gamma = 7.5, \psi = 1.25$) has minor effects on the model-implied moments. The risk-free rate increases marginally, while the risk premium is unaffected. The reason is that a lower elasticity of intertemporal substitution reduces the compensation associated with the permanent shock x but increases the compensation associated with the transitory shock z ; the two effects offset each other.

Overall, the calibrated model delivers realistic asset pricing moments, and partial information yields a lower risk-free rate, a higher risk premium, and a higher Sharpe ratio than under full information.

Figure 2 depicts the calibrated model-implied term structure of volatility and variance ratio for the output growth rate.⁴ Under full information, the term structure of output growth variance ratio is flat because it was calibrated accordingly. This shows that the model dynamics are flexible enough to match the observed term structure of output growth

⁴Full and partial information variance ratios are respectively defined as

$$VR_C(t, \tau) = \frac{\sigma_C^2(t, \tau)}{\sigma_C^2(t, 1)} \quad \text{and} \quad \widehat{VR}_C(t, \tau) = \frac{\widehat{\sigma}_C^2(t, \tau)}{\widehat{\sigma}_C^2(t, 1)}.$$

variance ratio. Note that a flat term structure of variance ratios could have also been obtained with an i.i.d. process for the output growth. However, the presence of the latent variables x and z in our model is key to obtain a realistic level of risk premium. With our calibration, learning does not significantly affect the overall level of output growth volatility; the volatility lies in the interval 2.6-3.2%.

Importantly, learning alters the timing of output growth risk across different horizons. We observe in Figure 2 that the variance ratios of output growth are decreasing with the horizon under partial information. The reason is the following. The transitory process z is mean-reverting but not highly persistent under our calibration. This implies that the horizon at which the filtered volatility of z diverges the most from the true volatility is relatively short. As a result, long horizon variance ratios lie below unity under partial information.

To further evaluate the model fit, we compare in Appendix A the model-implied expected output growth rates inferred from observed asset prices to the forecasts obtained from the Survey of Professional Forecasters. The results show that the theoretical link between equilibrium outcomes and beliefs about expected growth finds support in the data.

4.2 Term Structures of Market Risk Premium and Bond Yields

This section compares the term structures of market risk premium and risk-free bond yields obtained under full and partial information.

The left panel of Figure 3 shows that under full information the risk premium for the market dividend strip with short maturity (e.g., 1-year) is particularly low. Risk premia increase sharply up to the 3-year maturity and then remain flat at a level of about 4.2%. As a consequence of learning, the behavior of the dividend strip risk premium is the opposite under partial information. The risk premium at short maturity (e.g., 1-year) is about 6.2%, which is high. Risk premia decrease uniformly up to the 10-year maturity and then remain flat at a level of about 5%. To summarize, the slope of the term structure of market risk premia switches from positive in the full information economy to negative in the partial

Figure 3: Term Structures of Market Risk Premium and Bond Yields.

information economy with learning. The reason for the switch in sign is detailed in the discussion below.

Since it is typically easier for investors to acquire accurate information in less uncertain times, our economies with full and partial information can be thought of as bounds of a single economy with time-varying uncertainty (Bloom, 2009, 2014; Jurado, Ludvigson, and Ng, 2015)—where the slope of the term structure of market risk premium switches sign over time (van Binsbergen et al., 2013).

The right panel of Figure 3 depicts the zero-coupon risk-free bond yields. The term structure of bond yields is downward-sloping under full information, in contrast with what is observed in the data. Partial information and learning generate two results which make the model predictions conform with the data. First, the short-term bond yield is lower than under full information. Second, the slope of the term structure is positive, consistent with TIPS data. Indeed, between 2003 and 2018 the average yields on TIPS were 0.6%, 0.9%, 1.1%, and 1.4% at the 5-year, 7-year, 10-year, and 20-year maturities, respectively. The term structure of bond yields becomes upward-sloping under partial information for the following reason. Bonds are used to hedged against consumption growth risk. Under partial information, the agent perceives high consumption growth risk in the short term and low

consumption growth risk in the long term (left panel of Figure 2). Therefore, the demand for short-term bonds is larger than that for long term bonds, thereby explaining why short-term yields are lower than long-term yields.

To better understand the mechanism of learning on the term structure of market risk premium, we study the differences between the pricing of risk in the economies with full and partial information. Under full information the risk premium on the market dividend strip with maturity τ is given by⁵

$$RP_m^{Full}(t, \tau) = \sum_{i=y,x,z} \Lambda_i \sigma_i w_i(\tau),$$

where $w_i(\tau) = \partial_i \log S(t, \tau)$, $i = \{y, x, z\}$. That is, the risk premium is a sum of price sensitivities ($w_i(\tau)$, the only terms depending on τ) weighted by the product of the fundamental volatilities (σ_i) and the corresponding prices of risk (Λ_i).

Under partial information the risk premium on the market dividend strip with maturity τ is given by

$$RP_m^{Partial}(t, \tau) = \Lambda \sum_{i=\hat{y}, \hat{x}, \hat{z}} \hat{\sigma}_i w_i(\tau),$$

where $w_i(\tau) = \partial_i \log S(t, \tau)$, $i = \{\hat{y}, \hat{x}, \hat{z}\}$. That is, the risk premium is the product of the unique price of risk in the economy (Λ) and a sum of price sensitivities ($w_i(\tau)$, the only terms depending on τ) weighted by the fundamental volatilities ($\hat{\sigma}_i$).

Note that $w_y(\tau)$, $w_x(\tau)$, and $w_z(\tau)$ are respectively constant, increasing, and decreasing with the maturity τ . The same holds for $w_{\hat{y}}(\tau)$, $w_{\hat{x}}(\tau)$, and $w_{\hat{z}}(\tau)$.

The positive slope of the term structure of risk premium under full information is due to the fact that the price of risk Λ_x , which rewards for bearing variations in x , is larger than the price of risk Λ_z , which rewards for bearing variations in z . Even though $\sigma_z w_z(\tau)$ is steeper than $\sigma_x w_x(\tau)$ (dashed lines on middle and right panels of Figure 4), the price of risk Λ_x is large enough to dominate the impact of the transitory shock z . Therefore, the

⁵Results pertaining to the full information model are provided in Appendix C.

Figure 4: Price Elasticities.

term structure of risk premium is upward-sloping. Note that the positive slope of the term structure of risk premium obtains in the full information case, although the model has been calibrated to match an empirically observed flat term structure of output growth risk.

Let us consider now the partial information case and the role of learning. While the processes x , y , and z have independent increments, learning leads to an endogenous correlation structure between the dynamics of the filtered variables $\hat{x}$, $\hat{y}$, and $\hat{z}$. More precisely, learning implies perfect correlations among these variables because the investor updates her beliefs by observing a single source of information (the history of output). In turn, the unique source of risk commands a unique price of risk Λ . This shows that learning neutralizes the role of the prices of risk in determining the shape of the term structure of risk premium. As a result, the slope of the term structure of risk premium is solely driven by the magnitude and steepness of the price elasticities. As the solid lines on the middle and right panels of Figure 4 show, the range of values taken by the price elasticity with respect to $\hat{z}$ (i.e., $\hat{\sigma}_z w_{\hat{z}}(\tau)$) is substantially larger than that taken by the price elasticity with respect to $\hat{x}$ (i.e., $\hat{\sigma}_x w_{\hat{x}}(\tau)$). Formally, we have

$$\lim_{\tau \rightarrow 0} \hat{\sigma}_z w_{\hat{z}}(\tau) - \lim_{\tau \rightarrow \infty} \hat{\sigma}_z w_{\hat{z}}(\tau) > \lim_{\tau \rightarrow \infty} \hat{\sigma}_x w_{\hat{x}}(\tau) - \lim_{\tau \rightarrow 0} \hat{\sigma}_x w_{\hat{x}}(\tau)$$

when $\hat{\sigma}_z > \hat{\sigma}_x / \lambda_x$ and $\phi > 1/\psi$. Therefore, learning leads to a downward-sloping term structure of market risk premium via its impact on both the quantity of risk across horizons

(driven by the perception of downward-sloping cashflow growth risk) and the price of risk (driven by the endogenous correlation structure among the filtered variables).

4.3 Term Structures of Value and Growth Firms Risk Premia

This section investigates the term structure of risk premium for value and growth firms. We find that the model-implied term structure of risk premium is upward-sloping for value firms and downward-sloping for growth firms, consistent with the empirical findings of [Giglio et al. \(2020\)](#).

Our goal is to highlight how information processing and learning determine the valuation of a firm, the level of its risk premium, and the shape of its term structure of risk premium in the cross-section. Since the news signal precision p_i is the only source of cross-sectional heterogeneity in our model, cross-sectional differences in the aforementioned equilibrium quantities depend exclusively on the level of p_i . That is, information processing and learning are the only determinants of the heterogeneity observed across firms in our model. Equivalently, there would be no heterogeneity across firms if the agent had full information on all economic shocks.

Figure 5 displays the dividend yield (i.e., the inverse of the price-dividend ratio) and the risk premium of firm i as a function of the news signal precision $p_i \in [0, 1]$. We observe that both the dividend yield and the risk premium monotonically decrease with the signal precision. Since dividend yields determine the value and growth categories in equilibrium models (see, e.g., [Lettau and Wachter, 2007, 2011](#)), our model shows that rational learning endogenously gives rise to these categories. Firms featuring low news signal precision can be thought of as value firms and firms featuring high news signal precision can be thought of as growth firms.

To investigate the economic mechanism, we consider the two limiting cases for the value firm V and the growth firm G . That is, firms V and G 's news signals have precisions $p_V = 0$ and $p_G = 1$, respectively. This means that the news signal s_V does not provide valuable

Figure 5: Cross-Section of Dividend Yield and Risk Premium.

information on firm V 's expected dividend growth rate shock B_{x_V} , whereas the news signal s_G is perfectly informative about firm G 's expected dividend growth rate shock B_{x_G} (see Equations (18), (19), and (20)).

The different signal precisions imply that the correlations between the dividend growth rate and the agent's expected dividend growth component, $\hat{x}_i$ with $i \in \{V, G\}$, are different across firms. Since firm V 's news signal is uninformative, the agent updates her beliefs about $\hat{x}_V$ using a single piece of information, namely the dividend growth surprise. As a result, the agent's expected dividend growth component $\hat{x}_V$ is perfectly correlated to firm V 's dividend growth rate (see Proposition 5). In contrast, the agent updates her beliefs about $\hat{x}_G$ using the news signal only because it is perfectly informative. Therefore, the expected dividend growth component $\hat{x}_G$ is uncorrelated to firm G 's dividend growth rate (see Proposition 5).

To understand how these different correlations affect the equilibrium returns of firms V and G , let us consider a negative dividend growth surprise $d\hat{B}$. This negative surprise implies that firm V 's expected growth component $\hat{x}_V$ is updated downward, whereas firm G 's expected growth component $\hat{x}_G$ is unaffected by this negative shock. That is, the surprise impacts both the current dividend and the expected dividend component of firm V , whereas it only affects the current dividend of firm G . As a result, the risk premium and dividend

Figure 6: Term Structures of Value and Growth Firms Risk Premium.

yield (inverse of the price-dividend ratio) of firm V are larger than those of firm G . Thus, firms V and G are value and growth firms, respectively.⁶

The model-implied dividend yield and risk premium of the value firm V are 2.2% and 9.9%, respectively. The dividend yield of growth firm G is 1.9%, and its risk premium is 5.1%.⁷ Therefore, the model-implied value premium is $RP_V - RP_G = 4.8\%$, in line with its empirical counterpart estimated at 4.9% (Lettau and Wachter, 2007).

In our model, the quality of information on a given firm determines whether the firm belongs to the value or growth category. Accurate information on firm G makes it a growth firm, whereas less accurate information on firm V makes it a value firm. This theoretical relationship between the accuracy of information and the firm's characteristic is consistent with empirical evidence. Indeed, the findings Brennan and Subrahmanyam (1995), Botosan (1997), Brennan and Tamarowski (2000), and Doukas et al. (2005) among others show that high information disclosure or high analyst coverage is associated with high valuations (low

⁶To clear the equity market, we implicitly assume that there exists a third firm N with stock price, $P_N = P_m - P_G - P_V$, paying a dividend, $D_N = D_m - D_G - D_V$. This firm's dividend yield lies in-between the dividend yields of firms V and G .

⁷Note that the market dividend yield is 2.1%, which lies in-between the dividend yields of firms V and G .

dividend yields) and low expected returns.

Figure 6 depicts the term structure of risk premium for the value firm V (left panel) and the growth firm G (right panel). The risk premium of the value firm increases with maturity, whereas that of the growth firm decreases with maturity. Our model of learning therefore provides a theoretical explanation for the empirical findings of Giglio et al. (2020).

The economic mechanism is as follows. When the news signal about x_i is perfectly informative ($p_i = 1$), the agent updates her estimate $\hat{x}_i$ using the information provided by the signal only (see Proposition 5; $\hat{\sigma}_{x_i} = 0$ when $p_i = 1$). Since the news signal is uncorrelated to consumption, signal shocks are not priced in equilibrium. Therefore, the firm's risk premium obtained with a perfectly informative news signal does not depend on the risk of the filter $\hat{x}_i$ (see Proposition 8; $\hat{\sigma}_{x_i} = 0$ when $p_i = 1$); It only depends on the risk of the filters $\hat{x}$ and $\hat{z}$, as does the risk premium of the market. As detailed in Section 4.2, the joint impact of the filters $\hat{x}$ and $\hat{z}$ yields a downward-sloping effect on the term structure of risk premium. This explains why growth firm G 's term structure of risk premium is downward sloping, as the term structure of the market risk premium. In contrast, when the news signal is uninformative ($p_i = 0$), the agent updates her filter $\hat{x}_i$ using the information provided by the dividend growth surprise only (see Proposition 5). That is, the filter $\hat{x}_i$ is perfectly correlated to consumption, and therefore the risk of $\hat{x}_i$ contributes to the firm's risk premium when the signal is uninformative (see Proposition 8). As a negative shock hits the firm's dividend, the expected future dividend growth component $\hat{x}_i$ is revised downward, which means that risk is more concentrated in the long run than in the short run. This upward-sloping effect implied by the filter $\hat{x}_i$ dominates the downward-sloping effect implied by $\hat{x}$ and $\hat{z}$, thereby explaining why value firm V 's term structure of risk premium is upward sloping.

5 Conclusion

This paper shows that rational learning about expected economic growth helps explain the observed shape of the term structure of risk premium for value firms, growth firms, and the market together with the observed term structure of interest rates. More accurate information on a firm's future cashflow means less risk, and therefore a higher valuation and a lower risk premium in equilibrium. That is, firms featuring accurate information are growth firms, whereas firms featuring less accurate information are value firms. Moreover, learning using accurate information implies that risk is low in the long term, whereas learning using less accurate information yields high risk in the long term. This implies that the term structure of risk premium is downward sloping for growth firms and upward sloping for value firms, as observed in the data ([Giglio et al., 2020](#)). Learning also yields a decreasing term structure of market risk premium and an increasing term structure of interest rates, whereas the economy with full information predicts the opposite shapes. That is, information processing and learning help explain the shape of the observed term structures of risk premium in the cross-section of equities together with the observed term structure of interest rates. Furthermore, these results do not deteriorate the model's ability to generate realistic levels for the value premium, market risk premium, and risk-free rate.

References

- Abel, A. B. (1999). Risk premia and term premia in general equilibrium. *Journal of Monetary Economics* 43(1), 3–33.
- Ai, H., M. M. Croce, A. M. Diercks, and K. Li (2018). News shocks and the production-based term structure of equity returns. *The Review of Financial Studies* 31(7), 2423–2467.
- Bansal, R., R. F. Dittmar, and C. T. Lundblad (2005, 08). Consumption, dividends, and the cross section of equity returns. *Journal of Finance* 60(4), 1639–1672.
- Bansal, R., D. Kiku, and A. Yaron (2012). An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review* 1, 183–221.
- Bansal, R., S. Miller, D. Song, and A. Yaron (2019). The term structure of equity risk premia. NBER Working Paper No. 25690.
- Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59(4), 1481–1509.
- Belo, F., P. Collin-Dufresne, and R. S. Goldstein (2015). Dividend dynamics and the term structure of dividend strips. *Journal of Finance* 70(3), 1115–1160.
- Berrada, T. (2006). Incomplete information, heterogeneity, and asset pricing. *Journal of Financial Econometrics* 4(1), 136–160.
- Berrada, T. (2009). Bounded rationality and asset pricing with intermediate consumption. *Review of Finance* 13(4), 693–725.
- Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica* 77(3), 623–685.
- Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives* 28(2), 153–76.
- Botosan, C. A. (1997). Disclosure level and the cost of equity capital. *The Accounting Review* 72(3), 323–349.
- Brennan, M. J. and A. Subrahmanyam (1995). Investment analysis and price formation in securities markets. *Journal of Financial Economics* 38(3), 361 – 381.
- Brennan, M. J. and C. Tamarowski (2000). Investor relations, liquidity, and stock prices. *Journal of Applied Corporate Finance* 12(4), 26–37.
- Brennan, M. J. and Y. Xia (2001). Stock price volatility and equity premium. *Journal of Monetary Economics* 47(2), 249 – 283.
- Campbell, J., A. Lo, and C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton University Press, Princeton, NJ.

- Campbell, J. Y. and R. J. Shiller (1988). Stock prices, earnings, and expected dividends. *Journal of Finance* 43(3), 661–676.
- Chen, H. J. (2017). Do cash flows of growth stocks really grow faster? *The Journal of Finance* 72(5), 2279–2330.
- Croce, M. M., M. Lettau, and S. C. Ludvigson (2015). Investor information, long-run risk, and the term structure of equity. *Review of Financial Studies* 28(3), 706–742.
- Detemple, J. B. (1986). Asset pricing in a production economy with incomplete information. *Journal of Finance* 41(2), 383–91.
- Detemple, J. B. (1991). Further results on asset pricing with incomplete information. *Journal of Economic Dynamics and Control* 15(3), 425–453.
- Dew-Becker, I. (2017). How risky is consumption in the long-run? Benchmark estimates from a robust estimator. *The Review of Financial Studies* 30(2), 631–666.
- Doukas, J. A., C. F. Kim, and C. Pantzalis (2005). The two faces of analyst coverage. *Financial Management* 34(2), 99–125.
- Duffie, D. and L. G. Epstein (1992). Stochastic differential utility. *Econometrica* 60(2), 353–94.
- Dumas, B., A. Kurshev, and R. Uppal (2009). Equilibrium portfolio strategies in the presence of sentiment risk and excess volatility. *The Journal of Finance* 64(2), 579–629.
- Epstein, L. G. and S. E. Zin (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: A theoretical framework. *Econometrica* 57(4), 937–69.
- Eraker, B. and I. Shaliastovich (2008). An equilibrium guide to designing affine pricing models. *Mathematical Finance* 18(4), 519–543.
- Fama, E. F. and K. R. French (1992). The cross-section of expected stock returns. *The Journal of Finance* 47(2), pp. 427–465.
- Gennotte, G. (1986). Optimal portfolio choice under incomplete information. *Journal of Finance* 41(3), 733–46.
- Giglio, S., B. Kelly, and S. Kozak (2020). Equity term structures without dividend strips data. *Working Paper, Yale*.
- Golubov, A. and T. Konstantinidi (2019). Where is the risk in value? Evidence from a market-to-book decomposition. *The Journal of Finance* 74(6), 3135–3186.
- Gürkaynak, R. S., B. Sack, and J. H. Wright (2010). The tips yield curve and inflation compensation. *American Economic Journal: Macroeconomics* 2(1), 70–92.
- Hansen, L. P., J. C. Heaton, and N. Li (2008). Consumption strikes back? measuring long-run risk. *Journal of Political Economy* 116(2), 260–302.

- Hasler, M. and M. Khapko (2018). Term structures of equity and interest rates in time-varying expected growth models. *Working Paper, University of Texas at Dallas and University of Toronto*.
- Hasler, M. and R. Marfè (2016). Disaster recovery and the term structure of dividend strips. *Journal of Financial Economics* 122(1), 116 – 134.
- Jegadeesh, N., J. Kim, S. D. Krische, and C. M. C. Lee (2004). Analyzing the analysts: When do recommendations add value? *The Journal of Finance* 59(3), 1083–1124.
- Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *The American Economic Review* 105(3), 1177–1216.
- Kogan, L. and D. Papanicolaou (2014). Growth opportunities, technology shocks, and asset prices. *The Journal of Finance* 69(2), 675–718.
- Kreps, D. M. and E. L. Porteus (1979). Temporal von neumann-morgenstern and induced preferences. *Journal of Economic Theory* 20(1), 81–109.
- Lettau, M. and J. A. Wachter (2007). Why is long-horizon equity less risky? A duration-based explanation of the value premium. *Journal of Finance* 62(1), 55–92.
- Lettau, M. and J. A. Wachter (2011). The term structures of equity and interest rates. *Journal of Financial Economics* 101(1), 90–113.
- Liptser, R. S. and A. N. Shiryaev (2001). *Statistics of Random Processes II*. Springer Verlag, New York.
- Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica* 46(6), 1429–1445.
- Marfè, R. (2016). Corporate fraction and the equilibrium term structure of equity risk. *Review of Finance* 20(2), 855–905.
- Marfè, R. (2017). Income insurance and the equilibrium term structure of equity. *The Journal of Finance* 72(5), 2073–2130.
- Santos, T. and P. Veronesi (2006). Labor income and predictable stock returns. *The Review of Financial Studies* 19(1), 1–44.
- Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *The Journal of Political Economy* 111(6), 1183–1219.
- van Binsbergen, J., M. Brandt, and R. Koijen (2012). On the timing and pricing of dividends. *American Economic Review* 102(4), 1596–1618.
- van Binsbergen, J., W. Hueskes, R. Koijen, and E. Vrugt (2013). Equity yields. *Journal of Financial Economics* 110(3), 503 – 519.
- van Binsbergen, J. H. and R. Koijen (2017). The term structure of returns: Facts and theory. *Journal of Financial Economics* 124(1), 1 – 21.

- Veronesi, P. (1999). Stock market overreaction to bad news in good times: A rational expectations equilibrium model. *Review of Financial Studies* 12(5), 975–1007.
- Veronesi, P. (2000). How does information quality affect stock returns? *Journal of Finance* 55(2), 807–837.
- Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics* 24(3), 401–421.
- Xiong, W. and H. Yan (2010). Heterogeneous expectations and bond markets. *Review of Financial Studies* 23(4), 1433–1466.
- Yogo, M. (2006). A consumption-based explanation of expected stock returns. *The Journal of Finance* 61(2), 539–580.
- Zhang, L. (2005). The value premium. *The Journal of Finance* 60(1), 67–103.

Appendix

A Model-Implied Growth Forecasts and Survey Data

We verify that the theoretical link between equilibrium outcomes and beliefs about expected growth finds support in the data. The model predicts that empirically observable equilibrium outcomes such as the risk-free rate and the market price-dividend ratio are functions of the agent's filter estimates, $\hat{x}$ and $\hat{z}$. At the same time, the filters $\hat{x}$ and $\hat{z}$ are the drivers of the agent's belief about expected growth:

$$\frac{1}{dt}\mathbb{E}[d\log C_t|\mathcal{F}_t^o] = \mu + \hat{x}_t - \lambda_z \hat{z}_t.$$

First, we use the time series of the risk-free rate and the S&P 500 log price-dividend ratio obtained from Robert Shiller's website to infer the model-implied time series of $\hat{x}$ and $\hat{z}$, as follows:⁸

$$\min_{\{\hat{x}_t, \hat{z}_t\}} |r_t^{\text{data}} - r_t^{\text{model}}|^2 + |pd_t^{\text{data}} - pd_t^{\text{model}}|^2, \quad \forall t.$$

Second, we construct the model-implied time series of expected growth (up to a constant), as follows:

$$\hat{g}_t^{\text{model}} = \hat{x}_t - \lambda_z \hat{z}_t.$$

Third, we consider several measures to proxy for agents' beliefs about expected growth. Namely, we consider the mean and median forecasts of real GDP growth, industrial production growth, and corporate profits growth. Quarterly data available from Q4:1968 to Q2:2015 are obtained from the Survey of Professional Forecasters. We then regress these proxies of beliefs about expected growth on the model-implied time series, as follows:

$$\hat{g}_t^{\text{survey}} = \alpha + \beta \hat{g}_t^{\text{model}} + \epsilon_t.$$

Whether or not the data lend support to the economic mechanism of the model depends on the sign and significance of the coefficient β . If β is positive and significant, then the way the model predicts that beliefs about growth shape equilibrium prices is consistent with empirical evidence.

Table 2 reports the results from the regression. All of the six empirical measures of beliefs about expected growth are positively and significantly related to the model-implied beliefs inferred from actual prices. The coefficient β is statistically different from zero at the 1% level in all of the six cases and the R^2 ranges from 6% to 10%. Thus, we recover in actual data the model mechanism through which beliefs about growth drive prices.

⁸The risk-free rate and the log price-dividend ratio are affine in $\hat{x}$ and $\hat{z}$. Therefore, by observing the former we can infer the latter by simply solving a linear system.

Table 2: Model-Implied Growth Forecasts and Survey Data.

This table reports the regression coefficients, t-statistic, and R^2 from the regressions of several measures of expected growth forecasts on the model-implied expected growth:

$$\hat{g}_t^{\text{survey}} = \alpha + \beta \hat{g}_t^{\text{model}} + \epsilon_t.$$

The expected growth forecasts are either the cross-sectional mean or median from the Survey of Professional Forecasters (SPF) about real GDP, industrial production, and corporate profits. The data are quarterly from 1968:Q4 to 2015:Q2. The model-implied time series of expected growth is computed as

$$\hat{g}_t^{\text{model}} = \hat{x}_t - \lambda_z \hat{z}_t,$$

where $\hat{x}_t$ and $\hat{z}_t$ are extracted by minimizing the distance between the time series of the risk-free rate and the S&P 500 log price-dividend ratio obtained from Robert Shiller's website and their model-implied counterparts.

Growth Forecasts:	Gross Domestic Product		Industrial Production		Corporate Profits	
	mean	median	mean	median	mean	median
constant	0.029***	0.029***	0.035***	0.034***	0.094***	0.085***
t-stat	(13.88)	(13.65)	(8.31)	(8.09)	(8.53)	(7.89)
slope	0.133***	0.131***	0.243***	0.229***	0.762***	0.742***
t-stat	(4.06)	(3.90)	(3.64)	(3.40)	(4.35)	(4.33)
R^2 (%)	8.18	7.61	6.69	5.87	9.27	9.10

B Proofs

B.1 Proposition 1

Proof. The filtering equations follow from standard results in non-linear filtering theory, see [Liptser and Shiryaev \(2001\)](#). The steady-state volatilities $\hat{\sigma}_x$ and $\hat{\sigma}_z$ satisfy

$$\hat{\sigma}_x = \frac{\gamma_x - \lambda_z \gamma_{xz}}{\sqrt{v}}, \quad \hat{\sigma}_z = \frac{\sigma_z^2 + \gamma_{xz} - \lambda_z \gamma_z}{\sqrt{v}},$$

where the steady-state posterior variances γ_x and γ_z , and the steady-state posterior covariance γ_{xz} solve the following system of equations

$$\begin{aligned} 0 &= \sigma_x^2 - 2\lambda_x \gamma_x - v^{-1} (\gamma_x - \lambda_z \gamma_{xz})^2, \\ 0 &= \sigma_z^2 - 2\lambda_z \gamma_z - v^{-1} (\sigma_z^2 - \lambda_z \gamma_z + \gamma_{xz})^2, \\ 0 &= -(\lambda_x + \lambda_z) \gamma_{xz} - v^{-1} (\gamma_x - \lambda_z \gamma_{xz}) (\sigma_z^2 - \lambda_z \gamma_z + \gamma_{xz}). \end{aligned}$$

□

B.2 Proposition 2

Proof. Using the moment generating function of consumption under the full information filtration and the definition of full information consumption volatility in (13), we can compute the annualized variance of consumption as

$$\sigma_C^2(t, \tau) = \sigma_y^2 + \frac{\sigma_x^2 e^{-2\lambda_x \tau} (e^{2\lambda_x \tau} (2\lambda_x \tau - 3) + 4e^{\lambda_x \tau} - 1)}{2\lambda_x^3 \tau}. \quad (22)$$

Partial derivatives with respect to the horizon τ , volatility σ_x , and the mean-reversion speed λ_x are as follows:

$$\begin{aligned} \frac{\partial \sigma_C^2(t, \tau)}{\partial \tau} &= \frac{\sigma_x^2 e^{-2\lambda_x \tau} (1 + 3e^{2\lambda_x \tau} + 2\lambda_x \tau - 4e^{\lambda_x \tau} (\lambda_x \tau + 1))}{2\lambda_x^3 \tau^2}, \\ \frac{\partial \sigma_C^2(t, \tau)}{\partial \sigma_x} &= -\frac{\sigma_x (3 + e^{-2\lambda_x \tau} - 4e^{-\lambda_x \tau} - 2\lambda_x \tau)}{\lambda_x^3 \tau}, \\ \frac{\partial \sigma_C^2(t, \tau)}{\partial \lambda_x} &= \frac{\sigma_x^2 e^{-2\lambda_x \tau} (2\lambda_x \tau + e^{2\lambda_x \tau} (9 - 4\lambda_x \tau) - 4e^{\lambda_x \tau} (\lambda_x \tau + 3) + 3)}{2\lambda_x^4 \tau}. \end{aligned}$$

For $\tau > 0$ the following holds: $\frac{\partial \sigma_C^2(t, \tau)}{\partial \tau} > 0$, $\frac{\partial \sigma_C^2(t, \tau)}{\partial \sigma_x} > 0$, and $\frac{\partial \sigma_C^2(t, \tau)}{\partial \lambda_x} < 0$. The first inequality holds since

$$\begin{aligned} &1 + 3e^{2\lambda_x \tau} + 2\lambda_x \tau - 4e^{\lambda_x \tau} (\lambda_x \tau + 1) \\ &= 2(e^{\lambda_x \tau} - 1)(e^{\lambda_x \tau} - 2\lambda_x \tau - 1) + (e^{2\lambda_x \tau} - 2\lambda_x \tau - 1) \\ &> (e^{\lambda_x \tau} - 1)(e^{\lambda_x \tau} - 2\lambda_x \tau - 1) + (e^{2\lambda_x \tau} - 2\lambda_x \tau - 1) \\ &= 2e^{\lambda_x \tau} (e^{\lambda_x \tau} - \lambda_x \tau - 1) > 0 \text{ for } \tau > 0. \end{aligned}$$

The second inequality holds since

$$\begin{aligned} &e^{2\lambda_x \tau} (3 + e^{-2\lambda_x \tau} - 4e^{-\lambda_x \tau} - 2\lambda_x \tau) \\ &= 1 - 4e^{\lambda_x \tau} + e^{2\lambda_x \tau} (2 - 3\lambda_x \tau) \\ &< 1 - 4e^{\lambda_x \tau} + e^{\lambda_x \tau} (2 - 3\lambda_x \tau) \\ &= 1 - e^{\lambda_x \tau} (1 + \lambda_x \tau) < 0 \text{ for } \tau > 0. \end{aligned}$$

The third inequality, $\frac{\partial \sigma_C^2(t, \tau)}{\partial \lambda_x} < 0$, holds since $3 + 2\lambda_x \tau + e^{2\lambda_x \tau} (9 - 4\lambda_x \tau) - 4e^{\lambda_x \tau} (\lambda_x \tau + 3) < 0$.

Similarly, using the moment generating function of consumption under the partial information filtration and the definition of partial information consumption volatility in (13), we

can compute the agent's estimate of the annualized variance of consumption as

$$\begin{aligned}\widehat{\sigma}_C^2(t, \tau) &= \frac{e^{-2\lambda_x\tau}}{2\lambda_x^3\tau} \left(2\lambda_x\sigma_y \left((e^{\lambda_x\tau} - 1)^2 \sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} + \lambda_x^2\sigma_y\tau e^{2\lambda_x\tau} - \lambda_x\sigma_y (e^{\lambda_x\tau} - 1)^2 \right) \right. \\ &\quad \left. + \sigma_x^2 (e^{2\lambda_x\tau}(2\lambda_x\tau - 3) + 4e^{\lambda_x\tau} - 1) \right).\end{aligned}\quad (23)$$

Partial derivatives with respect to the horizon τ , volatility σ_x , and the mean-reversion speed λ_x are as follows:

$$\begin{aligned}\frac{\partial \widehat{\sigma}_C^2(t, \tau)}{\partial \tau} &= \frac{e^{-2\lambda_x\tau}}{2\lambda_x^3\tau^2} \left(2\lambda_x\sigma_y (e^{\lambda_x\tau} - 1) (1 + 2\lambda_x\tau - e^{\lambda_x\tau}) \left(\sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} - \lambda_x\sigma_y \right) \right. \\ &\quad \left. + \sigma_x^2 (1 + 3e^{2\lambda_x\tau} + 2\lambda_x\tau - 4e^{\lambda_x\tau}(\lambda_x\tau + 1)) \right), \\ \frac{\partial \widehat{\sigma}_C^2(t, \tau)}{\partial \sigma_x} &= \frac{e^{-2\lambda_x\tau}\sigma_x}{\lambda_x^3\tau} \left(\frac{\lambda_x\sigma_y (e^{\lambda_x\tau} - 1)^2}{\sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2}} + e^{\lambda_x\tau} (4 + e^{\lambda_x\tau}(2\lambda_x\tau - 3)) - 1 \right), \\ \frac{\partial \widehat{\sigma}_C^2(t, \tau)}{\partial \lambda_x} &= \frac{e^{-2\lambda_x\tau}}{2\lambda_x^4\tau\sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2}} \left(4\lambda_x^4\sigma_y^3\tau (e^{\lambda_x\tau} - 1) + 3\sigma_x^2 (-4e^{\lambda_x\tau} + 3e^{2\lambda_x\tau} + 1) \sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} \right. \\ &\quad + 2\lambda_x^2\sigma_y (e^{\lambda_x\tau} - 1) \left(\sigma_y (e^{\lambda_x\tau} - 1) \sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} + 2\sigma_x^2\tau \right) \\ &\quad - 2\lambda_x\sigma_x^2 \left(\tau (2e^{\lambda_x\tau} (e^{\lambda_x\tau} + 1) - 1) \sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} + 2\sigma_y (e^{\lambda_x\tau} - 1)^2 \right) \\ &\quad \left. - 2\lambda_x^3\sigma_y^2 (e^{\lambda_x\tau} - 1) \left(2\tau\sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} + \sigma_y (e^{\lambda_x\tau} - 1) \right) \right).\end{aligned}$$

For $\tau > 0$ we have $\frac{\partial \widehat{\sigma}_C^2(t, \tau)}{\partial \tau} > 0$, $\frac{\partial \widehat{\sigma}_C^2(t, \tau)}{\partial \sigma_x} > 0$, and $\frac{\partial \widehat{\sigma}_C^2(t, \tau)}{\partial \lambda_x} < 0$. The first inequality holds since

$$\begin{aligned}& 2\lambda_x\sigma_y \left(\sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} - \lambda_x\sigma_y \right) (e^{\lambda_x\tau} - 1) (1 + 2\lambda_x\tau - e^{\lambda_x\tau}) \\ & + \sigma_x^2 (1 + 3e^{2\lambda_x\tau} + 2\lambda_x\tau - 4e^{\lambda_x\tau}(\lambda_x\tau + 1)) \\ & \geq \min \left\{ \lambda_x\sigma_y \left(\sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} - \lambda_x\sigma_y \right), \sigma_x^2 \right\} \left(2(e^{\lambda_x\tau} - 1) (1 + 2\lambda_x\tau - e^{\lambda_x\tau}) \right. \\ & \quad \left. + (1 + 3e^{2\lambda_x\tau} + 2\lambda_x\tau - 4e^{\lambda_x\tau}(\lambda_x\tau + 1)) \right) \\ & = \min \left\{ \lambda_x\sigma_y \left(\sqrt{\lambda_x^2\sigma_y^2 + \sigma_x^2} - \lambda_x\sigma_y \right), \sigma_x^2 \right\} \left(e^{2\lambda_x\tau} - 2\lambda_x\tau - 1 \right) > 0 \text{ for } \tau > 0,\end{aligned}$$

where we use the fact that for $0 < a < b$ and $y > 0$ we have $ax + by \geq \min \{a, b\} (x + y)$. The second inequality holds since $4 + e^{\lambda_x\tau}(2\lambda_x\tau - 3) > 0$. Finally, by a lengthy and tedious calculation, one can show that $\frac{\partial \widehat{\sigma}_C^2(t, \tau)}{\partial \lambda_x} < 0$.

□

B.3 Proposition 3

Proof. Using the moment generating function of consumption under the full information filtration and the definition of consumption volatility in (13), we can compute the annualized

variance of consumption as

$$\sigma_C^2(t, \tau) = \sigma_y^2 + \frac{\sigma_z^2 (1 - e^{-2\lambda_z \tau})}{2\lambda_z \tau}. \quad (24)$$

Partial derivatives with respect to the horizon τ , volatility σ_z , and the mean-reversion speed λ_z are as follows:

$$\begin{aligned} \frac{\partial \sigma_C^2(t, \tau)}{\partial \tau} &= -\frac{\sigma_z^2 e^{-2\lambda_z \tau} (e^{2\lambda_z \tau} - 1 - 2\lambda_z \tau)}{2\lambda_z \tau^2}, \\ \frac{\partial \sigma_C^2(t, \tau)}{\partial \sigma_z} &= \frac{\sigma_z (1 - e^{-2\lambda_z \tau})}{\lambda_z \tau} \\ \frac{\partial \sigma_C^2(t, \tau)}{\partial \lambda_z} &= -\frac{\sigma_z^2 e^{-2\lambda_z \tau} (e^{2\lambda_z \tau} - 1 - 2\lambda_z \tau)}{2\lambda_z^2 \tau}. \end{aligned}$$

For $\tau > 0$ we have $\frac{\partial \sigma_C^2(t, \tau)}{\partial \tau} < 0$, $\frac{\partial \sigma_C^2(t, \tau)}{\partial \sigma_z} > 0$, and $\frac{\partial \sigma_C^2(t, \tau)}{\partial \lambda_z} < 0$.

Similarly, using the moment generating function of consumption under the partial information filtration, we can compute the agent's estimate of the annualized variance of consumption as

$$\begin{aligned} \hat{\sigma}_C^2(t, \tau) &= \frac{e^{-2\lambda_z \tau}}{2\lambda_z \tau} \left((e^{\lambda_z \tau} - 1) \left(2\sqrt{\sigma_y^2 (\sigma_y^2 + \sigma_z^2)} (e^{\lambda_z \tau} - 1) + \sigma_z^2 (e^{\lambda_z \tau} + 1) \right) \right. \\ &\quad \left. + 2\sigma_y^2 (e^{2\lambda_z \tau} (\lambda_z \tau - 1) + 2e^{\lambda_z \tau} - 1) \right). \end{aligned} \quad (25)$$

Partial derivatives with respect to the horizon τ , volatility σ_z , and the mean-reversion speed λ_z are as follows:

$$\begin{aligned} \frac{\partial \hat{\sigma}_C^2(t, \tau)}{\partial \tau} &= -\frac{e^{-2\lambda_z \tau}}{2\lambda_z \tau^2} \left(2\sigma_y \left(\sqrt{\sigma_y^2 + \sigma_z^2} - \sigma_y \right) (e^{\lambda_z \tau} - 1) (e^{\lambda_z \tau} - 2\lambda_z \tau - 1) \right. \\ &\quad \left. + \sigma_z^2 (e^{2\lambda_z \tau} - 2\lambda_z \tau - 1) \right), \\ \frac{\partial \hat{\sigma}_C^2(t, \tau)}{\partial \sigma_z} &= \frac{\sigma_z e^{-2\lambda_z \tau}}{\lambda_z \tau} (e^{\lambda_z \tau} - 1) \left(\frac{\sigma_y^2 (e^{\lambda_z \tau} - 1)}{\sqrt{\sigma_y^2 (\sigma_y^2 + \sigma_z^2)}} + e^{\lambda_z \tau} + 1 \right), \\ \frac{\partial \hat{\sigma}_C^2(t, \tau)}{\partial \lambda_z} &= -\frac{e^{-2\lambda_z \tau}}{2\lambda_z^2 \tau} \left(2\sigma_y \left(\sqrt{\sigma_y^2 + \sigma_z^2} - \sigma_y \right) (e^{\lambda_z \tau} - 1) (e^{\lambda_z \tau} - 2\lambda_z \tau - 1) \right. \\ &\quad \left. + \sigma_z^2 (e^{2\lambda_z \tau} - 2\lambda_z \tau - 1) \right). \end{aligned}$$

For $\tau > 0$ we have $\frac{\partial \hat{\sigma}_C^2(t, \tau)}{\partial \tau} < 0$, $\frac{\partial \hat{\sigma}_C^2(t, \tau)}{\partial \sigma_z} > 0$, and $\frac{\partial \hat{\sigma}_C^2(t, \tau)}{\partial \lambda_z} < 0$. The second inequality is

obvious, the first and last inequalities follow since

$$\begin{aligned}
& 2\sigma_y \left(\sqrt{\sigma_y^2 + \sigma_z^2} - \sigma_y \right) (e^{\lambda_z \tau} - 1) (e^{\lambda_z \tau} - 2\lambda_z \tau - 1) + \sigma_z^2 (e^{2\lambda_z \tau} - 2\lambda_z \tau - 1) \\
& \geq \min \left\{ 2\sigma_y \left(\sqrt{\sigma_y^2 + \sigma_z^2} - \sigma_y \right), \sigma_z^2 \right\} \left((e^{\lambda_z \tau} - 1) (e^{\lambda_z \tau} - 2\lambda_z \tau - 1) + (e^{2\lambda_z \tau} - 2\lambda_z \tau - 1) \right) \\
& = \min \left\{ 2\sigma_y \left(\sqrt{\sigma_y^2 + \sigma_z^2} - \sigma_y \right), \sigma_z^2 \right\} 2e^{\lambda_z \tau} (e^{\lambda_z \tau} - \lambda_z \tau - 1) > 0 \text{ for } \tau > 0,
\end{aligned}$$

where we use the fact that for $0 < a < b$ and $y > 0$ we have $ax + by \geq \min \{a, b\} (x + y)$.

□

B.4 Proposition 4

Proof. For the economy with only permanent shocks, taking the limits of (22) and (23) as horizon τ approaches zero or infinity gives the result in (14). Furthermore,

$$\hat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau) = \frac{e^{-2\lambda_x \tau} (e^{\lambda_x \tau} - 1)^2 \left(\sqrt{\sigma_y^2 (\lambda_x^2 \sigma_y^2 + \sigma_x^2)} - \lambda_x \sigma_y^2 \right)}{\lambda_x^2 \tau} > 0. \quad (26)$$

The derivative of the difference in (26) with respect to horizon is

$$\frac{\partial (\hat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau))}{\partial \tau} = \frac{e^{-2\lambda_x \tau} (e^{\lambda_x \tau} - 1) (1 + 2\lambda_x \tau - e^{\lambda_x \tau}) \left(\sqrt{\sigma_y^2 (\lambda_x^2 \sigma_y^2 + \sigma_x^2)} - \lambda_x \sigma_y^2 \right)}{\lambda_x^2 \tau^2}.$$

The sign of this derivative depends on the sign of $1 + 2\lambda_x \tau - e^{\lambda_x \tau}$ and the result in (16) for $l = x$ follows.

For the economy with transitory shocks only, taking the limits of (24) and (25) as horizon τ approaches zero or infinity gives the result in (15). Furthermore,

$$\hat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau) = \frac{\left(\sqrt{\sigma_y^2 (\sigma_y^2 + \sigma_z^2)} - \sigma_y^2 \right) e^{-2\lambda_z \tau} (e^{\lambda_z \tau} - 1)^2}{\lambda_z \tau} > 0. \quad (27)$$

The derivative of the difference in (27) with respect to horizon is

$$\frac{\partial (\hat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau))}{\partial \tau} = \frac{e^{-2\lambda_z \tau} (e^{\lambda_z \tau} - 1) (1 + 2\lambda_z \tau - e^{\lambda_z \tau}) \left(\sqrt{\sigma_y^2 (\sigma_y^2 + \sigma_z^2)} - \sigma_y^2 \right)}{\lambda_z \tau^2}.$$

The sign of this derivative depends on the sign of $1 + 2\lambda_z \tau - e^{\lambda_z \tau}$ and the result in (16) for $l = z$ follows.

□

B.5 Proposition 5

Proof. The market dividend is a deterministic function of consumption:

$$D_{m,t} = e^{-\beta_m t} C_t^\phi,$$

where $\beta_m \equiv \phi\mu - \mu_m$. The dynamics of the market dividend are obtained by applying Ito's lemma to the above equation.

To project firm i 's dividend D_i on the observable filtration, we define the process H_i as

$$H_{i,t} \equiv \log D_{i,t} - \bar{\phi} \log C_t.$$

The drift of H_i depends exclusively on the unobservable component x_i . Indeed,

$$dH_{i,t} = (\bar{\mu} - \bar{\phi}\mu + \bar{\phi}x_{i,t})dt + (\phi - \bar{\phi})\sqrt{v}dB_t.$$

Following [Liptser and Shiryaev \(2001\)](#), observing H_i and s_i yields the agent's estimate of x_i , $\hat{x}_{i,t} \equiv \mathbb{E}[x_{i,t} \mid \mathcal{F}_t^o]$. The dynamics of the filter $\hat{x}_i$ are given by

$$d\hat{x}_{i,t} = -\bar{\lambda}\hat{x}_{i,t}dt + \hat{\sigma}_{x_{i,t}}d\hat{B}_t + p_i\bar{\sigma}ds_{i,t},$$

where $\hat{\sigma}_{x_{i,t}} \equiv \frac{\bar{\phi}}{\sqrt{v}|\phi - \bar{\phi}|}\gamma_{x_{i,t}}$ and $\gamma_{x_{i,t}} \equiv \text{Var}[x_{i,t} \mid \mathcal{F}_t^o]$. The dynamics of H_i with respect to the observable filtration is

$$dH_{i,t} = (\bar{\mu} - \bar{\phi}\mu + \bar{\phi}\hat{x}_{i,t})dt + (\phi - \bar{\phi})\sqrt{v}d\hat{B}_t,$$

and the $\mathcal{F}^o$ -adapted dynamics of D_i follow directly from the definition of H_i . Note that the Brownian motion $\hat{B}$ driving the dynamics of D_i and $\hat{x}_i$ is exactly the same as that driving the output C , the permanent component $\hat{x}$, and the transitory component $\hat{z}$ in Equations (6) to (8).

The posterior variance $\gamma_{x_{i,t}}$ satisfies

$$\frac{d\gamma_{x_{i,t}}}{dt} = (1 - p_i^2)\bar{\sigma}^2 - 2\bar{\lambda}\gamma_{x_{i,t}} - \frac{\bar{\phi}^2}{v(\phi - \bar{\phi})^2}\gamma_{x_{i,t}}^2.$$

As in [Scheinkman and Xiong \(2003\)](#) and [Dumas et al. \(2009\)](#) among others, we focus on the stationary solution: $\frac{d\gamma_{x_{i,t}}}{dt} = 0$. The steady-state posterior variance is given by

$$\gamma_{x_i} = \frac{v(\phi - \bar{\phi})^2}{\bar{\phi}^2} \left(-\bar{\lambda} + \sqrt{\bar{\lambda}^2 + \frac{\bar{\phi}^2}{v(\phi - \bar{\phi})^2}(1 - p_i^2)\bar{\sigma}^2} \right). \quad (28)$$

□

B.6 Proposition 6

Proof. This proof follows closely [Eraker and Shaliastovich \(2008\)](#). We conjecture that the log wealth-consumption ratio is affine in the state variables $X_t = (\hat{y}_t, \hat{y}_{m,t}, \hat{y}_{i,t}, \hat{x}_t, \hat{x}_{i,t}, \hat{z}_t)^\top$, where $\hat{y}_t \equiv \log C_t - \hat{z}_t$, $\hat{y}_{m,t} \equiv \log D_{m,t} - \phi \hat{z}_t$, and $\hat{y}_{i,t} \equiv \log D_{i,t} - \phi \hat{z}_t$, so that

$$wc_t \equiv \log \frac{W_t}{C_t} = A_0 + A_1^\top X_t, \quad (29)$$

and use the fact that the state variables belong to the affine class, so that their dynamics can be written as:

$$\begin{aligned} dX_t &= \mu(X_t)dt + \Sigma(X_t)d\hat{B}_t \\ \mu(X_t) &= \mathcal{M} + \mathcal{K}X_t \\ \Sigma(X_t)\Sigma(X_t)^\top &= h + \sum_{i=1}^6 H^i X_t^i, \end{aligned}$$

where X_t is the vector of the filtered state variables, $\Sigma(X_t) \in \mathbb{R}^{6 \times 1}$ encodes the diffusions of the state variables, $\mathcal{M} \in \mathbb{R}^6$, $\mathcal{K} \in \mathbb{R}^{6 \times 6}$, $h \in \mathbb{R}^{6 \times 6}$, $H \in \mathbb{R}^{6 \times 6 \times 6}$, and $\hat{B}_t$ is a standard Brownian motion.

The dynamics of the state-price density then can be written as

$$d \log M_t = (\theta \log \delta - (\theta - 1) \log k_1 + (\theta - 1)(k_1 - 1)A_1'(X_t - \mu_X))dt - \Omega' dX_t, \quad (30)$$

where $\mu_X = (0, 0, 0, 0, 0, 0)^\top$, $\Omega = \gamma(1, 0, 0, 0, 0, 1)^\top + (1 - \theta)k_1 A_1$, and the coefficients $A_0 \in \mathbb{R}$ and $A_1 \in \mathbb{R}^6$ are the loadings defined in (29).

The coefficients $A_0 \in \mathbb{R}$, $A_1 \in \mathbb{R}^6$ solve the following system of equations

$$0 = \mathcal{K}^\top \chi - \theta(1 - k_1)A_1 + \frac{1}{2}\chi^\top H \chi, \quad (31)$$

$$0 = \theta(\log \delta + k_0 - (1 - k_1)A_0) + \mathcal{M}^\top \chi + \frac{1}{2}\chi^\top h \chi, \quad (32)$$

and the linearization coefficient $k_1 \in \mathbb{R}$ satisfies

$$\theta \log k_1 = \theta(\log \delta + (1 - k_1)A_1^\top \mu_X) + \mathcal{M}^\top \chi + \frac{1}{2}\chi^\top h \chi,$$

where $\chi = \theta \left((1 - \frac{1}{\psi})(1, 0, 0, 0, 0, 1)^\top + k_1 A_1 \right)$.

Solving (31) for the vector of loadings $A_1 \in \mathbb{R}^6$ gives

$$A_1^\top = \left(0, 0, 0, \frac{1 - 1/\psi}{1 - k_1(1 - \lambda_x)}, 0, -\frac{\lambda_z(1 - 1/\psi)}{1 - k_1(1 - \lambda_z)} \right).$$

Plugging this solution in equation (32) allows to solve for the coefficient A_0 .

From the arbitrage theory we know that the state-price density M_t satisfies

$$\frac{dM_t}{M_t} = -r_t dt - \Lambda_t^\top dB_t.$$

where r_t is the risk-free rate and Λ_t is the market price of risk vector.

Eraker and Shaliastovich (2008) show that from the expression for the state price density in (30), the risk free rate and market price of risk vector can be determined as follows:

$$\begin{aligned} r_t &= r_0 + r_1^\top X_t, \\ \Lambda_t &= \Sigma(X_t)^\top \Omega, \end{aligned}$$

where the vector $\Omega = \gamma(1, 0, 0, 0, 0, 1)^\top + (1 - \theta)k_1 A_1$ and the coefficients $r_0 \in \mathbb{R}$ and $r_1 \in \mathbb{R}^6$ solve the system of equations

$$\begin{aligned} r_1 &= (1 - \theta)(k_1 - 1)A_1 + \mathcal{K}^\top \Omega - \frac{1}{2}\Omega^\top H\Omega, \\ r_0 &= -\theta \log \delta + (\theta - 1)(\log k_1 + (k_1 - 1)A_1^\top \mu_X) + \mathcal{M}^\top \Omega - \frac{1}{2}\Omega^\top h\Omega. \end{aligned}$$

Solving for r_1, r_0 gives $r_1^\top = (0, 0, 0, 1/\psi, 0, -\lambda_z/\psi)$ and

$$r_0 = -\frac{1 - \gamma}{1 - 1/\psi} \log \delta + \frac{1/\psi - \gamma}{1 - 1/\psi} \log k_1 + \gamma\mu - \frac{1}{2}\Theta(\hat{\sigma}_y, \hat{\sigma}_x, \hat{\sigma}_z),$$

where

$$\begin{aligned} \Theta(\hat{\sigma}_y, \hat{\sigma}_x, \hat{\sigma}_z) &\equiv \\ &\frac{1}{\psi^2(k_1(\lambda_x - 1) + 1)^2(k_1(\lambda_z - 1) + 1)^2}(\gamma\psi((k_1(\lambda_z - 1) + 1)(k_1((\lambda_x - 1)\hat{\sigma}_y + \hat{\sigma}_x) + \hat{\sigma}_y) \\ &\quad - (k_1 - 1)\hat{\sigma}_z(k_1(\lambda_x - 1) + 1)) + k_1\lambda_z\hat{\sigma}_z(k_1(\lambda_x - 1) + 1) + k_1\hat{\sigma}_x(k_1(-\lambda_z) + k_1 - 1))^2, \end{aligned}$$

where $\hat{\sigma}_y \equiv \sqrt{v} - \hat{\sigma}_z$, and $\hat{\sigma}_x, \hat{\sigma}_z$ are defined in Proposition 1. Similarly, market price of risk can be written as

$$\Lambda = \gamma\hat{\sigma}_y + \left(\frac{(\gamma - 1/\psi)}{1/k_1 - (1 - \lambda_x)}\right)\hat{\sigma}_x + \left(\gamma - \frac{\lambda_z(\gamma - 1/\psi)}{1/k_1 - (1 - \lambda_z)}\right)\hat{\sigma}_z.$$

Finally, following Eraker and Shaliastovich (2008), the dynamics of the vector of state variables X_t under the risk neutral measure $\mathbb{Q}$ are given by

$$dX_t = (\mathcal{M}^\mathbb{Q} + \mathcal{K}^\mathbb{Q}X_t)dt + \Sigma(X_t)d\hat{B}_t^\mathbb{Q},$$

where $\hat{B}_t^\mathbb{Q} = \hat{B}_t + \int_0^t \Lambda_s ds$ is a $\mathbb{Q}$ -Brownian motion and the coefficients $\mathcal{M}^\mathbb{Q} \in \mathbb{R}^6$ and

$\mathcal{K}^{\mathbb{Q}} \in \mathbb{R}^{6 \times 6}$ satisfy

$$\mathcal{M}^{\mathbb{Q}} = \mathcal{M} - h\Omega, \quad (33)$$

$$\mathcal{K}^{\mathbb{Q}} = \mathcal{K} - H\Omega. \quad (34)$$

□

B.7 Proposition 7

Proof. The price of a zero-coupon bond satisfies

$$Q(t, \tau) = \mathbb{E}_t^{\mathbb{Q}} \left(e^{-\int_t^{t+\tau} r_s ds} \right) = e^{q_0(\tau) + q_1(\tau) X_t},$$

where $X_t = (\hat{y}_t, \hat{y}_{mt}, \hat{y}_{it}, \hat{x}_t, \hat{x}_{it}, \hat{z}_t)^\top$. Eraker and Shaliastovich (2008) show that the functions $q_0(\tau) \in \mathbb{R}$ and $q_1(\tau) \in \mathbb{R}^6$ solve the following system of Ricatti equations

$$\frac{\partial}{\partial \tau} q_1(\tau) = -r_1 + \mathcal{K}^{\mathbb{Q}^\top} q_1(\tau) + \frac{1}{2} q_1(\tau)^\top H q_1(\tau), \quad (35)$$

$$\frac{\partial}{\partial \tau} q_0(\tau) = -r_0 + \mathcal{M}^{\mathbb{Q}^\top} q_1(\tau) + \frac{1}{2} q_1(\tau)^\top h q_1(\tau), \quad (36)$$

with boundary conditions $q_0(0) = 0$ and $q_1(0) = (0, 0, 0, 0, 0, 0)^\top$. Coefficients $\mathcal{M}^{\mathbb{Q}} \in \mathbb{R}^6$ and $\mathcal{K}^{\mathbb{Q}} \in \mathbb{R}^{6 \times 6}$ are characterized in (33)–(34).

Solving (35) gives

$$q_1(\tau)^\top = \left(0, 0, 0, -\frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}), 0, \frac{1}{\psi} (1 - e^{-\lambda_z \tau}) \right).$$

Using these results in (36) allows to solve for the function q_0 .

□

B.8 Proposition 8

Proof. The price of a dividend strip satisfies

$$S_j(t, \tau) = \mathbb{E}_t^{\mathbb{Q}} \left(e^{-\int_t^{t+\tau} r_s ds} D_{j,t+\tau} \right) = e^{w_{0j}(\tau) + w_{1j}(\tau) X_t}, \text{ with } j \in \{m, i\}$$

and $X_t = (\hat{y}_t, \hat{y}_{m,t}, \hat{y}_{i,t}, \hat{x}_t, \hat{x}_{i,t}, \hat{z}_t)^\top$. Eraker and Shaliastovich (2008) show that the functions $w_{0j}(\tau) \in \mathbb{R}$ and $w_{1j}(\tau) \in \mathbb{R}^6$ solve the following system of Ricatti equations

$$\frac{\partial}{\partial \tau} w_{1j}(\tau) = -r_1 + \mathcal{K}^{\mathbb{Q}^\top} w_{1j}(\tau) + \frac{1}{2} w_{1j}(\tau)^\top H w_{1j}(\tau), \quad (37)$$

$$\frac{\partial}{\partial \tau} w_{0j}(\tau) = -r_0 + \mathcal{M}^{\mathbb{Q}^\top} w_{1j}(\tau) + \frac{1}{2} w_{1j}(\tau)^\top h w_{1j}(\tau), \quad (38)$$

with boundary conditions $w_{0j}(0) = 0$ and either $w_{1m}(0) = (0, 1, 0, 0, 0, \phi)^\top$ or $w_{1i}(0) = (0, 0, 1, 0, 0, \phi)^\top$. Coefficients $\mathcal{M}^\mathbb{Q} \in \mathbb{R}^6$ and $\mathcal{K}^\mathbb{Q} \in \mathbb{R}^{6 \times 6}$ are characterized in (33)–(34).

Solving (37) gives

$$\begin{aligned} w_{1m}(\tau)^\top &= \left(0, 1, 0, \frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}) (\phi \psi - 1), 0, \phi + \frac{1}{\psi} (1 - e^{-\lambda_z \tau}) (1 - \phi \psi) \right), \\ w_{1i}(\tau)^\top &= \left(0, 0, 1, \frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}) (\bar{\phi} \psi - 1), \frac{\bar{\phi} (1 - e^{-\bar{\lambda} \tau})}{\bar{\lambda}}, \phi + \frac{1}{\psi} (1 - e^{-\lambda_z \tau}) (1 - \bar{\phi} \psi) \right). \end{aligned}$$

Using these results in (38) allows to solve for the function w_{0j} . The risk premium of asset $j \in \{m, i\}$ dividend strip for with time to maturity τ is computed as

$$\text{RP}_j(t, \tau) = -\frac{1}{dt} \left\langle \frac{dM_t}{M_t}, \frac{dS_j(t, \tau)}{S_j(t, \tau)} \right\rangle.$$

□

B.9 Proposition 9

Proof. Following Eraker and Shaliastovich (2008) we consider an approximate equilibrium solution for the price-dividend ratio, which is obtained, as wealth-consumption ratio in Proposition 6, through the log-linearization of returns. Namely, the log price-dividend ratio is linear in the state variables,

$$pd_{j,t} \equiv \log \frac{P_{j,t}}{D_{j,t}} = A_{0j} + A_{1j}^\top X_t, \text{ with } j \in \{m, i\}$$

and $X_t = (\hat{y}_t, \hat{y}_{mt}, \hat{y}_{it}, \hat{x}_t, \hat{x}_{it}, \hat{z}_t)^\top$. The coefficients $A_{0j} \in \mathbb{R}$, $A_{1j} \in \mathbb{R}^6$ solve the following system of equations

$$0 = \mathcal{K}^\top \chi_j + (\theta - 1)(k_1 - 1)A_1 + (k_{1j} - 1)A_{1j} + \frac{1}{2} \chi_j^\top H \chi_j, \quad (39)$$

$$\begin{aligned} 0 &= \theta \ln \delta - (\theta - 1) (\ln k_1 + (k_1 - 1)A_1^\top \mu_X) - (\ln k_{1j} + (k_{1j} - 1)A_{1j}^\top \mu_X) \\ &\quad + \mathcal{M}^\top \chi_j + \frac{1}{2} \chi_j^\top h \chi_j, \end{aligned} \quad (40)$$

where $\chi_m = (0, 1, 0, 0, 0, \phi)^\top + k_{1m}A_{1m} - \Omega$, $\chi_i = (0, 0, 1, 0, 0, \phi)^\top + k_{1i}A_{1i} - \Omega$, $k_{1m} \in \mathbb{R}$ and $k_{1i} \in \mathbb{R}$ are the linearization coefficients for the stock return.

Solving (39) for the vector of loadings $A_{1j} \in \mathbb{R}^6$ gives

$$\begin{aligned} A_{1m}^\top &= \left(0, 0, 0, \frac{\phi - 1/\psi}{1 - k_{1m}(1 - \lambda_x)}, 0, -\frac{\lambda_z(\phi - 1/\psi)}{1 - k_{1m}(1 - \lambda_z)} \right) \\ A_{1i}^\top &= \left(0, 0, 0, \frac{\bar{\phi} - 1/\psi}{1 - k_{1i}(1 - \lambda_x)}, \frac{\bar{\phi}}{1 - k_{1i}(1 - \bar{\lambda})}, -\frac{\lambda_z(\bar{\phi} - 1/\psi)}{1 - k_{1i}(1 - \lambda_z)} \right). \end{aligned}$$

Plugging this solution in equation (40) allows to solve for k_{1m} and k_{1j} . Then we obtain the intercept A_{0j} as

$$A_{0j} = \log \frac{k_{1j}}{1 - k_{1j}} - A_{1j}^\top \mu_X.$$

The risk premium of asset j is computed as

$$\text{RP}_j(t) = -\frac{1}{dt} \left\langle \frac{dM_t}{M_t}, \frac{dP_{j,t}}{P_{j,t}} \right\rangle.$$

□

C Asset Prices in the Full Information Economy

Proposition C.1. *The equilibrium state-price density in the full information economy has dynamics given by*

$$\frac{dM_t}{M_t} = -r_t dt - \Lambda^\top dB_t,$$

where $B_t = (B_{y,t}, B_{x,t}, B_{x_i,t}, B_{z,t})^\top$. The risk-free rate satisfies

$$r_t = r_0 + r_x x_t + r_z z_t,$$

with

$$\begin{aligned} r_0 &= -\frac{1-\gamma}{1-1/\psi} \log \delta + \frac{1/\psi - \gamma}{1-1/\psi} \log k_1 + \gamma \mu \\ &\quad - \frac{1}{2} \left(\gamma^2 \sigma_y^2 + \left(\frac{(\gamma - 1/\psi)}{1/k_1 - (1 - \lambda_x)} \right)^2 \sigma_x^2 + \left(\gamma - \frac{\lambda_z(\gamma - 1/\psi)}{1/k_1 - (1 - \lambda_z)} \right)^2 \sigma_z^2 \right), \\ r_x &= \frac{1}{\psi}, \\ r_z &= -\frac{\lambda_z}{\psi}, \end{aligned}$$

and the market price of risk vector is

$$\Lambda^\top = \left(\gamma \sigma_y, \frac{(\gamma - 1/\psi)}{1/k_1 - (1 - \lambda_x)} \sigma_x, 0, \left(\gamma - \frac{\lambda_z(\gamma - 1/\psi)}{1/k_1 - (1 - \lambda_z)} \right) \sigma_z \right).$$

Proof. The proof is similar to the proof of analogous proposition for the partial information economy, Proposition 6. We conjecture that the log wealth-consumption ratio is affine in the state variables $X_t = (y_t, y_{m,t}, y_{i,t}, x_t, x_{i,t}, z_t)^\top$, where $y_{m,t} \equiv \log D_{m,t} - \phi z_t$ and $y_{i,t} \equiv \log D_{i,t} - \phi z_t$. We then use the fact that the state variables belong to the affine class, so that

their dynamics can be written as:⁹

$$\begin{aligned} dX_t &= \mu(X_t)dt + \Sigma(X_t)dB_t \\ \mu(X_t) &= \mathcal{M} + \mathcal{K}X_t \\ \Sigma(X_t)\Sigma(X_t)^\top &= h + \sum_{i=1}^6 H^i X_t^i. \end{aligned}$$

Moreover, following [Eraker and Shaliastovich \(2008\)](#), the dynamics of the vector of state variables X_t under the risk neutral measure $\mathbb{Q}$ are given by

$$dX_t = (\mathcal{M}^\mathbb{Q} + \mathcal{K}^\mathbb{Q}X_t)dt + \Sigma(X_t)dB_t^\mathbb{Q},$$

where the coefficients can be identified analogously to (33)–(34). □

Proposition C.2. *The price of the zero-coupon bond with time to maturity τ in the full information economy is given by*

$$Q(t, \tau) = \mathbb{E} \left[\frac{M_{t+\tau}}{M_t} \mid \mathcal{F}_t \right] = e^{q_0(\tau) + q_x(\tau)x_t + q_z(\tau)z_t},$$

where

$$q_x(\tau) = -\frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}), \quad q_z(\tau) = \frac{1}{\psi} (1 - e^{-\lambda_z \tau}).$$

and $q_0(\tau)$ solves

$$\frac{\partial}{\partial \tau} q_0(\tau) = -r_0 + \mathcal{M}^\mathbb{Q}^\top q_1(\tau) + \frac{1}{2} q_1(\tau)^\top h q_1(\tau)$$

with $q_1(\tau)^\top \equiv (0, 0, 0, q_x(\tau), 0, q_z(\tau))$.

Proof. Analogous to the proof of Proposition 7. □

Proposition C.3. *The price of a (market or firm) dividend strip with time to maturity τ in the full information economy is given by*

$$S_j(t, \tau) = \mathbb{E} \left[\frac{M_{t+\tau}}{M_t} D_{j,t+\tau} \mid \mathcal{F}_t \right] = D_{j,t} e^{w_{0j}(\tau) + w_{xj}(\tau)x_t + w_{zj}(\tau)z_t + w_{x_{ij}}(\tau)x_{i,t}}, \quad j \in \{m, i\}$$

where

$$w_{xj}(\tau) = \frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}) (\phi_j \psi - 1), \quad w_{zj}(\tau) = \frac{1}{\psi} (1 - e^{-\lambda_z \tau} (1 - \phi_j \psi)),$$

⁹Note that in this appendix we use X to denote the vector of state variables in the full information economy.

with $\phi_m = \phi$ and $\phi_i = \bar{\phi}$,

$$w_{x_i m}(\tau) = 0, \quad w_{x_i i}(\tau) = \frac{\bar{\phi}(1 - e^{\bar{\lambda}\tau})}{\bar{\lambda}},$$

and $w_{0j}(\tau)$ solves

$$\frac{\partial}{\partial \tau} w_{0j}(\tau) = -r_0 + \mathcal{M}^{\mathbb{Q}\top} w_{1j}(\tau) + \frac{1}{2} w_{1j}(\tau)^\top h w_{1j}(\tau), \quad j \in \{m, i\}$$

with $w_{1m}(\tau)^\top = (0, 1, 0, w_{xm}(\tau), w_{x_i m}(\tau), w_{zm}(\tau))$, $w_{1i}(\tau)^\top = (0, 0, 1, w_{xi}(\tau), w_{x_i i}(\tau), w_{zi}(\tau))$ and $w_{0j}(0) = 0$. The risk premium of the dividend strip for asset $j \in \{m, i\}$ is

$$RP_j(t, \tau) = -\frac{1}{dt} \left\langle \frac{dM_t}{M_t}, \frac{dS_j(t, \tau)}{S_j(t, \tau)} \right\rangle = (\phi \sigma_y, w_{xj}(\tau) \sigma_x, w_{x_{ij}}(\tau) \sigma_{x_i}, (\phi + w_{zj}(\tau)) \sigma_z)^\top \Lambda.$$

Proof. Analogous to the proof of Proposition 8. □

Proposition C.4. The (market or firm) stock price in the full information economy is

$$P_{j,t} = \int_0^\infty \mathbb{E}_t \left[\frac{M_{t+\tau}}{M_t} D_{j,t+\tau} \mid \mathcal{F}_t \right] d\tau \approx D_{j,t} e^{A_{0j} + A_{xj}x_t + A_{zj}z_t + A_{x_{ij}}x_{i,t}}, \quad j \in \{m, i\},$$

where

$$A_{xj} = \frac{\phi_j - 1/\psi}{1 - k_{1j}(1 - \lambda_x)}, \quad A_{zj} = -\frac{\lambda_z(\phi_j - 1/\psi)}{1 - k_{1j}(1 - \lambda_z)}$$

with $\phi_m = \phi$ and $\phi_i = \bar{\phi}$,

$$A_{x_{im}} = 0, \quad A_{x_{ii}} = \frac{\bar{\phi}}{1 - k_{1j}(1 - \bar{\lambda})},$$

and A_{0j} is satisfies

$$A_{0j} = \log \frac{k_{1j}}{1 - k_{1j}} - A_{1j}^\top \mu_X,$$

where $A_{1j}^\top = (0, 0, 0, A_{xj}, A_{x_{ij}}, A_{zj})$ and the linearization coefficient k_{1j} solves a full information analogue of (40). The risk premium of asset j is given by

$$RP_j(t) = -\frac{1}{dt} \left\langle \frac{dM_t}{M_t}, \frac{dP_{j,t}}{P_{j,t}} \right\rangle = (\phi \sigma_y, A_{xj} \sigma_x, A_{x_{ij}} \sigma_{x_i}, (\phi + A_{zj}) \sigma_z)^\top \Lambda.$$

Proof. Analogous to the proof of Proposition 9. □