



LTI@UniTO
Long-Term Investors

Collegio Carlo Alberto
UNIVERSITÀ DEGLI STUDI DI TORINO

Learning and Term Structures

Michael Hasler
Mariana Khapko
Roberto Marfé

No. 1
September 2018

LTI WORKING PAPERS

www.carloalberto.org/lti/working-papers

© 2018 by Michael Hasler, Mariana Khapko and Roberto Marfé.
. Any opinions expressed here are those of the authors and not those of LTI.

Learning and Term Structures*

Michael Hasler[†]

Mariana Khapko[‡]

Roberto Marfè[§]

August 20, 2018

Abstract

We study the impact of rational learning on the term structures of consumption growth risk, interest rates, and equity risk premia in general equilibrium. Learning yields lower interest rates and a larger equity risk premium than in an otherwise identical economy with full information. In opposition to the full information economy, learning implies an upward-sloping term structure of interest rates and a downward-sloping term structure of equity risk premia. Therefore, incomplete information and rational learning helps to explain jointly the empirically observed term structures of interest rates and equity risk premia.

*Research support from Long-Term Investors@UniTo (LTI@UniTO) is gratefully acknowledged.

[†]University of Toronto, 105 St. George Street, Toronto, ON, M5S 3E6, Canada, Michael.Hasler@rotman.utoronto.ca; www.utoronto.hasler

[‡]University of Toronto, 105 St. George Street, Toronto, ON, M5S 3E6, Canada, Mariana.Khapko@rotman.utoronto.ca; www.utoronto.khapko

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

1 Introduction

The factors driving economic fundamentals are unobservable. We study the role of incomplete information and learning in an asset-pricing model with time-varying economic growth.¹ Investors, unable to directly observe economic fundamentals, use a Bayesian model to update their beliefs about future growth prospects. We show that rational learning changes agents' perception of consumption growth risk across different horizons and helps to understand the empirical patterns of the term structures of equity risk premia and interest rates (van Binsbergen, Brandt, and Koijen, 2012), thus overcoming the limitations of the full information framework (Beeler and Campbell, 2012).

We analyze an economy in which (i) investors operate under partial information and (ii) aggregate consumption is driven by two sources of risk: a long-lasting component and a contemporaneous shock. The long-lasting component of consumption growth must be small to generate a realistic model-implied volatility of consumption growth. Such a component is difficult to detect statistically, even in large samples. Therefore, we follow Croce, Lettau, and Ludvigson (2015) and consider a partial information model in which investors face a learning problem. They observe changes in consumption, but they cannot observe the long-lasting component driving the change. Unlike in Croce et al. (2015), the representative agent in our model engages in *fully rational* learning.

Furthermore, we follow Marfè (2017) and assume that the long-lasting component of consumption growth depends on two latent factors. The first is the usual long-run component considered in the long-run risk literature. The second is a mean-reverting, transitory component that captures business cycle fluctuations. The long-run component implies that consumption growth risk tends to increase with the time horizon, whereas the transitory

¹The importance of time variation in economic growth for asset pricing is highlighted by Bansal and Yaron (2004), who show that a small but persistent component in the drivers of consumption growth generates large risk premiums and high Sharpe ratios if investors have a preference for early resolution of uncertainty. An extensive literature has followed this work. See, for instance, Bansal, Dittmar, and Lundblad (2005), Bansal, Dittmar, and Kiku (2009), Bansal, Kiku, and Yaron (2010, 2012, 2016), Drechsler and Yaron (2011), Colacito and Croce (2011, 2013), Segal, Shaliastovich, and Yaron (2015), and Schorfheide, Song, and Yaron (2018) among others.

component has the exact opposite impact. Since the impact of the transitory component dampens that of the long-run component, the term structure of consumption growth risk is about flat, as in the data (Marfè, 2017; Dew-Becker, 2017).

Consistent with the empirical findings of van Binsbergen et al. (2012), we show that partial information and rational learning implies a downward-sloping term structure of equity risk premia, whereas the full information model yields an upward-sloping term structure. That is, rational learning explains jointly the timing of consumption risk and that of equity risk premia. The equity risk premium under partial information is sizable and larger than under full information. Moreover, the slope of the term structure of interest rates is positive under partial information, whereas it is negative under full information. Also, the short-term real bond yield is smaller under the partial information than under full information. These results provide evidence that the predictions of the model with partial information and learning are consistent with data.

The economic mechanism is as follows. Under full information, each priced source of risk has a specific effect on the slope of risk premia. Specifically, the long-run component, the transitory component, and the contemporaneous shock to consumption have respectively an upward-sloping, a downward-sloping, and a neutral effect on the term structure of equity risk premia. The upward-sloping effect dominates even if consumption growth risk is flat across different horizons. Under partial information, the investor needs to infer the latent components using only one source of information; the realized consumption growth rate. As a result, learning yields an endogenous, perfect correlation between shocks to realized consumption growth, shocks to the long-run component, and shocks to the transitory component. This implies that the economy, as perceived by the investor, is driven by a unique priced source of risk. In such a case, the slope of the term structure of equity risk premia depends on the price sensitivities to the filtered long-run component and the filtered transitory component. Because prices are more sensitive to the transitory component than to the long-run component, the impact of the transitory component dominates and yields a

downward-sloping term structure of equity risk premia. Moreover, since bonds are used to hedge equity risk, short-term bonds are more expensive and therefore feature lower yields than long-term bonds.

This paper is related to the long-run risk literature, which was launched by [Bansal and Yaron \(2004\)](#). In this literature, the long-run component is highly persistent, which yields highly volatile consumption growth rates over long horizons. However, empirical evidence documents that consumption growth risk is about the same at short and long horizons ([Marfè, 2017](#); [Dew-Becker, 2017](#)). This suggests that the high empirical level of the equity premium is unlikely to be rationalized by highly volatile long-horizon cash-flows ([Beeler and Campbell, 2012](#)). In addition, long-run risk models imply an upward-sloping term structure of equity risk premia, which is inconsistent with the empirical findings of [van Binsbergen et al. \(2012\)](#), [van Binsbergen, Hueskes, Koijen, and Vrugt \(2013\)](#), [van Binsbergen and Koijen \(2017\)](#), and [Weber \(2017\)](#), as well as a downward-sloping term structure of interest rates. In our model, the existence of the transitory component dampens the upward-sloping impact of the long-run component. Therefore, the model can explain jointly the flat term structure of consumption growth risk, the downward-sloping term structure of equity risk premia, the upward-sloping term structure of interest rates, and the high equity risk premium.

The most closely related paper is that of [Croce et al. \(2015\)](#). Similar to their work, we study the impact of information and learning on the shape of the term structure of equity risk premia. However, while we focus on fully rational learning, they depart from the standard Bayesian learning framework to explain the timing of equity risk premia. Indeed, they show that the observed timing of equity risk premia and the high risk premium can be explained simultaneously if the representative agent learns about the latent factors under bounded rationality. In contrast, we show that rational learning can also generate a downward-sloping term structure of equity risk premia and a high equity premium when the model-implied timing of consumption growth risk is designed to match its empirical counterpart. That is, when the long-lasting component of consumption growth depends on both a long-run

component and a transitory component.

2 Economic Fundamentals

In this section we describe the economy and discuss the implications of learning on the term structure of consumption growth risk. The only information available to the representative agent is the one generated by the aggregate consumption process. Importantly, the underlying factors driving the consumption growth dynamics are not directly observable. This introduces learning into the decision problem of the agent and has implications for the agent's perception of cash-flow risk at different horizons.

2.1 Consumption Dynamics

The aggregate consumption, 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 written

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

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

where the mean-reverting process z satisfies

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

The three Brownian motions, B_y , B_x , and B_z are independent. The agent observes only the level of consumption C and, through Bayesian updating, filters out the unobservable factors

$\theta = (x, z)'$. The full filtration generated by observing all three Brownian shocks is denoted by \mathbf{F} .

The aim of considering both components, y and z , is to introduce some flexibility in modeling the timing of risk (Marfè, 2017). The first component, y , is an integrated process which depends on the time integral of x . That is, shocks in x accumulate and therefore permanently affect future consumption levels. For this reason, we call x the stochastic drift of the permanent component. The second component, z , depends on the current value of z . Since the process z is mean-reverting, shocks in z dissipate as time passes and therefore have a transitory impact on consumption. For this reason, we call z the transitory component. Without the z component, the aggregate consumption follows the standard dynamics considered in the literature on incomplete information and learning (e.g. Genotte, 1986; Detemple, 1986) as well as in the long-run risk literature pioneered by Bansal and Yaron (2004). In this case, consumption features, as we will show, an upward-sloping term structure of consumption growth risk 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 consumption growth risk. The existence of both x and z therefore provides flexibility in the modeling of the timing of consumption growth risk.

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

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

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 an increment of a standard Brownian motion.

2.2 Bayesian Learning

The expected growth rate of consumption varies over time due to shocks that come from two sources: the drift of the permanent component x_t defined in (3) and the transitory component z_t defined in (4). The agent only has access to information generated by the observation of the realized aggregate consumption path in (5), and thus 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 $\mathbf{F}^o = \{\mathcal{F}_t^o\}_{t \geq 0}$, defined as the flow of information generated by the path of consumption. In other words, the agent needs to filter out through Bayesian updating the unobservable components $\theta = (x, z)'$ by observing the history of consumption only. Proposition 1 provides the dynamics of the filtered state variables.

Proposition 1. *With respect to the agent's observation filtration, the dynamics of consumption C , and the filtered state variables \hat{x}_t , and \hat{z}_t 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 $\hat{x}_t \equiv \mathbb{E}[x_t | \mathcal{F}_t^o]$, $\hat{z}_t \equiv \mathbb{E}[z_t | \mathcal{F}_t^o]$, \hat{B}_t is an \mathcal{F}_t^o -Brownian motion, and

$$\hat{\sigma}_{x,t} = \frac{\gamma_{x,t} - \lambda_z \gamma_{xz,t}}{\sqrt{v}}, \quad \hat{\sigma}_{z,t} = \frac{\sigma_z^2 + \gamma_{xz,t} - \lambda_z \gamma_{z,t}}{\sqrt{v}}.$$

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-consumption, $\log C_t$, projected on the observable filtration, while Equations (8) and (7) 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 accord with the literature (e.g., Scheinkman and Xiong, 2003; Dumas, Kurshev, and Uppal, 2009), we replace Γ_t with its steady-state (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 consumption growth rates to reach the most precise variance estimate of the unobserved components. We will use $\hat{\sigma}_x$ and $\hat{\sigma}_z$ to denote the corresponding steady-state volatilities of the two filter estimates. The steady-state volatilities $\hat{\sigma}_x$ and $\hat{\sigma}_z$ are characterized in Appendix B.1.

2.3 Timing of Consumption Growth Risk

The goal here is to study how cash flow risk varies across different horizons. To this end, we follow Belo, Collin-Dufresne, and Goldstein (2015) and Marfè (2017) and compute an

annualized measure of consumption growth volatility under the full filtration \mathbf{F} :

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

or under the observation filtration \mathbf{F}^o :

$$\widehat{\sigma}_C(t, \tau) = \sqrt{\frac{1}{\tau} \log \left(\frac{\mathbb{E}_t[C_{t+\tau}^2 | \mathcal{F}_t^o]}{\mathbb{E}_t[C_{t+\tau} | \mathcal{F}_t^o]^2} \right)} \quad (14)$$

where τ denotes the horizon.

To compare risk across horizons, we also look at the corresponding term structures of variance ratios:

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

with reference of one year.

We study how learning affects the perception of risk across horizons and, then, how learning alters the shape of the term structures of cash flow risk. To build intuition, our analysis focuses first on the simplified models with either a permanent shock only or a transitory shock only, and then considers the general case that accounts for both permanent and transitory shocks.

2.3.1 The Case of Permanent Shocks Only

In this subsection, we assume that consumption is an integrated process with a drift driven by the process x only. That is, $d \log C_t = dy_t$.

Under the full filtration \mathbf{F} , the term structure of risk is monotone increasing. The more volatile (i.e., the larger $\sigma_x > 0$) or the more persistent (i.e., the smaller $\lambda_x > 0$) the drift x , the higher the level of the term structure of consumption growth risk. The same properties are also shared by the term structure of risk under the partial information filtration \mathbf{F}^o . The

aforementioned results are formalized in Proposition 2 below.

Proposition 2. *The following properties hold for any horizon $\tau > 0$:*

$$\begin{aligned} \partial_\tau \sigma_C(t, \tau) &> 0, & \partial_\tau \widehat{\sigma}_C(t, \tau) &> 0, \\ \partial_{\sigma_x} \sigma_C(t, \tau) &> 0, & \partial_{\sigma_x} \widehat{\sigma}_C(t, \tau) &> 0, \\ \partial_{\lambda_x} \sigma_C(t, \tau) &< 0, & \partial_{\lambda_x} \widehat{\sigma}_C(t, \tau) &< 0. \end{aligned}$$

Proof. See Appendix B.2.

These results obtain because x is the instantaneous drift of an integrated process y . That is, fluctuations in x accumulate over time and contribute to the integrated path of y . Thus, the longer the horizon, the larger the accumulated variation of x and so the larger the variance of y relative to such an horizon.

The same reasoning also applies to the term structure of consumption growth risk under the observation filtration. However, the term structures under \mathbf{F} and \mathbf{F}^o are not equal. They are both increasing and share the same short-run and long-run limits but the consumption growth variance perceived by the agent under the partial information filtration is larger than that under the full information filtration. The difference is a hump-shaped function of the horizon. These results are formalized in Proposition 3 below.

Proposition 3. *The following properties hold:*

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

Moreover, for any finite horizon $\tau > 0$:

$$\widehat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau) > 0,$$

and

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

where

$$\bar{\tau}_x \equiv -\frac{1}{2\lambda_x} \left(1 + 2\mathcal{L}\left(-1, -\frac{1}{2\sqrt{e}}\right) \right) > 0$$

and $\mathcal{L}(k, z)$ is k -th solution of the Lambert- W (or product logarithm) function.

Proof. See Appendix B.3.

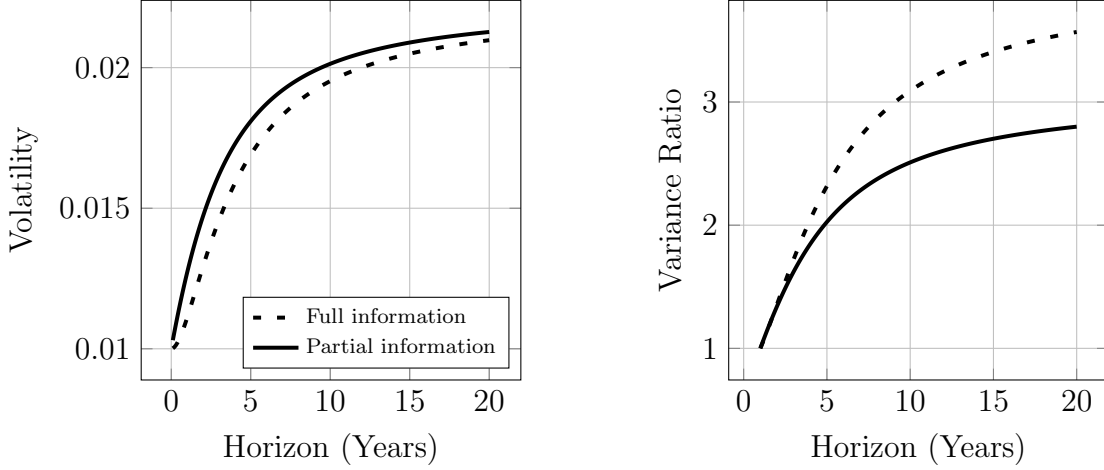
Why does the term structure of consumption growth risk under the partial information filtration lie above that obtained under the full information filtration? Because the agent observes only the level of consumption, uncertainty is generated by a unique Brownian motion under her observation filtration. As a result, the filtered variables are driven by a unique Brownian motion and, hence, are instantaneously perfectly correlated. Positive (negative) shocks to the level y are perceived to come together with positive (negative) shocks to its expected growth x . Such positive correlation increases $\widehat{\sigma}_C^2(t, \tau)$ relative to $\sigma_C^2(t, \tau)$ because y and x are instead uncorrelated under \mathbf{F} .

Moreover, the difference between the perceived and true consumption growth risk is a hump shaped function of the horizon: It increases up to a threshold $\bar{\tau}_x$, and decreases afterwards. At the horizon $\bar{\tau}_x$, the divergence between the agent's perception of consumption growth risk under the full and partial information is maximal. We note that the threshold $\bar{\tau}_x$ is decreasing in λ_x . Consequently, the difference between the perceived and the true consumption growth risk increases for a longer horizon when the persistence of x is high, or in other words, when the mean-reversion speed λ_x is low.

Figure 1 illustrates the term structure of consumption growth risk in the model with only permanent shocks.

Figure 1: Term Structure of Consumption Growth Risk with Permanent Shocks Only.

Parameter values are $\mu = 0.025$, $\sigma_y = 0.01$, $\lambda_x = 0.5$, $\sigma_x = 0.01$.



2.3.2 The Case of Transitory Shocks Only

In this subsection, we assume that the drift of consumption is only driven by transitory shocks z . That is, $d \log C_t = dy_t + dz_t$, where $dy_t = \mu dt + \sigma_y dB_{y,t}$.

Both under the full and partial information filtrations, the term structure of consumption growth risk is monotone decreasing. The reason is that shocks to z affect consumption in a transitory way; as time passes, the impact of these shocks on consumption weakens. Therefore, risk is higher in the short term than in the long term. The more volatile (i.e., the larger $\sigma_z > 0$) or the more persistent (i.e., the smaller $\lambda_z > 0$) the process z , the higher is the level of the term structure of consumption growth risk. These results are summarized in Proposition 4 below.

Proposition 4. *The following properties hold for any horizon $\tau > 0$:*

$$\begin{aligned} \partial_\tau \sigma_C(t, \tau) &< 0, & \partial_\tau \widehat{\sigma}_C(t, \tau) &< 0, \\ \partial_{\sigma_z} \sigma_C(t, \tau) &> 0, & \partial_{\sigma_z} \widehat{\sigma}_C(t, \tau) &> 0, \\ \partial_{\lambda_z} \sigma_C(t, \tau) &< 0, & \partial_{\lambda_z} \widehat{\sigma}_C(t, \tau) &< 0. \end{aligned}$$

Proof. See Appendix B.4.

Similar to the case of permanent shocks only, the term structures of consumption growth risk under \mathbf{F} and \mathbf{F}° are not equal. At any finite horizon, the risk perceived under the partial information filtration is higher. In addition, the difference between the partial and full information term structures is a hump-shaped function of the horizon. These results are summarized in Proposition 5 below.

Proposition 5. *The following properties hold:*

$$\lim_{\tau \rightarrow 0} \sigma_C^2(t, \tau) = \lim_{\tau \rightarrow 0} \widehat{\sigma}_C^2(t, \tau) = \sigma_y^2 + \sigma_z^2, \quad (17)$$

$$\lim_{\tau \rightarrow \infty} \sigma_C^2(t, \tau) = \lim_{\tau \rightarrow \infty} \widehat{\sigma}_C^2(t, \tau) = \sigma_y^2.$$

Moreover, for any finite horizon $\tau > 0$:

$$\widehat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau) > 0,$$

and

$$\partial_\tau (\widehat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau)) \begin{cases} > 0 & \tau < \bar{\tau}_z, \\ < 0 & \tau > \bar{\tau}_z, \end{cases} \quad (18)$$

where

$$\bar{\tau}_z \equiv -\frac{1}{2\lambda_z} \left(1 + 2\mathcal{L}\left(-1, -\frac{1}{2\sqrt{e}}\right) \right) > 0$$

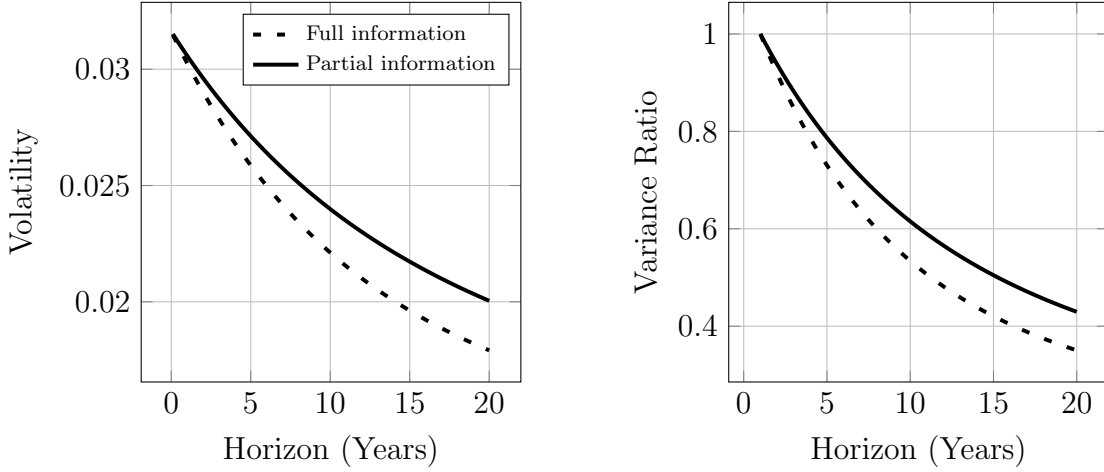
and $\mathcal{L}(k, z)$ is k -th solution of the Lambert- W function.

Proof. See Appendix B.5.

Figure 2 illustrates the term structure of consumption growth risk in the model with only transitory shocks.

Figure 2: Term Structure of Consumption Growth Risk with Transitory Shocks Only.

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



2.3.3 The Case of Permanent and Transitory Shocks

When aggregate consumption is driven by both permanent and transitory shocks, the term structure of consumption growth risk can be either increasing or decreasing depending on which of the two shocks dominates. The limits of the annualized variance of consumption growth at short and long horizon, respectively, are

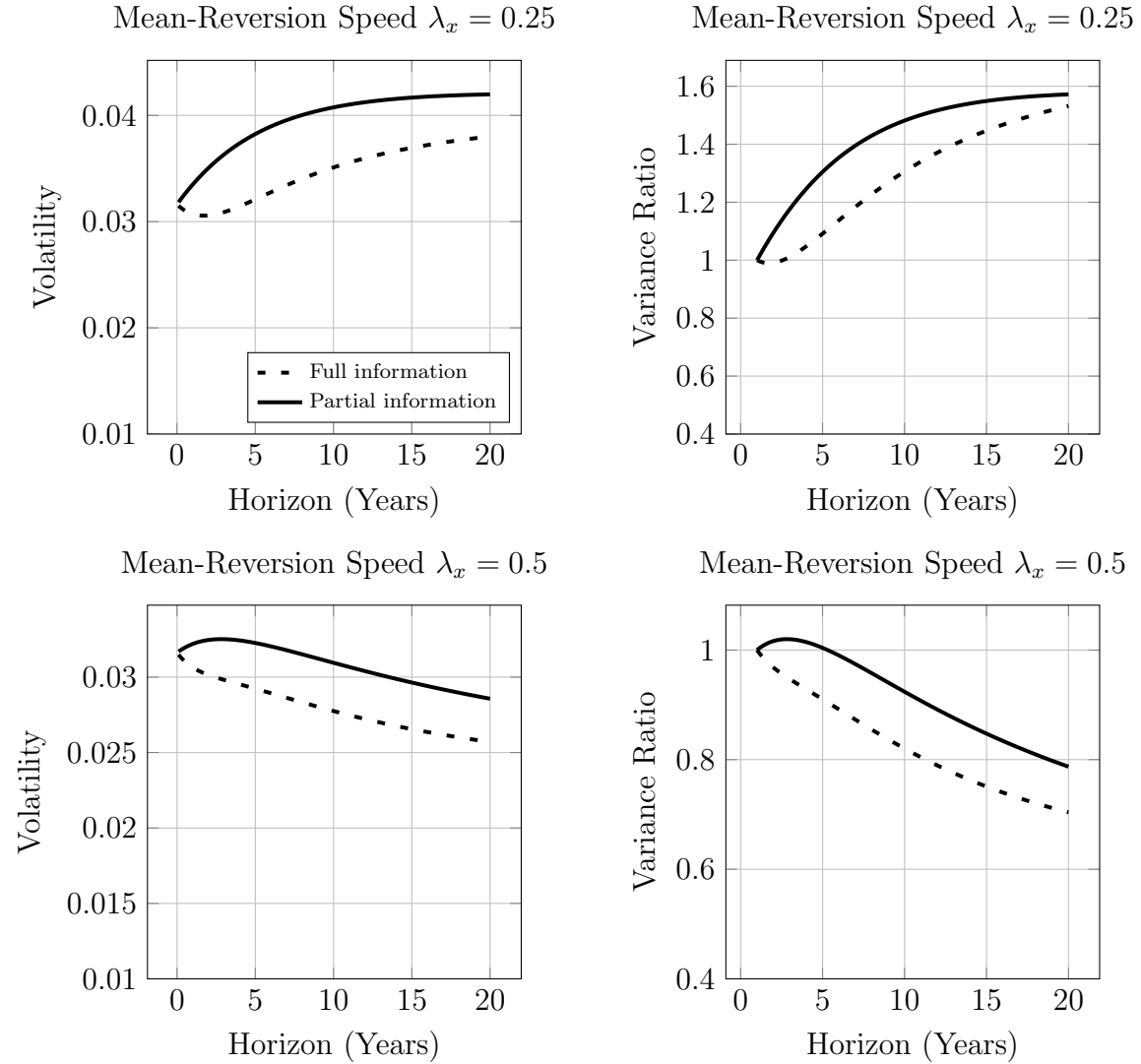
$$\begin{aligned} \lim_{\tau \rightarrow 0} \sigma_C^2(t, \tau) &= \lim_{\tau \rightarrow 0} \widehat{\sigma}_C^2(t, \tau) = \sigma_y^2 + \sigma_z^2, \\ \lim_{\tau \rightarrow \infty} \sigma_C^2(t, \tau) &= \lim_{\tau \rightarrow \infty} \widehat{\sigma}_C^2(t, \tau) = \sigma_y^2 + \frac{\sigma_x^2}{\lambda_x^2}.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

At the short end, consumption 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 consumption growth risk at the long end. The lower the mean-reversion speed of the permanent shock, λ_x , the higher the consumption growth volatility at the long end. Figure 3 illustrates the term structures of consumption growth risk in the model with both permanent and transitory shocks for different values of the mean-reversion speed λ_x .

Importantly, in the economy with both permanent and transitory shocks, partial in-

Figure 3: Term Structure of Consumption Growth Risk with Permanent and Transitory Shocks.

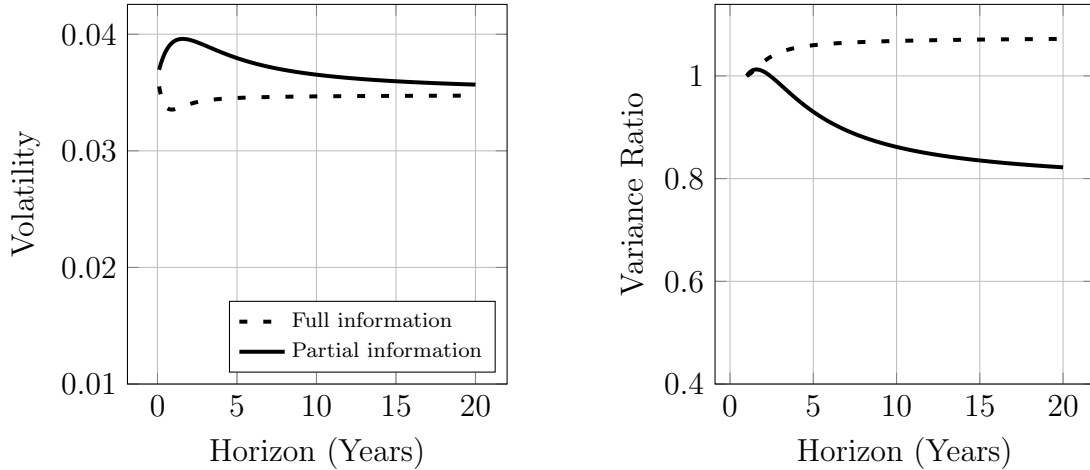
Parameter values are $\mu = 0.025$, $\sigma_y = 0.01$, $\sigma_x = 0.01$, $\lambda_z = 0.1$, $\sigma_z = 0.03$.



formation and learning can alter the shape of the term structure of consumption growth risk perceived by the agent. Figure 4 illustrates the case when the true term structure of consumption growth risk is decreasing up to one year and then flat. However the agent, unable to directly observe the shocks that drive consumption, perceives the term structure as increasing for up to one year and decreasing afterwards.

Figure 4: Term Structure of Consumption Growth Risk with Permanent and Transitory Shocks.

Parameter values are $\mu = 0.025$, $\sigma_y = 0.01$, $\lambda_x = 1.5$, $\sigma_x = 0.05$, $\lambda_z = 0.6$, $\sigma_z = 0.035$.



3 Asset Pricing

In this section we study the role of learning in the context of a general equilibrium asset pricing model. We focus on the equilibrium term structures of dividend strips and interest rates. We compare the shape of these term structures when those are derived either under full information or under partial information.

We consider an endowment economy (Lucas, 1978), in which the endowment process follows the dynamics in (1) and equals the representative agent’s aggregate consumption in equilibrium. As we have seen in Section 2, the permanent shock x induces an upward-sloping effect on the term structure of consumption growth risk, whereas the transitory shock z induces a downward-sloping effect. The two shocks jointly allow for a flexible shape of the term structure of consumption growth risk.

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

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

where δ is the time discount factor, γ is the coefficient of risk aversion, ψ is the elasticity of intertemporal substitution, and $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$.

Note that expectations are taken under the observation filtration \mathbf{F}^o . Thus, the dynamics of aggregate consumption depend on the filter estimates \hat{x} and \hat{z} , as provided in equations (6)–(8). The only source of uncertainty is the \mathcal{F}_t^o –Brownian motion \hat{B}_t .

In order to derive the price of dividend strips and equity, we assume a simple dynamics for dividends. In accord with most of the literature (e.g., [Abel, 1999](#); [Bansal and Yaron, 2004](#)) we define dividends as levered consumption:

$$D_t = e^{-\beta_d t} C_t^\phi,$$

where $\phi \geq 1$ is the leverage parameter and β_d is a parameter that determines the growth rate of dividends. Note that we do not need to alter the learning problem of the agent: Observing the dividend process does not bring any additional information compared to observing only the path of consumption. This is because the dividend process is a deterministic function of consumption.

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. The discrete time (continuously compounded) log-return on aggregate wealth W (e.g., the claim on the aggregate consumption stream $\{C_t\}_{t \geq 0}$) can be expressed as

$$\log R_{t+1} = \log \frac{W_{t+1} + C_{t+1}}{W_t} = \log (e^{w c_{t+1}} + 1) - w c_t + \log \frac{C_{t+1}}{C_t},$$

where $wc \equiv \log(W/C)$. A log-linearization of the first summand around the mean log wealth-consumption ratio leads to

$$\log R_{t+1} \approx k_0 + k_1 wc_{t+1} - wc_t + \log \frac{C_{t+1}}{C_t},$$

where the endogenous constants k_0 and k_1 satisfy

$$k_0 = -\log((1 - k_1)^{1-k_1} k_1^{k_1}) \quad \text{and} \quad k_1 = e^{\mathbb{E}(wc_t | \mathcal{F}^o)} / (1 + e^{\mathbb{E}(wc_t | \mathcal{F}^o)}).$$

Campbell, Lo, and MacKinlay (1997) and Bansal et al. (2012) provide evidence of the high accuracy of this log-linearization, which we assume exact hereafter. We follow Eraker and Shaliastovich (2008) and consider the continuous time counterpart defined as:

$$d \log R_t = k_0 dt + k_1 d(wc_t) - (1 - k_1) wc_t dt + d \log C_t. \quad (19)$$

Recursive preferences lead to the following Euler equation, which enables us to characterize the state-price density, M , that prices any asset in the economy:

$$\mathbb{E} \left[\exp \left(\log \frac{M_{t+\tau}}{M_t} + \int_t^{t+\tau} d \log R_s \right) \mid \mathcal{F}_t^o \right] = 1.$$

The state-price density satisfies

$$d \log M_t = \theta \log \delta dt - \frac{\theta}{\psi} d \log C_t - (1 - \theta) d \log R_t.$$

Proposition 6 below characterizes the state-price density, the risk-free rate, and the price of risk in our 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

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

and the market price of risk equals

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

where $\widehat{\sigma}_y \equiv \sqrt{v} - \widehat{\sigma}_z$, and $\widehat{\sigma}_x, \widehat{\sigma}_z$ are defined in Appendix B.1. $\Theta(\widehat{\sigma}_y, \widehat{\sigma}_x, \widehat{\sigma}_z)$ is defined in the Appendix B.6.

Proof. See Appendix B.6.

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

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

$$B(t, \tau) = \mathbb{E} \left[\frac{M_{t+\tau}}{M_t} \mid \mathcal{F}_t^o \right] = e^{q_0(\tau) + q_{\widehat{x}}(\tau)\widehat{x}_t + q_z(\tau)\widehat{z}_t},$$

where $q_0(\tau)$ is derived in Appendix B.7

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

The yield to maturity τ is defined as $YTM(t, \tau) = -(1/\tau) \log B(t, \tau)$.

Proof. See Appendix B.7.

Proposition 8 below characterizes the dividend strip price and its return moments.

Proposition 8. *The equilibrium price of the dividend strip with time to maturity τ is given by*

$$S(t, \tau) = \mathbb{E} \left[\frac{M_{t+\tau}}{M_t} D_{t+\tau} \mid \mathcal{F}_t^o \right] = e^{-\beta_d t + \phi \hat{y}_t + w_0(\tau) + w_{\hat{x}}(\tau) \hat{x}_t + w_{\hat{z}}(\tau) \hat{z}_t},$$

where $w_0(\tau)$ is derived in Appendix B.8 and

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

The return premium of the dividend strip with time to maturity τ is given by

$$RP(t, \tau) = -\frac{1}{dt} \left\langle \frac{dM_t}{M_t}, \frac{dS(t, \tau)}{S(t, \tau)} \right\rangle = (\phi \hat{\sigma}_y + w_{\hat{x}}(\tau) \hat{\sigma}_x + w_{\hat{z}}(\tau) \hat{\sigma}_z) \Lambda.$$

The return volatility of the dividend strip with time to maturity τ is given by

$$Vol(t, \tau) = \sqrt{\frac{1}{dt} \left\langle \frac{dS(t, \tau)}{S(t, \tau)} \right\rangle} = |\phi \hat{\sigma}_y + w_{\hat{x}}(\tau) \hat{\sigma}_x + w_{\hat{z}}(\tau) \hat{\sigma}_z|.$$

Proof. See Appendix B.8.

The log return on equity, $\log R^e$, is defined in a similar way as the return on aggregate wealth in (19):

$$d \log R_t^e = k_{0,d} dt + k_{1,d} d(pd_t) - (1 - k_{1,d}) pd_t dt + d \log D_t,$$

where $pd_t \equiv \log P_t/D_t$ and $k_{0,d}, k_{1,d}$ are endogenous constants. The equity price can be approximated as an exponential affine function of the state variables. Proposition 9 below characterizes the equity price and its return moments.

Proposition 9. *The equilibrium price of equity is given by*

$$P_t = \int_0^\infty \mathbb{E}_t \left[\frac{M_{t+\tau}}{M_t} D_{t+\tau} \mid \mathcal{F}_t^o \right] d\tau \approx D_t e^{A_d + B_{\hat{x},d} \hat{x}_t + B_{\hat{z},d} \hat{z}_t},$$

where A_d is derived in Appendix B.9 and

$$B_{\hat{x},d} = \frac{\phi - 1/\psi}{1 - k_{1,d}(1 - \lambda_x)},$$

$$B_{\hat{z},d} = - \frac{\lambda_z(\phi - 1/\psi)}{1 - k_{1,d}(1 - \lambda_z)}.$$

The equity risk premium is given by

$$RP(t) = - \frac{1}{dt} \left\langle \frac{dM_t}{M_t}, \frac{dP_t}{P_t} \right\rangle = (\phi \hat{\sigma}_y + B_{\hat{x},d} \hat{\sigma}_x + (\phi + B_{\hat{z},d}) \hat{\sigma}_z) \Lambda.$$

The equity return volatility is given by

$$\text{Vol}(t) = \sqrt{\frac{1}{dt} \left\langle \frac{dP_t}{P_t} \right\rangle} = \left| \phi \hat{\sigma}_y + B_{\hat{x},d} \hat{\sigma}_x + (\phi + B_{\hat{z},d}) \hat{\sigma}_z \right|.$$

Proof. See Appendix B.9.

4 Results

This section studies the asset pricing predictions of the model. We focus on the effect of learning for the term structures of equity risk premia and real interest rates.

4.1 Model Calibration

In order to understand the role of learning on asset prices from a term structure perspective, it is important to parametrize the consumption dynamics so that it matches the empirical properties of consumption growth risk across different horizons. To do so we calibrate the consumption dynamics parameters as follows. First, we match the short term (e.g. one-year) empirical level of volatility. Second, we match the empirical observation that consumption 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 consumption growth rates in the U.S. are approximatively flat around unity. Also, [Dew-Becker \(2017\)](#) documents that robust estimators of long-run consumption growth volatility are very close to estimates of the one-year volatility.

Therefore, we set the consumption growth parameter $\mu = 2.5\%$ and obtain the other consumption dynamics parameters $\Theta = \{\sigma_y, \sigma_x, \lambda_x, \sigma_z, \lambda_z\}$ by minimizing the following objective:

$$\Theta^* = \arg \min \left\{ [\sigma_C(t, 1) - 3\%]^2 + \alpha \int_0^{50} [VR_C(t, \tau) - 1]^2 d\tau \right\}$$

where α is a weighting constant. This minimization procedure yields: $\sigma_y = 0.033$, $\sigma_x = 0.040$, $\lambda_x = 1.346$, $\sigma_z = 0.033$, and $\lambda_z = 0.549$.

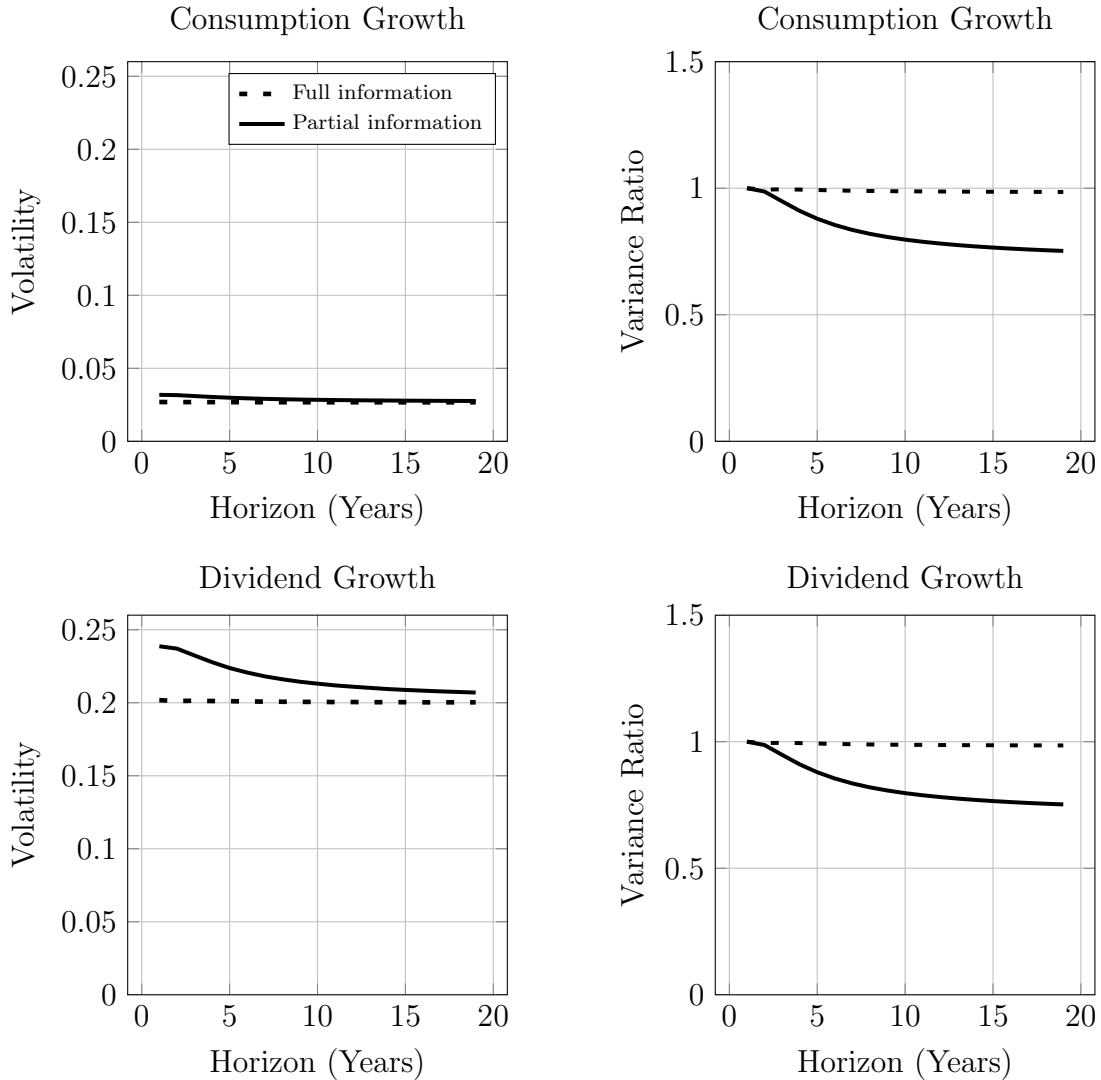
In addition we set $\beta_d = \mu(\phi - 1)$ such that the long-run growth rate of dividends equals that of consumption. The parameter ϕ captures the excessive volatility of dividends relative to consumption: We use $\phi = 7.5$ to approximatively match the one-year volatility of shareholders' remuneration in the U.S. which is about 20%, as documented in [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. Finally, we choose a relative risk aversion, $\gamma = 7.5$, an elasticity of intertemporal substitution, $\psi = 1.5$, and a time discount factor, $\delta = 0.99$.

Figure 5 shows the calibrated model-implied term structures of volatility and variance ratios for both consumption growth and dividend growth. Under full information, the term structure of consumption growth variance ratios is about flat because it was calibrated accordingly. This shows that the model dynamics are flexible enough to match the observed term structure of consumption growth variance ratios together with a 3% consumption growth volatility. Note that these flat variance ratios could also be obtained by assuming an i.i.d. process for consumption. However, the presence of the latent variables x_t and z_t in our model is key for asset pricing because the agent has to learn about these variables and their estimates are priced state variables in equilibrium. We note that, given our calibration, learning does not significantly affect the overall level of consumption growth volatility and dividend growth volatility. Indeed, the former lies in the interval 2.5-3%, while the latter lies in the interval 21-24%.

Importantly, learning alters the timing of consumption growth risk and dividend growth risk across different horizons. We observe that the variance ratios of both consumption growth and dividend 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. As shown in what follows, learning also affects the timing of equity risk premia via the impact of the priced risk factors \hat{x} and \hat{z} .

Figure 5: Term Structures of Consumption Growth Risk and Dividend Growth Risk Calibrated to U.S. Data.



4.2 Asset Pricing Moments

In this section, we highlight how partial information and learning affects the asset pricing moments, the term structure of equity risk premia, and the term structure of interest rates.

Table 1 compares the average model-implied and empirical risk-free rate, equity risk premium, equity return volatility, and dividend yield.

Under full information, the risk-free rate and equity premium are about 2.5% and 4.1%, respectively. These values are somewhat too high and too low respectively in comparison

Table 1: Model-Implied and Empirical Asset Pricing Moments.

	Data	Preference Setting					
	1931-2009	$\gamma = 7.5, \psi = 1.5$		$\gamma = 10, \psi = 1.5$		$\gamma = 7.5, \psi = 1.25$	
		Information		Information		Information	
		Full	Partial	Full	Partial	Full	Partial
Risk-free rate (%)	0.6	2.5	2.1	2.5	1.9	2.9	2.4
Equity premium (%)	6.2	4.1	5.2	5.6	6.9	4.0	5.2
Return volatility (%)	19.8	20.3	22.8	20.2	23.0	20.0	22.9
Sharpe ratio (%)	31.3	20.4	22.8	27.7	30.1	20.0	22.7
Dividend yield (%)	3.4	2.1	2.1	3.5	3.7	2.4	2.4

with actual data. The return volatility is quite large (20.3%) and, hence, the Sharpe ratio (20.4%) is below its empirical counterpart. Also, the dividend yield (2.1%) is lower than in actual data.

Under partial information, learning yields a lower risk-free rate (2.1%) and a higher equity premium (5.2%) than under full information. This corresponds to a decrease and an increase respectively of about 25%. Therefore, the model-implied risk-free rate and equity premium better match their empirical counterparts under partial information.

An increase in risk aversion ($\gamma = 10, \psi = 1.5$) modifies the results as follows. The risk-free rate decreases slightly and the equity premium increases substantially. This result obtains under both full and partial information. In the latter case, increasing risk aversion to ten is sufficient to match the empirical equity premium. Also, the Sharpe ratio and the dividend yield are quite 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 equity 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_t . That is, these two effects offset each other.

While our main goal to study the role of learning on the term structures, it is important to note that the model provides a good fit to the main asset pricing moments, and learning helps to solve the equity premium and risk-free rate puzzles.

4.3 Term Structures

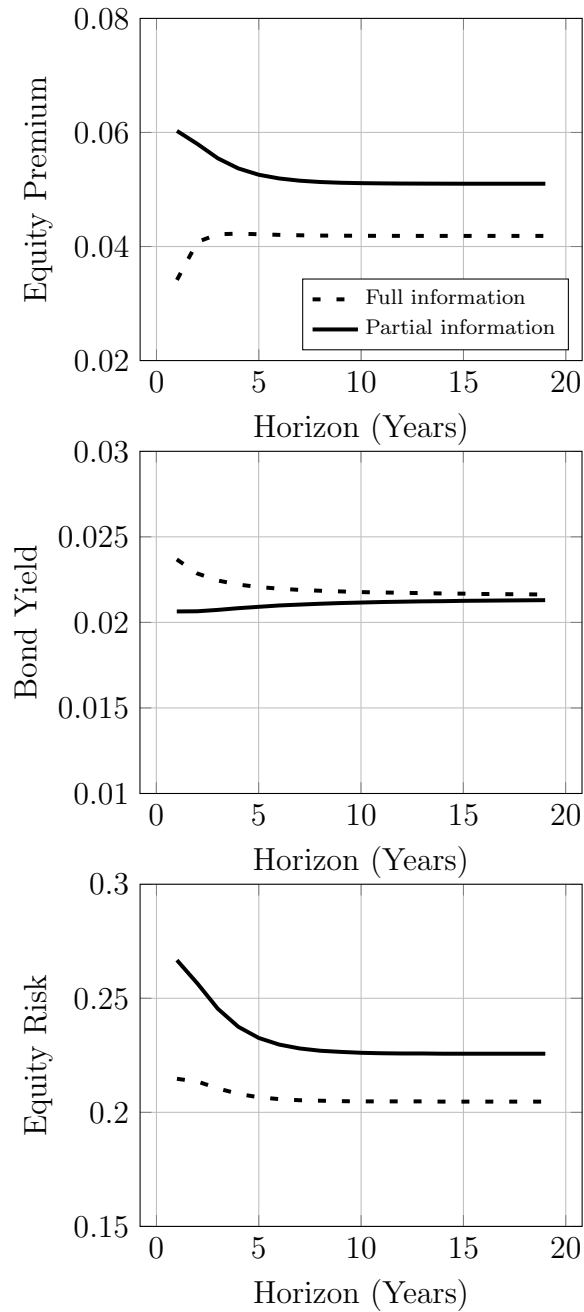
This section studies the term structures of equity risk, equity risk premia, and interest rates. The former are computed using the instantaneous return volatility and risk premium of dividend strips with different maturities. The latter is obtained by computing the yield of zero-coupon bonds with different maturities.

Figure 6 depicts the term structures under both full information and partial information with learning. The upper panel shows that, under full information, the risk premium for the dividend strip with short maturity (e.g. one-year) is close to zero. Risk premia increase sharply up to the three-year maturity and then are flat around a level of about 4.2%. As a consequence of learning, the behavior of the dividend strip risk premium is opposite under partial information. We note that the risk premium at short maturity (e.g. one-year) is about 6.2%. Risk premia decrease uniformly up to the ten-year maturity and then are almost flat around a level of about 5.0%. To summarize, accounting for partial information and learning switches the slope of the term structure of equity risk premia from positive to negative.

The middle panel of Figure 6 shows the zero-coupon bond yields. Under full information we observe that bond yields are downward-sloping, in contrast with actual data. Under partial information we find that learning produces two results which make the model predictions more conform with the data. First, the short-term bond yield is lower than under full information and the slope of the term structure is positive, consistent with TIPs data.

The lower panel of Figure 6 depicts the dividend strip return volatility. Equity risk is downward-sloping under both full information and partial information with learning. In the latter case, the level of volatility is somewhat higher and the negative slope in the first five years is steeper than under full information. This result comes from the fact that the term

Figure 6: Term Structures of Equity Premia, Equity Risk, and Bond Yields.



structure of consumption growth risk is downward-sloping under the partial information (see Figure 5).

4.3.1 Why Does Learning Switch the Slope of the Term Structures?

Under full information the risk premium on the dividend strip with maturity τ is given by²

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

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 dividend strip with maturity τ is given by

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

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 ($\hat{\sigma}_i$).

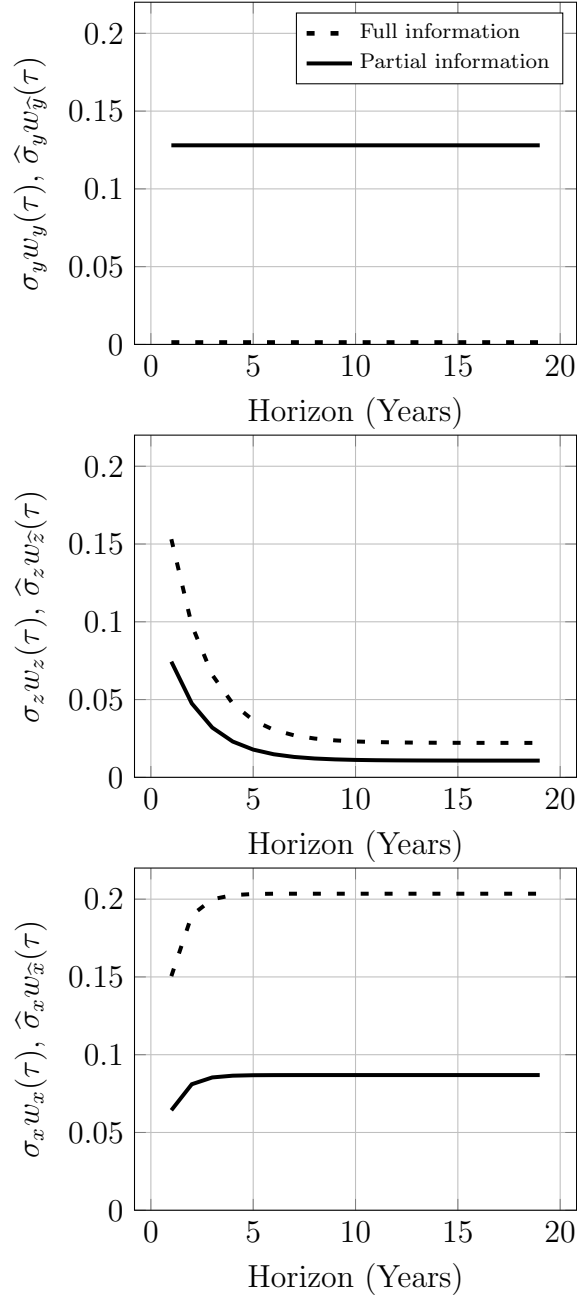
Note that $w_y(\tau)$, $w_x(\tau)$, and $w_z(\tau)$ are respectively constant, increasing, and decreasing with the maturity τ . The same holds for $w_{\hat{y}}(\tau)$, $w_{\hat{x}}(\tau)$, and $w_{\hat{z}}(\tau)$.

The positive slope of the term structure of equity risk premia under full information is due to the fact that the price of risk Λ_x , which compensates for variation in x , is larger than the price of risk Λ_z , which compensates for variation in z . Even if $w_z(\tau)$ is steeper than $w_x(\tau)$, the price of risk Λ_x is large enough to dominate the impact of the transitory shock z . Therefore, the term structure of risk premia is upward-sloping. Note that such a positive slope obtains in the full information case, although the model has been calibrated to match the empirically observed flat term structure to consumption growth risk.

Consider now the partial information case and the role of learning. While the processes y , x , and z have independent increments, learning leads to an endogenous correlation structure between the dynamics of the filtered variables, \hat{y} , \hat{x} , and \hat{z} . More precisely, learning implies

²Details about the full information model are provided in Appendix C.

Figure 7: Price Elasticities.



perfect correlations among these filtered variables because the investor updates her beliefs by observing a single source of information (the history of consumption). In turn, the unique source of risk commands a unique price of risk Λ . Therefore, learning neutralizes the role of the prices of risk in determining the shape of the term structure, and its slope is solely driven by the magnitude and the steepness of the price elasticities. In our case, the range of values

taken by the price elasticity with respect to \widehat{z} (i.e., $\widehat{\sigma}_z w_z(\tau)$) is substantially larger than the range of values taken by the price elasticity with respect to \widehat{x} (i.e., $\widehat{\sigma}_x w_x(\tau)$). Formally, we have

$$\lim_{\tau \rightarrow 0} \widehat{\sigma}_z w_z(\tau) - \lim_{\tau \rightarrow \infty} \widehat{\sigma}_z w_z(\tau) > \lim_{\tau \rightarrow \infty} \widehat{\sigma}_x w_x(\tau) - \lim_{\tau \rightarrow 0} \widehat{\sigma}_x w_x(\tau)$$

when

$$\widehat{\sigma}_z > \frac{\widehat{\sigma}_x}{\lambda_x},$$

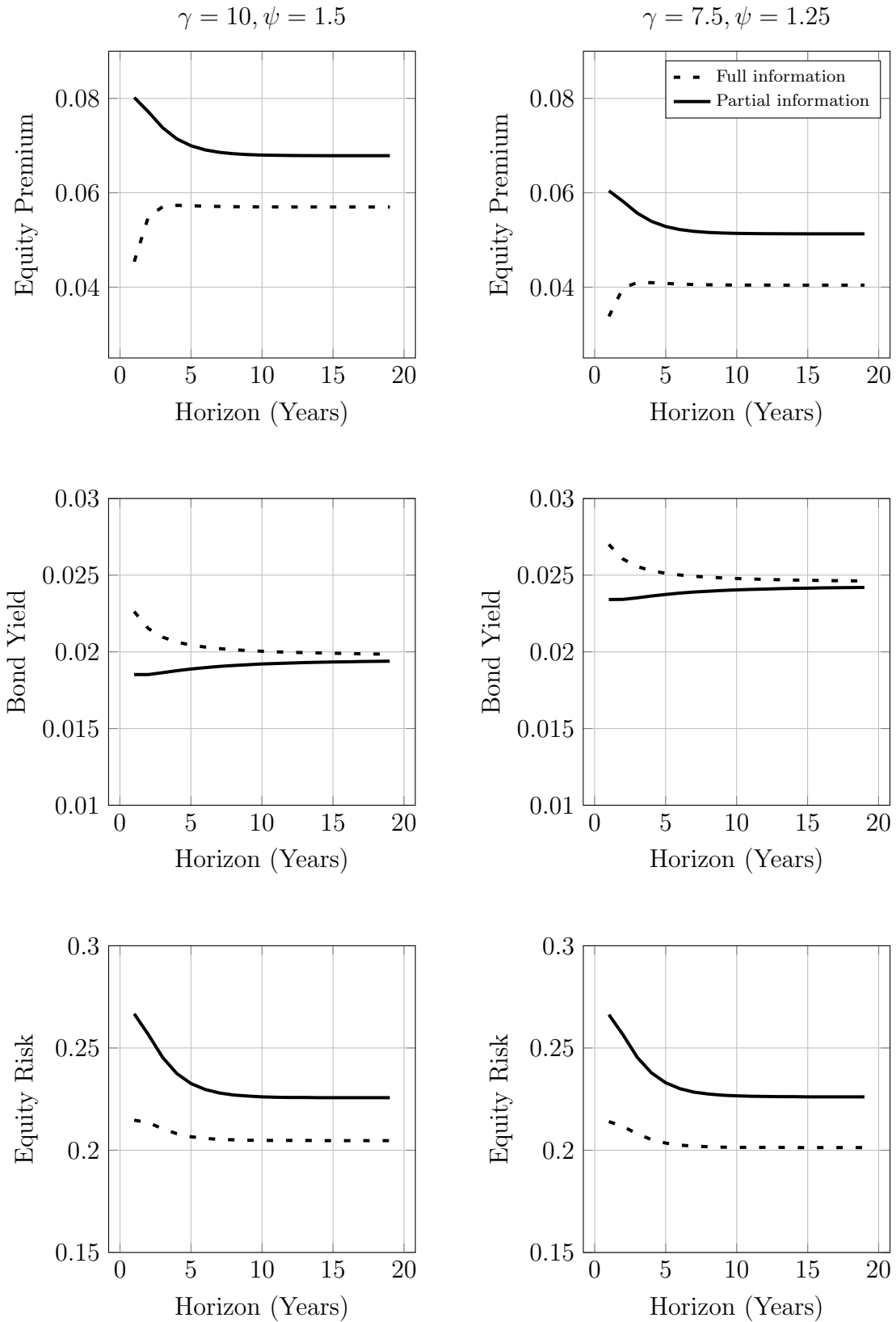
provided that $\phi > 1/\psi$. Therefore, learning leads to a downward-sloping term structure of equity risk premia via its impact on both the quantity of risk across horizons (driven by the perception of downward-sloping consumption growth risk) and the price of risk (driven by the endogenous correlation structure among the filtered variables).

4.3.2 Robustness: The Impact of Preferences

For robustness, we investigate the impact of preference parameters on the term structures of equity risk premia, equity return volatility, and interest rates. We compare the baseline setting ($\gamma = 7.5, \psi = 1.5$) with two alternative parametrizations: ($\gamma = 10, \psi = 1.5$) and ($\gamma = 7.5, \psi = 1.25$). Note that, for each parametrization, the representative agent has a preference for the early resolution of uncertainty.

Figure 8 depicts the term structures. An increase in relative risk aversion increases all the three prices of risk under the full information case as well as the unique price of risk under the partial information case with learning. However, an increase in relative risk aversion does not alter the relative size of price elasticities. In turn, we observe an increase in the level of equity risk premia while preserving the sign switch of the slope of the term structure of equity risk premia implied by learning. That is, with a risk aversion $\gamma = 10$ the model predicts simultaneously a high equity premium (about 7%) and a downward-sloping term structure of dividend strip risk premia. This is a remarkable result of learning, which obtains in absence of any stochastic volatility in the fundamentals; all fundamental processes were

Figure 8: The Impact of Preferences on the Term Structures.



assumed to be homoscedastic for the sake of highlighting the impact of learning on the term structures.

An increase in relative risk aversion decreases slightly the level of the risk-free rate, but it does not alter the impact of learning. The slope of the term structure of bond yields switches from negative to positive under partial information. That is, learning yields an increasing term structure of interest rates, as in the data. Finally, an increase in relative risk aversion does not affect substantially the term structure of dividend strip return volatility because the relative size of price elasticities are weakly sensitive to changes in risk aversion.

Consider now the case of a decrease in the elasticity of intertemporal substitution. Under full information, the price of risk for the permanent component x decreases and the price of risk for the transitory component z increases. Since the former is much larger than the latter in our calibration, the whole effect reduces the level of the equity premium. Under partial information, the unique price of risk decreases. At the same time, price elasticities with respect to x and z move in opposite directions but with the former being smaller than the latter. In turn, learning still switches the sign of the term structure of equity risk premia.

A lower elasticity of intertemporal substitution leads bond yields to be more sensitive to the long-run consumption growth rate. In turn, bond yields are slightly higher than in the baseline economy under both full information and partial information with learning. However, even in this case the effect of learning is preserved. Bond yields are decreasing under full information and increasing under partial information. Finally, the term structure of dividend strip return volatility is barely affected by a decrease in the elasticity of intertemporal substitution. The relative size of price elasticities is unaffected and the downward-sloping shape is driven by the price elasticity with respect to z under both full and partial information.

5 Conclusion

This paper highlights the impact of learning about the expected consumption growth rate on the term structures of equity risk premia and interest rates. While the term structures of equity risk premia and interest rates would be respectively upward-sloping and downward-sloping if the agent had full information about economic fundamentals, partial information and learning implies a switch of sign in the slope of these term structures. Indeed, the fact that the agent faces incomplete information and has to actually filter out unobservable economic fundamentals implies that the term structure of equity risk premia becomes downward-sloping, while the term structure of interest rates becomes upward-sloping in equilibrium. These results show that incomplete information about economic fundamentals and learning are key determinants of the empirically observed shape of the term structures of equity risk premia and interest rates.

References

- Abel, A. B. (1999, February). Risk premia and term premia in general equilibrium. *Journal of Monetary Economics* 43(1), 3–33.
- Bansal, R., R. Dittmar, and D. Kiku (2009, March). Cointegration and consumption risks in asset returns. *Review of Financial Studies* 22(3), 1343–1375.
- 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 (2010, September). Long run risks, the macroeconomy, and asset prices. *American Economic Review* 100(2), 542–46.
- 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., D. Kiku, and A. Yaron (2016). Risks for the long run: Estimation with time aggregation. *Journal of Monetary Economics* 82, 52 – 69.
- 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.
- Beeler, J. and J. Y. Campbell (2012). The long-run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review* 1, 141–182.
- 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.
- Campbell, J., A. Lo, and C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton University Press.
- Campbell, J. Y. and R. J. Shiller (1988). Stock prices, earnings, and expected dividends. *Journal of Finance* 43(3), 661–676.
- Colacito, R. and M. M. Croce (2011). Risks for the long run and the real exchange rate. *Journal of Political Economy* 119(1), 153–181.
- Colacito, R. and M. M. Croce (2013). International asset pricing with recursive preferences. *The Journal of Finance* 68(6), 2651–2686.
- 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.
- 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.

- Drechsler, I. and A. Yaron (2011). What's vol got to do with it. *Review of Financial Studies* 24(1), 1–45.
- 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.
- Gennotte, G. (1986). Optimal portfolio choice under incomplete information. *Journal of Finance* 41(3), 733–46.
- Hasler, M. and R. Marfè (2016). Disaster recovery and the term structure of dividend strips. *Journal of Financial Economics* 122(1), 116 – 134.
- Kreps, D. M. and E. L. Porteus (1979). Temporal von neumann-morgenstern and induced preferences. *Journal of Economic Theory* 20(1), 81–109.
- 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. (2017). Income insurance and the equilibrium term structure of equity. *The Journal of Finance* 72(5), 2073–2130.
- Scheinkman, J. A. and W. Xiong (2003). Overconfidence and speculative bubbles. *The Journal of Political Economy* 111(6), 1183–1219.
- Schorfheide, F., D. Song, and A. Yaron (2018). Identifying long-run risks: A bayesian mixed-frequency approach. *Econometrica* 86(2), 617–654.
- Segal, G., I. Shaliastovich, and A. Yaron (2015). Good and bad uncertainty: Macroeconomic and financial market implications. *Journal of Financial Economics* 117(2), 369 – 397.
- 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.

Weber, M. (2017). Cash flow duration and the term structure of equity returns. *Journal of Financial Economics* (forthcoming).

Weil, P. (1989). The equity premium puzzle and the risk-free rate puzzle. *Journal of Monetary Economics* 24(3), 401–421.

Appendix

A Notation Summary

Parameter	Interpretation
μ	Average growth in log consumption
$\sigma_y > 0$	Volatility of the permanent component of consumption growth
$\lambda_x > 0$	Mean reversion of the stochastic drift of the permanent component
$\sigma_x > 0$	Volatility of of the stochastic drift of the permanent component
$\lambda_z > 0$	Mean reversion of the transitory component
$\sigma_z > 0$	Volatility of of the transitory component
β_d	Parameter that determines the growth rate of dividends
$\phi \geq 1$	Leverage
$\delta \in (0, 1)$	Rate of time preference
$\gamma > 0$	Relative risk aversion
$\psi > 0$	Elasticity of intertemporal substitution

B Proofs

B.1 Proposition 1

Proof. See Liptser and Shiryaev (2001).

The steady-state volatilities $\hat{\sigma}_x$ and $\hat{\sigma}_z$ satisfy

$$\hat{\sigma}_x = \frac{\bar{\gamma}_x - \lambda_z \bar{\gamma}_{xz}}{\sqrt{v}}, \quad \hat{\sigma}_z = \frac{\sigma_z^2 + \bar{\gamma}_{xz} - \lambda_z \bar{\gamma}_z}{\sqrt{v}},$$

where the steady-state posterior variances $\bar{\gamma}_x$ and $\bar{\gamma}_z$, and the steady-state posterior covariance $\bar{\gamma}_{xz}$ solve the following system of equations

$$\begin{aligned} 0 &= \sigma_x^2 - 2\lambda_x \bar{\gamma}_x - v^{-1} (\bar{\gamma}_x - \lambda_z \bar{\gamma}_{xz})^2, \\ 0 &= \sigma_z^2 - 2\lambda_z \bar{\gamma}_z - v^{-1} (\sigma_z^2 - \lambda_z \bar{\gamma}_z + \bar{\gamma}_{xz})^2, \\ 0 &= -(\lambda_x + \lambda_z) \bar{\gamma}_{xz} - v^{-1} (\bar{\gamma}_x - \lambda_z \bar{\gamma}_{xz}) (\sigma_z^2 - \lambda_z \bar{\gamma}_z + \bar{\gamma}_{xz}). \end{aligned}$$

□

B.2 Proposition 2

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

$$\sigma_C^2(t, \tau) = \sigma_y^2 + \frac{\sigma_x^2 e^{-2\lambda_x \tau} (e^{2\lambda_x \tau} (2\lambda_x \tau - 3) + 4e^{\lambda_x \tau} - 1)}{2\lambda_x^3 \tau}. \quad (20)$$

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 consumption volatility in (14), we can compute the agent's estimate of the annualized variance of consumption as

$$\begin{aligned}\hat{\sigma}_C^2(t, \tau) &= \frac{e^{-2\lambda_x \tau}}{2\lambda_x^3 \tau} \left(2\lambda_x \sigma_y \left((e^{\lambda_x \tau} - 1)^2 \sqrt{\lambda_x^2 \sigma_y^2 + \sigma_x^2} + \lambda_x^2 \sigma_y \tau e^{2\lambda_x \tau} - \lambda_x \sigma_y (e^{\lambda_x \tau} - 1)^2 \right) \right. \\ &\quad \left. + \sigma_x^2 (e^{2\lambda_x \tau} (2\lambda_x \tau - 3) + 4e^{\lambda_x \tau} - 1) \right).\end{aligned}\tag{21}$$

Partial derivatives with respect to the horizon τ , volatility σ_x , and the mean-reversion speed

λ_x are as follows:

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

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

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

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. Taking the limits of (20) and (21) as horizon τ approaches zero or infinity gives the result in (15). Furthermore,

$$\widehat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau) = \frac{e^{-2\lambda_x \tau} (e^{\lambda_x \tau} - 1)^2 \left(\sqrt{\sigma_y^2 (\lambda_x^2 \sigma_y^2 + \sigma_x^2)} - \lambda_x \sigma_y^2 \right)}{\lambda_x^2 \tau} > 0. \quad (22)$$

The derivative of the difference in (22) with respect to horizon is

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

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

□

B.4 Proposition 4

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

$$\sigma_C^2(t, \tau) = \sigma_y^2 + \frac{\sigma_z^2 (1 - e^{-2\lambda_z \tau})}{2\lambda_z \tau}. \quad (23)$$

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 and the definition of consumption volatility in (14), we can compute the agent's estimate of the annualized variance of consumption as

$$\begin{aligned} \widehat{\sigma}_C^2(t, \tau) &= \frac{e^{-2\lambda_z \tau}}{2\lambda_z \tau} \left((e^{\lambda_z \tau} - 1) \left(2\sqrt{\sigma_y^2 (\sigma_y^2 + \sigma_z^2)} (e^{\lambda_z \tau} - 1) + \sigma_z^2 (e^{\lambda_z \tau} + 1) \right) \right. \\ &\quad \left. + 2\sigma_y^2 (e^{2\lambda_z \tau} (\lambda_z \tau - 1) + 2e^{\lambda_z \tau} - 1) \right). \end{aligned} \quad (24)$$

Partial derivatives with respect to the horizon τ , volatility σ_z , and the mean-reversion speed

λ_z are as follows:

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

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.5 Proposition 5

Proof. Taking the limits of (23) and (24) as horizon τ approaches zero or infinity gives the result in (17). Furthermore,

$$\hat{\sigma}_C^2(t, \tau) - \sigma_C^2(t, \tau) = \frac{\left(\sqrt{\sigma_y^2 (\sigma_y^2 + \sigma_z^2)} - \sigma_y^2 \right) e^{-2\lambda_z \tau} (e^{\lambda_z \tau} - 1)^2}{\lambda_z \tau} > 0. \quad (25)$$

The derivative of the difference in (25) 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 (18) follows. □

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 $X_t = (\widehat{y}_t, \widehat{y}_{d,t}, \widehat{x}_t, \widehat{z}_t)^\top$, where $\widehat{y}_t \equiv \log C_t - \widehat{z}_t$ and $\widehat{y}_{d,t} \equiv \phi \widehat{y}_t - \beta_d t^3$, so that

$$wc_t \equiv \log \frac{W_t}{C_t} = A + B^\top X_t, \quad (26)$$

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

$$\begin{aligned} dX_t &= \mu(X_t)dt + \Sigma(X_t)d\widehat{B}_t \\ \mu(X_t) &= \mathcal{M} + \mathcal{K}X_t \\ \Sigma(X_t)\Sigma(X_t)^\top &= h + \sum_{i=1}^4 H^i X_t^i, \end{aligned}$$

where $\mathcal{M} \in \mathbb{R}^4$, $\mathcal{K} \in \mathbb{R}^{4 \times 4}$, $h \in \mathbb{R}^{4 \times 4}$, $H \in \mathbb{R}^{4 \times 4 \times 4}$, and \widehat{B}_t is a standard Brownian motion.

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

$$d \log M_t = (\theta \log \delta - (\theta - 1) \log k_1 + (\theta - 1)(k_1 - 1)B'(X_t - \mu_X))dt - \Omega' dX_t, \quad (27)$$

where $X_t = (\widehat{y}_t, \widehat{y}_{d,t}, \widehat{x}_t, \widehat{z}_t)^\top$, $\mu_X = (0, 0, 0, 0)^\top$, $\Omega = \gamma(1, 0, 0, 1)^\top + (1 - \theta)k_1 B$, and the coefficients $A \in \mathbb{R}$ and $B \in \mathbb{R}^4$ are the loadings defined in (26).

The coefficients $A \in \mathbb{R}$, $B \in \mathbb{R}^4$ solve the following system of equations

$$0 = \mathcal{K}^\top \chi - \theta(1 - k_1)B + \frac{1}{2}\chi^\top H \chi, \quad (28)$$

$$0 = \theta(\log \delta + k_0 - (1 - k_1)A) + \mathcal{M}^\top \chi + \frac{1}{2}\chi^\top h \chi, \quad (29)$$

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

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

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

³Note that the dividend dynamics on the full filtration can be equivalently written as

$$d \log D_t = dy_{d,t} + \phi dz_t,$$

where

$$dy_{d,t} = (\mu_d + \phi x_t) dt + \phi \sigma_y dB_{y,t},$$

with $\mu_d \equiv \phi \mu - \beta_d$.

Solving (28) for the vector of loadings $B \in \mathbb{R}^4$ gives

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

Plugging this solution in equation (29) allows to solve for the coefficient A .

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 (27), 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 X_t is the vector of state variables, $\Sigma(X_t) \in \mathbb{R}^{4 \times 1}$ encodes the diffusions of the state variables, vector $\Omega = \gamma(1, 0, 0, 1)^\top + (1 - \theta)k_1 B$ and the coefficients $r_0 \in \mathbb{R}$ and $r_1 \in \mathbb{R}^4$ solve the system of equations

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

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

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

where $\widehat{B}_t^{\mathbb{Q}} = \widehat{B}_t + \int_0^t \Lambda_s ds$ is a \mathbb{Q} -Brownian motion and the coefficients $\mathcal{M}^{\mathbb{Q}} \in \mathbb{R}^4$ and $\mathcal{K}^{\mathbb{Q}} \in \mathbb{R}^{4 \times 4}$ satisfy

$$\mathcal{M}^{\mathbb{Q}} = \mathcal{M} - h\Omega, \quad (30)$$

$$\mathcal{K}^{\mathbb{Q}} = \mathcal{K} - H\Omega. \quad (31)$$

□

B.7 Proposition 7

Proof. Price of a zero-coupon bond can be determined from

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

where $X_t = (\widehat{y}_t, \widehat{x}_t, \widehat{z}_t)^\top$. Eraker and Shaliastovich (2008) show that the functions $q_0(\tau) \in \mathbb{R}$ and $q_1(\tau) \in \mathbb{R}^4$ solve the following system of Riccati equations

$$\frac{\partial}{\partial \tau} q_1(\tau) = -r_1 + \mathcal{K}^{\mathbb{Q}\top} q_1(\tau) + \frac{1}{2} q_1(\tau)^\top H q_1(\tau), \quad (32)$$

$$\frac{\partial}{\partial \tau} q_0(\tau) = -r_0 + \mathcal{M}^{\mathbb{Q}\top} q_1(\tau) + \frac{1}{2} q_1(\tau)^\top h q_1(\tau), \quad (33)$$

with boundary conditions $q_0(0) = 0$ and $q_1(0) = (0, 0, 0, 0)^\top$. Coefficients $\mathcal{M}^{\mathbb{Q}} \in \mathbb{R}^4$ and $\mathcal{K}^{\mathbb{Q}} \in \mathbb{R}^{4 \times 4}$ are characterized in (30)–(31).

Solving (32) gives

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

Using these results in (33) allows to solve for function q_0 .

□

B.8 Proposition 8

Proof. Price of a dividend strip can be determined from

$$S(t, \tau) = \mathbb{E}_t^{\mathbb{Q}} \left(e^{-\int_t^{t+\tau} r_s ds} D_{t+\tau} \right) = e^{w_0(\tau) + w_1(\tau) X_t}.$$

Eraker and Shaliastovich (2008) show that the functions $w_0(\tau) \in \mathbb{R}$ and $w_1(\tau) \in \mathbb{R}^4$ solve the following system of Ricatti equations

$$\frac{\partial}{\partial \tau} w_1(\tau) = -r_1 + \mathcal{K}^{\mathbb{Q}\top} w_1(\tau) + \frac{1}{2} w_1(\tau)^\top H w_1(\tau), \quad (34)$$

$$\frac{\partial}{\partial \tau} w_0(\tau) = -r_0 + \mathcal{M}^{\mathbb{Q}\top} w_1(\tau) + \frac{1}{2} w_1(\tau)^\top h w_1(\tau), \quad (35)$$

with boundary conditions $w_0(0) = 0$ and $w_1(0) = (0, 1, 0, \phi)^\top$. Coefficients $\mathcal{M}^{\mathbb{Q}} \in \mathbb{R}^4$ and $\mathcal{K}^{\mathbb{Q}} \in \mathbb{R}^{4 \times 4}$ are characterized in (30)–(31).

Solving (34) gives

$$w_1(\tau)^\top = \left(0, 1, \frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}) (\phi \psi - 1), \frac{1}{\psi} (1 - e^{-\lambda_z \tau} (1 - \phi \psi)) \right).$$

Using these results in (35) allows to solve for function w_0 .

□

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 equilibrium price-dividend ratio is linear in the state variables,

$$pd_t = \equiv \log \frac{P_t}{D_t} = A_d + B_d^\top X_t.$$

The coefficients $A_d \in \mathbb{R}$, $B_d \in \mathbb{R}^4$ solve the following system of equations

$$0 = \mathcal{K}^\top \chi_d + (\theta - 1)(k_1 - 1)B + (k_{1,d} - 1)B_d + \frac{1}{2} \chi_d^\top H \chi_d, \quad (36)$$

$$0 = \theta \ln \delta - (\theta - 1) (\ln k_1 + (k_1 - 1)B^\top \mu_X) - (\ln k_{1,d} + (k_{1,d} - 1)B_d^\top \mu_X) + \mathcal{M}^\top \chi_d + \frac{1}{2} \chi_d^\top h \chi_d, \quad (37)$$

where $\chi_d = (0, 1, 0, \phi)^\top + k_{1,d} B_d - \Omega$ and $k_{1,d} \in \mathbb{R}$ is the linearization coefficient for the stock return.

Solving (36) for the vector of loadings $B_d \in \mathbb{R}^4$ gives

$$B_d^\top = \left(0, 0, \frac{\phi - 1/\psi}{1 - k_{1,d}(1 - \lambda_x)}, -\frac{\lambda_z(\phi - 1/\psi)}{1 - k_{1,d}(1 - \lambda_z)} \right).$$

Plugging this solution in equation (37) allows to solve for $k_{1,d}$. Then we obtain the intercept

A_d as

$$A_d = \log \frac{k_{1,d}}{1 - k_{1,d}} - B_d^\top \mu_X.$$

□

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 $B_t = (B_{y,t}, B_{x,t}, B_{z,t})^\top$. 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

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

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 $X_t = (y_t, y_{d,t}, x_t, z_t)^\top$, where $y_{d,t} \equiv \phi y_t - \beta_d t$ and use the fact that the state variables belong to the affine class, so that their dynamics can be written as:

$$\begin{aligned} dX_t &= \mu(X_t) dt + \Sigma(X_t) dB_t \\ \mu(X_t) &= \mathcal{M} + \mathcal{K} X_t \\ \Sigma(X_t) \Sigma(X_t)^\top &= h + \sum_{i=1}^4 H^i X_t^i. \end{aligned}$$

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 (30)-(31).

□

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

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

where

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

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

Proof. Analogous to the proof of Proposition 7.

□

Proposition C.3. *The equilibrium price of the dividend strip with time to maturity τ in the full information economy is given by*

$$S(t, \tau) = \mathbb{E} \left[\frac{M_{t+\tau}}{M_t} D_{t+\tau} \mid \mathcal{F}_t \right] = e^{-\beta_d t + \phi y_t + w_0(\tau) + w_x(\tau)x_t + w_z(\tau)z_t},$$

where

$$\begin{aligned} w_x(\tau) &= \frac{1}{\lambda_x \psi} (1 - e^{-\lambda_x \tau}) (\phi \psi - 1) \\ w_z(\tau) &= \frac{1}{\psi} (1 - e^{-\lambda_z \tau} (1 - \phi \psi)). \end{aligned}$$

and $w_0(\tau)$ solves

$$\frac{\partial}{\partial \tau} w_0(\tau) = -r_0 + \mathcal{M}^{\mathbb{Q}\top} w_1(\tau) + \frac{1}{2} w_1(\tau)^\top h w_1(\tau)$$

with $w_1(\tau)^\top = (0, 1, w_x(\tau), w_z(\tau))$ and $w_0(0) = 0$. The return premium of the dividend strip with time to maturity τ is given by

$$RP(t, \tau) = -\frac{1}{dt} \left\langle \frac{dM_t}{M_t}, \frac{dS(t, \tau)}{S(t, \tau)} \right\rangle = (\phi\sigma_y, w_x(\tau)\sigma_x, w_z(\tau)\sigma_z)^\top \Lambda.$$

The return volatility of the dividend strip with time to maturity τ is given by

$$Vol(t, \tau) = \sqrt{\frac{1}{dt} \left\langle \frac{dS(t, \tau)}{S(t, \tau)} \right\rangle} = \|(\phi\sigma_y, w_x(\tau)\sigma_x, w_z(\tau)\sigma_z)^\top\|,$$

where $\|\cdot\|$ is the Euclidean norm.

Proof. Analogous to the proof of Proposition 8. □

Proposition C.4. The equilibrium price of equity in the full information economy is given by

$$P_t = \int_0^\infty \mathbb{E}_t \left[\frac{M_{t+\tau}}{M_t} D_{t+\tau} \mid \mathcal{F}_t \right] d\tau \approx D_t e^{A_d + B_{x,d}x_t + B_{z,d}z_t},$$

where

$$B_{x,d} = \frac{\phi - 1/\psi}{1 - k_{1,d}(1 - \lambda_x)},$$

$$B_{z,d} = -\frac{\lambda_z(\phi - 1/\psi)}{1 - k_{1,d}(1 - \lambda_z)}.$$

and A_d is satisfies

$$A_d = \log \frac{k_{1,d}}{1 - k_{1,d}} - B_d^\top \mu_X,$$

where $B_d^\top = (0, 0, B_{x,d}, B_{z,d})$ and the linearization coefficient $k_{1,d}$ solves a full information analogue of (37). The return premium of equity is given by

$$RP(t) = -\frac{1}{dt} \left\langle \frac{dM_t}{M_t}, \frac{dP_t}{P_t} \right\rangle = (\phi\sigma_y, B_{x,d}\sigma_x, (\phi + B_{z,d})\sigma_z)^\top \Lambda.$$

The return volatility of equity is given by

$$Vol(t) = \sqrt{\frac{1}{dt} \left\langle \frac{dP_t}{P_t} \right\rangle} = \|(\phi\sigma_y, B_{x,d}\sigma_x, (\phi + B_{z,d})\sigma_z)^\top\|,$$

where $\|\cdot\|$ is the Euclidean norm.

Proof. Analogous to the proof of Proposition 9. □