NBER WORKING PAPER SERIES

INVESTOR INFORMATION, LONG-RUN RISK, AND THE TERM STRUCTURE OF EQUITY

Mariano M. Croce
Martin Lettau
Sydney C. Ludvigson

Working Paper 12912
http://www.nber.org/papers/w12912

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2007

Previously circulated as "Investor Information, Long-Run Risk, and the Duration of Risky Cash-Flows." This material is based upon work supported by the National Science Foundation under Grant No. 0617858 to Lettau and Ludvigson. Ludvigson also acknowledges financial support from the Alfred P. Sloan Foun- dation and the CV Starr Center at NYU. The authors thank Dave Backus, John Y. Campbell, Timothy Cogley, Michael Gallmeyer, Lars Hansen, John Heaton, Dana Kiku, Thomas Sargent, Jay Shanken, Stijn Van Nieuwerburgh and seminar participants at the 2006 Society for Economic Dynamics conference, the summer 2006 NBER Asset Pricing meeting, the Western Finance Association 2007 meetings, the American Finance 2009 meetings, Emory, NYU, Texas A&M, UCLA, and UNC Chapel Hill for helpful comments. Any errors or omissions are the responsibility of the authors. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Investor Information, Long-Run Risk, and the Term Structure of Equity
Mariano M. Croce, Martin Lettau, and Sydney C. Ludvigson
NBER Working Paper No. 12912
February 2007, Revised August 2014
JEL No. E44,G10,G12

ABSTRACT

We study the role of information in asset pricing models with long-run cash flow risk. When investors can distinguish short- from long-run consumption risks (full information), the model generates a sizable equity risk premium only if the equity term structure slopes up, contrary to the data. In general, the short- and long-run components are unidentified. We propose a sparsity-based bounded rationality model of long-run risk that is both parsimonious and fully identified from historical data. In contrast to full information, the model generates a sizable market risk premium simultaneously with a downward sloping equity term structure, as in the data.

Mariano M. Croce
Kenan-Flagler Business School,
UNC at Chapel Hill
mmc287@gmail.com

Martin Lettau
Haas School of Business
University of California, Berkeley
545 Student Services Bldg. #1900
Berkeley, CA 94720-1900
and NBER
lettau@haas.berkeley.edu

Sydney C. Ludvigson
Department of Economics
New York University
19 W. 4th Street, 6th Floor
New York, NY 10002
and NBER
sydney.ludvigson@nyu.edu

1 Introduction

We study the role of information in asset pricing models with long-run cash flow risk. The idea that long-run cash flow risk can have important affects on asset prices is exemplified by the work of Bansal and Yaron (2004), who show that a small but persistent common component in the time-series processes of consumption and dividend growth is capable of generating large risk premia and high Sharpe ratios.[1]

A crucial aspect of the long-run risk theory is that the persistent component in consumption growth must be small. That is, it must account for only a small fraction of its short-run variability. If this were not the case, the model-implied annualized volatilities of consumption and dividend growth, as well as the correlation between the two, would be implausibly large and both series implausibly persistent. But with the persistent component necessarily small, it is (both in theory and in empirical work) difficult to detect statistically, even in large samples. Despite this, a maintained assumption in the theoretical literature is that investors can directly observe the persistent component and distinguish its innovations from the more volatile transitory shocks to consumption and dividend growth. We refer to this assumption as the *full information* paradigm. While this is a natural starting place and an important case to understand, in this paper we consider an alternative specification in which market participants face a signal extraction problem: they can observe the change in consumption and dividends each period, but they cannot observe the individual components of that change.

Information about long-run cash flow risk is likely to be limited. Indeed, in reality the long-run components are unobservable and the parameters of a general dynamic system for

[1] An extensive literature has followed this work. See Parker (2001); Parker and Julliard (2004); Colacito and Croce (2004); Bansal, Dittmar and Kiku (2009); Hansen, Heaton and Li (2008); Kiku (2005); Malloy, Moskowitz and Vissing-Jorgensen (2009); Bansal, Dittmar and Lundblad (2005); Bansal, Kiku and Yaron (2007); Hansen and Sargent (2006).

consumption and dividend growth that embeds these components cannot be identified from historical data using standard estimation techniques. Thus, the full information assumption takes the amount of information investors have very seriously: market participants must not only understand that a small predictable component in cash flow growth exists, they must also be able to decompose each period's innovation into its component sources and have complete knowledge of how the shocks to these sources vary and co-vary with one another, even though the data give us no guide for identifying these components separately.

We propose an alternative model of behavior in which a representative decision maker optimizes based on a cash flow model that is sparse in the sense that it ignores cross-equation restrictions that are difficult if not impossible to infer in finite samples, and fully identified from historical data. We refer to this specification as the *bounded rationality limited information* model. The cash flow model serves as a signal extraction tool, allowing the investor to form an estimate of the long-run components in dividend and consumption growth, given historical data.

As an illustration of the potential importance of the information structure, we study the implications of models with long-run consumption risk for jointly matching evidence for a sizable equity risk premium simultaneously with a downward sloping term structure of equity. The term structure of aggregate equity may be computed by recognizing that an equity index claim is a portfolio of zero-coupon dividend claims (strips) with different maturities. A downward sloping equity term structure means that expected excess returns on strips that pay dividends in the near future are higher than that for the market index, an average over all strips.

We study a cash flow model in which the aggregate dividend growth rate is differentially exposed to two systematic risk components driven by aggregate consumption growth, in addition to a purely idiosyncratic component uncorrelated with aggregate consumption. One systematic risk component is a small but highly persistent (long-run) component in consumption growth as in Bansal and Yaron (2004), while the second is a transitory (short-run) i.i.d. component with much larger variance. The purely idiosyncratic component is volatile, as required to match the evidence that dividend growth is substantially more volatile than consumption growth and not highly correlated with it. In addition, we follow the existing literature on long-run risk by employing the recursive utility specification developed by Epstein and Zin (1989, 1991) and Weil (1989). Under standard parameterizations of the utility function, investors have a preference for early resolution of uncertainty, implying that shocks to the priced long-run component of cash flows command large risk premia even if they are far less volatile over short horizons than are shocks to the priced short-run component.

We use the framework just described to study the term structure of equity. We show that, with full information, expected excess returns on strips that pay a dividend in the far future are higher than the market risk premium, implying an upward sloping term structure, contrary to the historical data. Specifically, the full information paradigm cannot simultaneously generate both a high equity risk premium and a downward sloping term structure, unless the long-run component in dividend growth is a source of *insurance,* rather than risk. Insurance means that innovations in the long-run component (holding other shocks fixed) generate positive covariance between the pricing kernel and returns, so that the long-run component generates a negative risk premium. In a long-run insurance model, more than 100% of the equity premium must be attributable to short-run consumption risk, and under reasonable parameterizations of the magnitude of this risk, the equity premium is small and the level of the term structure too low. We use approximate analytical solutions to the model to show that this is result is quite general. By contrast, under the bounded rationality limited information model, the equity risk premium can be sizable while the equity term structure slopes down even if, under full information, the long-run component in dividend growth would be a source of risk, rather than insurance.

The intuition for this result is as follows. When investors with a preference for early resolution of uncertainty can observe the long-run component in cash flows–in which a small shock today has a large impact on long-run growth rates–the long-run is correctly inferred to be more risky than the short-run, implying that long maturity equity strips must in equilibrium command high risk premia. By contrast, short maturity strips command low risk premia because they depend on exposure to the short-run consumption shock, which does not generate as large a risk premium as the long-run shock under standard calibrations.[2]

Under the bounded rationality limited information model proposed here, the opposite can occur because the decision maker's estimate of the long-run component in dividend growth will be "contaminated" by shocks to the i.i.d. components in consumption and dividend growth (including the idiosyncratic component), which cannot be distinguished from shocks to the persistent components. Shocks to the short-run component in consumption growth generate high risk premia on short maturity strips, both because the exposure of dividend growth to the short-run consumption shock contributes positively to the risk premium on the one-period strip even under full information, and because the decision maker erroneously revises upward her estimate of the long-run component in consumption and dividend growth in response to short-run (i.i.d.) shocks. This latter effect makes short-run shocks appear more risky under the bounded rationality limited information model than under full information, leading to higher risk premia on short maturity strips under the former than the latter.

On the other hand, long maturity strips have low risk premia under the bounded rationality limited information model because shocks to the persistent component in consumption growth (which drive the persistent component in dividend growth) are too small to be distinguished from the large idiosyncratic dividend shocks. Dividend growth as a whole appears close to i.i.d. and shocks to the long-run component are close to being unpriced under the

[2] In Section 4 we consider general calibrations of the full information model, including non-standard ones. Greater exposure to short-run risk can raise the premium on the one-period strip, but longer maturity strips will have lower risk premia only if the model is one of long-run insurance rather than long-run risk.

bounded rationality limited information model, even though they command substantial risk premia under full information. The end result is that long maturity strips have low risk premia under the bounded rationality limited information model, while they have high risk premia under full information.

The rest of this paper is organized as follows. The next section discusses related literature discussed here, as well as the empirical evidence for a downward sloping equity term structure. Section 3 presents the asset pricing model, the model for cash flows, and the information assumptions. Section 4 presents the theoretical results on the equity premium, level and slope of the term structure, beginning with approximate log-linear analytical solutions to the full and bounded rationality limited information models. We then move on to illustrate the role of the information structure in driving equilibrium outcomes. Section 5 concludes the paper.

2 Related Literature

A growing body of literature documents evidence that the term structure of the stock market is downward sloping (e.g., van Binsbergen, Brandt and Koijen (2010), Ang and Ulrich (2011), van Binsbergen, Hueskes, Koijen and Vrugt (2012), and Boguth, Carlson, Fisher and Simutin (2011)). As mentioned, a downward sloping equity term structure means that risk premia on strips that pay a dividend in the near future are higher than that for the market index, an average over all strips. The magnitude of these negatives spreads found in the data is substantial. These findings are consistent with those showing that short duration individual stocks that make up the equity index have higher expected returns than long duration individual stocks (Cornell (1999, 2000); Dechow, Sloan and Soliman (2004); Da (2005); van Binsbergen et al. (2010)).

A number of recent papers address issues related to those studied here. Hansen and Sargent (2006), also concerned about the agent's ability to observe the long-run risk component in aggregate cash flows, study these models in a robust control framework but do not study implications for the equity term structure.

Our paper is related to a recent literature that seeks to reconcile the cross-sectional properties of equity returns simultaneously with the cash flow duration properties of value and growth assets (Lettau and Wachter (2007) and Lettau and Wachter (2009)).[3] None of these studies investigate the role of the information structure on asset prices, a focus of this paper.

Finally, our work builds on an earlier literature that studies the effect of information quality and learning on asset prices. Since the cash-flow specifications in our model can be represented by linear Gaussian state space models, the filtering problem our agents solve is a special case of Bayesian updating used in many learning models, such as Veronesi (2000), Li (2005), Ai (2010), and Johannes, Lochstoer and Mou (2013). All of these models consider learning about the cash-flow or production process, but they differ in their modeling of preferences and/or technologies and/or the specific variables about which agents learn. In fact, Veronesi (2000), Li (2005), and Ai (2010) are special cases of our model.[4] None of these papers focus on the implications of learning for the term structure of equity, the goal of this

[3]Two other papers study duration indirectly. Lustig and Van Nieuwerburgh (2006) study a model with heterogenous agents and housing collateral constraints and find that conditional expected excess returns are hump-shaped in their measure of duration. Zhang (2005) shows that, when adjustment costs are asymmetric and the price of risk varies over time, growth assets can be less risky than assets in place (value stocks), consistent with the cash flow and return properties of value and growth assets. But the Zhang model does not account for the finding of Fama and French (1992) that value stocks do not have higher CAPM betas than growth stocks.

[4]Veronesi (2000) and Li (2005) study endowment economies with constant relative risk-aversion utility (a special case of our recursive preferences), while Ai (2010) and Johannes et al. (2013) study economies where agent's have recursive preferences, but with consumption equal to dividends (a special case of our model where consumption and dividend growth are imperfectly correlated).

paper.

Although not modeled explicitly as such, we can think of our information friction about the consumption process as the result of limited information about deeper variables of the economy. Examples include Kaltenbrunner and Lochstoer (2010), where the persistent component of consumption growth arises from consumption smoothing, or Croce (2014), where there is a persistent component in the growth of TFP. In both studies, the primitive shock is a technology shock and limited information could arise about the technology process itself, which could be subject to shocks with different degrees of persistence in either the first or second moments.

3 The Asset Pricing Model

Consider a representative agent who maximizes utility defined over aggregate consumption. To model utility, we use the more flexible version of the power utility model developed by Epstein and Zin (1989, 1991) and Weil (1989), also employed by other researchers who study the importance of long-run risks in cash flows (Bansal and Yaron (2004), Hansen et al. (2008) and Malloy et al. (2009)).

Let C_t denote consumption and $R_{C,t}$ denote the simple gross return on the portfolio of all invested wealth, which pays C_t as its dividend. The Epstein-Zin-Weil objective function is defined recursively as:

$$U_t = \left[(1-\delta)C_t^{\frac{1-\gamma}{\theta}} + \delta\left(E_t\left[U_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{\theta}}\right]^{\frac{\theta}{1-\gamma}}$$

where γ is the coefficient of risk aversion and the composite parameter $\theta = \frac{1-\gamma}{1-1/\Psi}$ implicitly defines the elasticity of intertemporal substitution (EIS) Ψ.

Let D_t denote the aggregate dividend at time t, and let P_t^D denote the ex-dividend price measured at the end of time t of a claim to the asset that pays the aggregate dividend stream $\{D_t\}_{t=1}^{\infty}$. Let P_t^C denote the ex-dividend price of a claim to the aggregate consumption stream. From the first-order condition for optimal consumption choice and the definition of returns

$$E_t\left[M_{t+1}R_{c,t+1}\right] = 1, \qquad R_{c,t+1} = \frac{P_{t+1}^{C} + C_{t+1}}{P_t^{D}} \tag{1}$$

$$E_t\left[M_{t+1}R_{d,t+1}\right] = 1, \qquad R_{d,t+1} = \frac{P_{t+1}^{D} + D_{t+1}}{P_t^{D}} \tag{2}$$

where M_{t+1} is the stochastic discount factor (SDF), given under Epstein-Zin-Weil utility as

$$M_{t+1} = \left(\delta\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{1}{\psi}}\right)^{\theta} R_{c,t+1}^{\theta-1}. \tag{3}$$

The return on a one-period risk-free asset whose value is known with certainty at time t is given by $R_{f,t+1} \equiv \left(E_t\left[M_{t+1}\right]\right)^{-1}$.

3.1 The Cash Flow Model

Equities are modeled as claims to the aggregate dividend stream. We are interested in a model for equity cash flows that allows dividend growth rates to be potentially exposed to both transitory and persistent sources of consumption risk, as well as to purely idiosyncratic shocks that command no risk premium. We use lower case letters denote log variables, e.g., $\log(C_t) \equiv c_t$. Denote the conditional means of the log difference in consumption and dividends as $x_{c,t}$ and $x_{d,t}$, respectively. Consider a general system of equations for log consumption, c_t, and log dividends, d_t, taking the form

$$\Delta c_{t+1} = \mu_c + \underbrace{x_{c,t}}_{\text{LR risk}} + \sigma \underbrace{\varepsilon_{c,t+1}}_{\text{SR risk}} \tag{4}$$

$$\Delta d_{t+1} = \mu_d + x_{d,t} + \phi_c \sigma \varepsilon_{c,t+1} + \sigma_d \sigma \varepsilon_{d,t+1} \tag{5}$$

$$x_{c,t} = \rho x_{c,t-1} + \sigma_{xc} \sigma \varepsilon_{xc,t} \tag{6}$$

$$x_{d,t} = \rho_d x_{d,t-1} + \sigma_{xd} \sigma \varepsilon_{xd,t} \tag{7}$$

$$\left(\varepsilon_{c,t+1}, \varepsilon_{d,t+1}, \varepsilon_{xc,t}, \varepsilon_{xd,t}\right) \sim N.i.i.d\left(\mathbf{0}, \mathbf{\Omega}\right), \tag{8}$$

where $\mathbf{\Omega}$ is an unrestricted symmetric covariance matrix,

$$\mathbf{\Omega} = \begin{bmatrix} 1 & \rho_{c,d} & \rho_{c,xc} & \rho_{c,xd} \\ \rho_{c,d} & 1 & \rho_{d,xc} & \rho_{d,xd} \\ \rho_{c,xc} & \rho_{d,xc} & 1 & \rho_{xc,xd} \\ \rho_{c,xd} & \rho_{d,xd} & \rho_{xc,xd} & 1 \end{bmatrix},$$

where $\rho_{i,j}$ denotes the correlation coefficient between the shocks ε_i and ε_j.

We use the term short-run "SR" risk to refer to the i.i.d. consumption shock, and long-run "LR" risk to refer to the persistent conditional mean $x_{c,t}$. This "long-run" terminology is used in the literature because even small innovations in $x_{c,t}$, if sufficiently persistent, will have large affects on cash flows in the long-run, resulting in high risk premia when investors prefer early resolution of uncertainty. The model of Bansal and Yaron (2004) is a special case of the system (4)-(8) in which consumption and dividend growth contain a single, common predictable component:

$$\Delta c_{t+1} = \mu_c + \underbrace{x_{c,t}}_{\text{LR risk}} + \sigma \underbrace{\varepsilon_{c,t+1}}_{\text{SR risk}} \tag{9}$$

$$\Delta d_{t+1} = \mu_d + x_{d,t} + \phi_c \sigma \varepsilon_{c,t+1} + \sigma_d \sigma \varepsilon_{d,t+1} \tag{10}$$

$$x_{c,t} = \rho x_{c,t-1} + \sigma_{xc} \sigma \varepsilon_{xc,t} \tag{11}$$

$$x_{d,t} = \phi_x x_{c,t} \tag{12}$$

$$\varepsilon_{c,t+1}, \varepsilon_{d,t+1}, \varepsilon_{xc,t} \sim N.i.i.d\left(0, 1\right), \tag{13}$$

This special case imposes the parameter restrictions $\rho_d = \rho$, $\sigma_{xd} = \phi_x \sigma_{xc}$, $\varepsilon_{xd,t} = \varepsilon_{xc,t}$, $x_{c,0} = x_{d,0} = 0$, $\rho_{xc,xd} = 1$, and $\rho_{c,d} = \rho_{c,xc} = \rho_{c,xd} = \rho_{d,xc} = \rho_{d,xd} = 0$. The Bansal and Yaron (2004) specification also sets $\phi_c = 0$ so that dividend growth is not exposed to short-run consumption risk.

The representative agent in the Bansal and Yaron (2004) is assumed to have the power to observe both the conditional means $x_{c,t}$ and $x_{d,t}$ as well as all the parameters of the cash flow process (4)-(8). These assumptions imply that the agent takes the cash flow model (9)-(13) as given, if the true data-generating process embeds the appropriate parameter restrictions. We refer to this assumption as *full information* ("FI" for short). In reality, the variables $x_{c,t}$ and $x_{d,t}$ are unobservable and the parameters of the general system (4)-(8) cannot be identified from historical data using standard estimation techniques (see below). We define *limited information* as a state of the world in which this reality holds for the decision maker in the model: investors cannot directly observe the latent variables $x_{c,t}$ and $x_{d,t}$ and they do not know, nor can they identify from data, the parameters of the data generating cash flow process.

3.2 Limited Information

We assume that investors in a limited information state of the world can observe all historical data, even asset price information. Because asset prices are endogenous outcomes conditional on the information in cash flows, asset price data are redundant once the information in historical consumption and dividend data has been taken into account. Therefore adding asset return data to the information set that includes consumption and dividend data leads to the same equilibrium allocations as the model where the information set includes only the history consumption and dividend data.

Armed with historical data on dividends and consumption, how could one estimate the parameters of the system together with estimates of the latent variables $x_{c,t}$ and $x_{d,t}$? A standard approach would be to write the dynamic system (4)-(8) in state space form and apply Maximum Likelihood estimation simultaneously with the Kalman filter to estimate both the latent state variables $x_{c,t}$ and $x_{d,t}$ and the parameters of the general dynamic system (4)-(8). Without further restrictions on the parameter space, however, the system

(4)-(8) is unidentified. That is, more than one set of parameter values can give rise to the same value of the likelihood function and the data give no guide for choosing among these.

To see why note that, given historical data and knowledge of the system (4)-(8), an investor with limited information can estimate the Wold representation for this system, which will exist as long as Δc_{t+1} and Δd_{t+1} follow covariance-stationary processes. For arbitrary parameter values (limiting to stationarity of the data), the dynamic system (4)-(7) has a Wold representation that is a first-order vector autoregressive-moving average process, or $VARMA\left(1,1\right)$:

$$\begin{bmatrix} \Delta c_{t+1} \\ \Delta d_{t+1} \end{bmatrix} = \begin{bmatrix} \mu_c\left(1-\rho\right) \\ \mu_d\left(1-\rho_d\right) \end{bmatrix} + \begin{bmatrix} \rho & 0 \\ 0 & \rho_d \end{bmatrix} \begin{bmatrix} \Delta c_t \\ \Delta d_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v^V_{c,t+1} \\ v^V_{d,t+1} \end{bmatrix} - \underbrace{\begin{bmatrix} b_{cc} & b_{cd} \\ b_{dc} & b_{dd} \end{bmatrix}}_{\mathbf{b}} \begin{bmatrix} v^V_{c,t} \\ v^V_{d,t} \end{bmatrix}. \tag{14}$$

The parameters ρ, ρ_d, b_{cc},..., b_{dd} and as well as those in the variance-covariance matrix of $v_{c,t+1}$ and $v_{d,t+1}$ are functions of the deep parameters of the cash flow system (4)-(8). The system (4)-(8) has 15 unknown parameters (including six unknown parameters in $\mathbf{\Omega}$). Estimation of (14) identifies 11 parameters (including three from the covariance matrix of the VARMA innovations), four short of what's needed for identification.[5] That is, given an infinite sample of data on consumption and dividend growth, the parameters of the dynamic system (4)-(8) can only be observed in certain combinations as the estimates ρ, ρ_d, b_{cc},..., b_{dd} and the variance-covariance matrix of $v_{c,t+1}$ and $v_{d,t+1}$, and this information is not enough to separately identify the deep parameters of (4)-(8).

As we demonstrate in the last section of the paper, under common parameterizations of the long-run risk cash-flow process, a limited information VARMA model generates an equity term structure slope that is very similar to the full information case, counterfactually

[5] One could restrict $\rho_{d,c}$ and $\rho_{d,xc}$ to zero to account for the fact that one shock to dividends captures purely idiosyncratic risk. In this case, the full system is underidentified by two parameters rather than four.

implying that it slopes up. Quite different implications for the term structure can be generated, however, if the off-diagonal elements of the **b** matrix are presumed to be zero, so that the system (14) collapses to a pair of first-order univariate autoregressive moving average (ARMA(1,1)) processes, each with 4 free parameters:

$$\Delta c_{t+1} = \mu_c (1-\rho) + \rho \Delta c_t + v^A_{c,t+1} - b_c v^A_{c,t} \tag{15}$$

$$\Delta d_{t+1} = \mu_d (1-\rho_d) + \rho_d \Delta d_t + v^A_{d,t+1} - b_d v^A_{d,t}. \tag{16}$$

Under parameter values typically employed in the long-run risk literature, the off-diagonal elements b_{cd} and b_{dc} are in fact close to zero and impossible to distinguish from zero with statistical tests, both in samples of the size currently available as well as in samples considerably larger.[6] In the data, consumption and dividend growth are only modestly correlated, and dividend growth is considerably more volatile (Table 1). Typical parameterizations of long-run risk models are calibrated to match these facts, so the idiosyncratic component of dividend growth is specified as highly volatile. This is the primary reason why the off-diagonal elements of **b** are so small in benchmark long-run risk models.

3.2.1 A Bounded Rationality Limited Information Model

With these implications of long-run risk models in mind, we propose a bounded rationality model of behavior in which the decision maker considers a simplified representation of (14) that is "sparse" in the sense that the small and difficult to infer off-diagonal elements of

[6] This statement is confirmed by Monte Carlo experiments. Specifically, under the benchmark calibration in column 3 of Table 1 (discussed below), the population (large sample) values for b_{cd} and b_{dc} 0.004 and -0.038, respectively. In finite samples, there is a significant downward bias in b_{cd} but the standard errors are large: the average (across 1000 artificial samples of size equal to that of our historical dataset) maximum likelihood point estimates (standard errors) for b_{cd} and b_{dc} are 0.005 (0.004) and -0.126 (0.111), respectively.

the **b** matrix are set to zero.[7] Even if the true data generating process implies small but non-zero values for the off-diagonal elements of the **b** matrix, this sparsity is nevertheless optimal in a forecasting sense, as explained below.

In the present context, it is the off-diagonal elements of the **b** matrix where the natural sparsity arises, since these values will necessarily be very small with standard error bands that include zero. We assume that the decision maker sets those elements to zero and estimates two univariate ARMA(1,1) specifications, one each for consumption and dividend growth. The ARMA parameters are functions of the primitive parameters of the dynamic system (4)-(7). The innovations $v^A_{c,t+1}$ and $v^A_{d,t+1}$ are in general correlated and are composites of the underlying innovations in (4)-(7). We refer to this model of behavior as the *bounded rationality limited information* model hereafter ("BRLI" for short), to emphasize that this specification embeds both a change in the information structure and a behavioral assumption, vis-a-vis the full information model.[8]

One way to interpret this restriction on the **b** matrix is to recognize that movements in consumption growth comprise too small a part of the volatility of dividends (given the noise created by the large idiosyncratic component) to be informative, so the tiny contemporaneous correlation between the VARMA innovations $v^V_{c,t}$ and $v^V_{d,t}$ is effectively ignored by setting the off-diagonal parameters to zero.

This model of behavior may be motivated by statistical considerations. Given the small

[7] Different forms of sparsity-based bounded rationality models have been proposed in the literature. See, for example, Gabaix (2011).

[8] Anderson, Hansen and Sargent (1998) study risk premia for a claim to aggregate consumption in an asset pricing model where the true data generating process for consumption growth follows an $ARMA(1,1)$. We note that if the true data generating process were an $ARMA(1,1)$, the limited and full information specifications in our paper would coincide. As we explain below, this case, can match the evidence for a downward sloping term structure only if fluctuations in expected consumption growth are a source of insurance rather than risk.

(and poorly identified) values of the off-diagonal elements of the **b** matrix, forecasts of consumption and dividend growth are actually improved by using the more parsimonious ARMA processes in place of an estimated VARMA specification. This is true in samples as large or even considerably larger than that currently available.[9] The reason is that estimation of the VARMA requires 11 unknown parameters to be identified, compared to just 4 unknown parameters in each individual ARMA estimation. The estimation of these additional parameters creates sufficient noise so as render the VARMA statistically inferior as a forecasting model for Δc_{t+1} and Δd_{t+1}, and therefore inferior as a model for estimating the latent conditional means $x_{c,t}$ and $x_{d,t}$.[10]

Even in infinite samples (where the VARMA model would provide superior forecasts), the welfare costs of using the more parsimonious ARMA processes in place of the full VARMA system are small. Consider two consumption sequences $\left\{C_i^A\right\}_{i=0}^{\infty}$ and $\left\{C_i^V\right\}_{i=0}^{\infty}$, where the former is the optimal sequence when the data are generated by (4)-(8) but the agent uses the two ARMA processes (15) and (16) as a cash-flow model, while the latter is the optimal sequence when the data are generated by (4)-(8) but the agent uses the VARMA system (14) as a cash-flow model. Then under our "refined calibration" (discussed below) with cash-flow statistics given in column (d) of Table 1, the welfare cost of receiving $\left\{C_i^A\right\}_{i=0}^{\infty}$ compared to $\left\{C_i^V\right\}_{i=0}^{\infty}$ amounts to 1.1% of time-0 monthly consumption.[11]

[9] The Web Appendix of this paper contains a detailed description of these statistical tests, conducted in both in the model using Monte Carlo experiments, and using historical data.

[10] The mean-square forecast error is increasing in both the bias and the variance of the forecast. Thus, investors face a tradeoff between the unbiased but noisy predictors $\Delta\widehat{c}_{t+1}^{VARMA}$ and $\Delta\widehat{d}_{t+1}^{VARMA}$, and the biased but less noisy predictors $\Delta\widehat{c}_{t+1}^{ARMA}$, and $\Delta\widehat{d}_{t+1}^{ARMA}$. If the improvement in forecast accuracy from eliminating bias is out-weighed by greater forecast noise, the $VARMA$ model will produce inferior forecasts.

[11] We define the welfare cost Λ of receiving $\left\{C_i^A\right\}_{i=0}^{\infty}$ rather than $\left\{C_i^V\right\}_{i=0}^{\infty}$ as the increment to lifetime utility (in consumption units) needed to make the investor indifferent between $\left\{C_i^A\right\}_{i=0}^{\infty}$ and $\left\{C_i^V\right\}_{i=0}^{\infty}$. This measure tells us what constant multiple $1+\Lambda$ of consumption in every period must be given to an investor with the stream $\left\{C_i^A\right\}_{i=0}^{\infty}$ to provide her with the same lifetime utility U as an investor with the stream

Notice that, in the limited information models considered here, there is no learning or Bayesian updating on the *parameters*, although there is learning about the latent conditional means of consumption and dividend growth based on the noisy information from the cash-flow model. The parameters in the general system (4)-(8) are unidentified, so even with an infinite amount of data the parameters of the true data generating process can never be learned. This contrasts with much of the learning literature, e.g., Veronesi (2000), Li (2005), and Ai (2010), where these papers entertain the possibility that agent's have some additional information–a signal–about dividends, consumption, or some component of output that can be used to learn about parameters.[12]

The cash flow model (15)-(16) serves as a signal extraction tool, allowing the investor to form an estimate of the long-run components in dividend and consumption growth, given historical data. This is immediately evident by noting that the pair of $ARMA(1,1)$ processes may be equivalently expressed in terms of the following pair of innovations representations

$\left\{C_i^V\right\}_{i=0}^{\infty}$. For the EZW preferences explored in this paper, this is given by

$$1+\Lambda=\frac{U_0\left(\left\{C_i^V\right\}_{i=0}^{\infty}\right)}{U_0\left(\left\{C_i^A\right\}_{i=0}^{\infty}\right)}.$$

Under the refined calibration, with $\Psi=1$, this ratio has an analytical solution. See Croce (2007).

[12] Johannes et al. (2013) show that learning about parameters can generate predictability of excess returns. The model considered here has no scope for generating time-varying risk premia. Adding time-varying consumption volatility to the model could in principle generate predictable variation in excess stock returns, e.g., (Bansal and Yaron (2004)). Even with a very highly persistent stochastic volatility process for consumption growth, however, long-run-risk models without parameter learning generate tiny magnitudes of forecastability in returns (see Bansal, Kiku and Yaron (2012), Ludvigson (2012)).

derived from the Kalman filter:

$$\Delta c_{t+1} = \mu_c + \widehat{x}^A_{c,t} + v^A_{c,t+1} \tag{17}$$

$$\widehat{x}^A_{c,t+1} = \rho \widehat{x}^A_{c,t} + K^A_c v^A_{c,t+1} \tag{18}$$

$$\Delta d_{t+1} = \mu_d + \widehat{x}^A_{d,t} + v^A_{d,t+1} \tag{19}$$

$$\widehat{x}^A_{d,t+1} = \rho \widehat{x}^A_{d,t} + K^A_d v^A_{d,t+1}, \tag{20}$$

where $K^A_c \equiv \rho - b_c$ and $K^A_d \equiv \rho - b_d$ and $\widehat{x}^A_{c,t}$ and $\widehat{x}^A_{d,t}$ denote optimal linear forecasts based on the history of consumption and dividend data separately, i.e., $\widehat{x}^A_{c,t} \equiv \widehat{E}\left(x_{c,t}|\mathbf{z}^t_c\right)$, and $\widehat{x}^A_{d,t} \equiv \widehat{E}\left(x_{d,t}|\mathbf{z}^t_d\right)$, where $\mathbf{z}^t_c \equiv (\Delta c_t, \Delta c_{t-1}, ..., \Delta c_1)'$ and $\mathbf{z}^t_d \equiv (\Delta d_t, \Delta d_{t-1}, ..., \Delta d_1)'$ .The optimal forecasts are functions of the observable ARMA parameters and innovations:

$$\widehat{x}^A_{c,t} = -\rho\mu_c + \rho\Delta c_t - b_c v^A_{c,t}$$

$$\widehat{x}^A_{d,t} = -\rho\mu_d + \rho\Delta d_t - b_d v^A_{d,t}.$$

Notice that the Kalman gain parameters K^A_c and K^A_d govern how much the estimated long-run components $\widehat{x}^A_{c,t}$ and $\widehat{x}^A_{d,t}$ respond to ARMA innovations $v^A_{c,t}$ and $v^A_{d,t}$, where the latter are non-linear functions of the primitive shocks in (8). The innovations representations above contain the same information about the latent variables $x_{c,t}$ and $x_{d,t}$ as do the ARMA processes (15) and (16).

3.3 Model Solution

Under full information, solutions to the model's equilibrium price-consumption and price-dividend ratios are found by iterating on the Euler equations (1) and (2), assuming that individuals observe the consumption and dividend processes. This means that under the special case (9)-(11), investors know that $x_{d,t} = \phi_x x_{c,t}$, thus the solution delivers a policy function for the price-consumption and price-dividend ratios as a function of a single state variable $x_{c,t}$.

Under the BRLI model, equilibrium price-consumption and price-dividend ratios are found by iterating on the Euler equations (1) and (2) assuming market participants observe (15)-(16), even though the data are actually generated by the dynamic system (4)-(8) with distinct short- and long-run components. In this case, the policy function for the price-consumption ratio is a function of the single state variable $\widehat{x}_{c,t}^{A}$, while that for the price-dividend ratio is a function of two state variables $\widehat{x}_{c,t}^{A}$ and $\widehat{x}_{d,t}^{A}$. For each specification, we simulate histories for consumption and dividend growth from the true data generating process and use solutions to the policy functions to generate equilibrium paths for asset prices.[13] The process is iterated forward to obtain simulated histories for asset returns. The Web Appendix explains how we solve for these functional equations numerically on a grid of values for the state variables.

4 Theoretical Results

This section presents theoretical results on the level and term structure of equity for both the full and BRLI models discussed above. We begin by illustrating the general nature of the challenge posed by evidence on the equity term structure for the full information paradigm, by considering the role played by key model parameter values. This can be accomplished by examining an approximate log-linear solution of the model, similar to Campbell (2003). We do this in Section 4.2, after introducing the concept of zero-coupon equity in Section 4.1. We then move on to show how assumptions about behavior and information structure matter, using numerical solutions.

[13] A minor complication is that the policy functions for the limited information specifications are a function of the current innovations in (15) and (16), whereas the actual innovations are generated from (9)-(11). However, the moving average representations are invertible, and their innovations can be recovered from $\sum_i b_c^i \left(\Delta c_{t-i} - \rho \Delta c_{t-i-1} - \mu_c\right)$ and $\sum_i b_d^i \left(\Delta d_{t-i} - \rho \Delta d_{t-i-1} - \mu_d\right)$, respectively.

4.1 Zero-Coupon Equity

An equity claim can be represented as a portfolio of zero-coupon dividend claims with different maturities (strips). Let $P_t^{(n)}$ denote the price of an asset at time t that pays the aggregate dividend n periods from now, and let $R_t^{(n)}$ be the one-period return on a zero-coupon equity strip that pays the aggregate dividend in n periods:

$$R_{t+1}^{(n)} = \frac{P_{t+1}^{(n-1)}}{P_t^{(n)}}.$$

Zero-coupon equity claims are priced under no-arbitrage according to the following Euler equation:

$$\begin{aligned} E_t\left[M_{t+1}R_{t+1}^{(n)}\right] &= 1 \Longrightarrow \\ P_t^{(n)} &= E_t\left[M_{t+1}P_{t+1}^{(n-1)}\right] \\ P_t^{(0)} &= D_t. \end{aligned}$$

Denote the log return on the n-period strip $r_{t+1}^{(n)} = \ln\left(R_{t+1}^{(n)}\right)$. Plotting $E\left(r_{t+1}^{(n)} - r_{f,t+1}\right)$ against n produces a yield curve, or term structure, of zero-coupon dividend strips. Since the aggregate market is the claim to all future dividends, the market price-dividend ratio $P_t^d/D_t = \sum_{n=1}^{\infty} P_t^{(n)}/D_t$ and the return on the market index, $R_{d,t+1}$ is the average return over all strips. We denote the log excess return on the market index as $r_{d,t+1}^{ex} \equiv r_{d,t+1} - r_{f,t+1}$.

4.2 Analytical Solutions

In this section we examine an approximate log-linear solution of the full and BRLI models to illustrate the role of key parameters in determining both the level and slope of the equity term structure.

4.2.1 Full Information

Two parameters in (4)-(8) play an important play role in determining whether risk premia are increasing or decreasing with maturity. These are the parameter $\rho_{xc,xd}$ that gives the correlation between the long-run innovations $\varepsilon_{xc,t}$ and $\varepsilon_{xd,t}$, and the parameter $\rho_{c,xc}$ that gives correlation between the short- and long-run consumption shocks $\varepsilon_{c,t}$ and $\varepsilon_{xc,t}$. For this reason, we free up restrictions embedded in (9)-(13) by allowing the parameter $\rho_{xc,xd} \neq 1$ and $\rho_{c,xc} \neq 0$. We also allow the short-run consumption shock $\varepsilon_{c,t+1}$ to be correlated with $\varepsilon_{xd,t+1}$, but only through its correlation with $\varepsilon_{xc,t+1}$, implying $\rho_{c,xd} = \rho_{c,xc} \cdot \rho_{xc,xd}$.[14] The parameter ϕ_x is defined $\phi_x \equiv \sigma_{xd}/\sigma_{xc}$, consistent with the definition of ϕ_x in (9)-(13).

We derive approximate log-linear solutions for the spread of the term structure and the equity market risk premium. Let $V_t(\cdot)$ denote the conditional variance of the generic argument "$\cdot$". Define the spread, S, of the log equity term structure (adjusted for a Jensen's inequality term) as

$$S \equiv \lim_{n\to\infty} E_t[r_{t+1}^{(n)ex} + .5V_t(r_{t+1}^{(n)ex}) - \left(r_{t+1}^{(1)ex} + .5V_t(r_{t+1}^{(1)ex})\right)], \tag{21}$$

where the superscript "ex" denotes the excess return over the log risk-free rate. S gives the difference in expected excess return between equity that pays a dividend in the infinite future and equity that pays a dividend in one period. From the loglinear approximation to the full information model, S can be shown to equal

$$S = \left(\phi_x \rho_{xc,xd} - \frac{1}{\Psi}\right) \underbrace{\left[\gamma \frac{\rho_{c,x_c}}{\sigma_{xc}} + \kappa_c \left(\frac{\gamma - 1/\Psi}{1 - \rho\kappa_c}\right)\right] \frac{\sigma_{xc}^2 \sigma^2}{1-\rho}}_{>0} \tag{22}$$

where $\kappa_c \equiv \frac{\overline{P^C/C}}{1+\overline{P^C/C}}$.

[14] To avoid clutter in the formulas, we maintain, as a benchmark, other restrictions imposed in (9)-(13) on the remaining correlations. Specifically we set $\rho_{d,xd} = \rho_{d,xc} = \rho_{c,d} = 0$. Freeing up these correlations does not change the conclusions of this section.

The parameter ϕ_x controls the exposure of dividend growth to the persistent component in consumption growth. The exposure of dividend growth to long-run consumption risk affects the slope of the term structure because it affects expected *future* dividend growth. Equation (22) shows that, the lower is ϕ_x, the less persistent is dividend growth and the less upward sloping is the term structure. This happens because, when dividend growth has little persistence, only the expected growth rates of dividends paid in the near future are revised significantly in response to an innovation; those paid in the far future are relatively unaffected. Thus a lower exposure of dividend growth to long-run consumption risk raises risk premia on short maturity strips relative to those on long maturity strips, driving down S. This effect must be multiplied by the correlation $\rho_{xc,xd}$, since that controls the extent to which a movement in expected dividend growth is correlated with the pricing kernel, and therefore priced.

The EIS Ψ affects the slope of the term structure by affecting expected future returns, rather than expected future dividend growth. The lower is Ψ, the more the expected risk-free rate increases in response to an increase in expected consumption growth. A positive innovation in expected consumption growth does two things. First, it leads to an increase in the expected future risk-free rate (increasingly so with smaller values of Ψ), which is associated with a capital loss for the asset today. Second, it leads to a decline in the stochastic discount factor. The two combined produce a positive contemporaneous correlation between the pricing kernel and returns, reducing the overall risk premium on the asset. This effect is stronger for assets that pay a dividend in the far future because shocks to expected consumption growth are persistent and cumulate over time. Consequently, the lower is the EIS Ψ, the lower are risk premia on long-duration assets relative to short-duration assets, and the less upward sloping is the zero-coupon-equity curve.

For the rest of this discussion, we maintain the assumption that $\gamma > 1/\Psi$. If we also assume for the moment that $\rho_{c,x_c} \geq 0$, then the term in the square brackets of (22) is positive, and it is possible to generate a downward sloping term structure of equity (a negative spread, $S < 0$) by setting $\phi_x \rho_{xc,xd} < 1/\Psi$. In the Bansal and Yaron (2004) parametrization, $\rho_{xc,xd} = 1$, so a downward slope could be generated by simply changing parameter values so that $\phi_x < 1/\Psi$.[15]

Under full information, however, this strategy for obtaining a downward sloping term structure presents a different problem. When $\phi_x \rho_{xc,xd} < 1/\Psi$, equivalently with $S < 0$, the model becomes one of long-run *insurance* rather than long-run *risk*. That is, innovations in $x_{d,t}$ (holding other shocks fixed) generate a *positive* correlation between the pricing kernel and returns, so that the marginal contribution of the long-run component to the market risk premium is negative. In a long-run "insurance" model, the full information term structure slopes down, but the overall equity premium for the market will be low or negative. This can be confirmed by examining the loglinear approximate solution for the log equity premium (adjusted for Jensen's inequality terms), given by

$$E_t(r^{ex}_{d,t+1}) + .5V_t(r^{ex}_{d,t+1}) = E_t\left[r^{(1)ex}_{t+1} + .5V_t\left(r^{(1)ex}_{t+1}\right)\right] + \kappa_d \frac{(1-\rho)}{1-\rho\kappa_d} S, \tag{23}$$

where $\kappa_d \equiv \frac{\overline{P^D/D}}{1+\overline{P^D/D}} > 0$ and the first term on the right-hand-side is excess return on the one-period zero coupon equity strip. It is evident that if the spread S of the equity term structure is negative, the equity premium on the left-hand-side will be small or negative, depending on the size of the risk premium on the one-period strip. But the size of the premium on the one-period strip–and therefore the level of the term structure–depends only on the dividend claim's exposure to short-run consumption risk:

$$E_t\left[r^{(1)ex}_{t+1} + .5V_t\left(r^{(1)ex}_{t+1}\right)\right] = \left[k_c \frac{\gamma - 1/\Psi}{1-\rho\kappa_d} \sigma_{xc}\rho_{c,xc} + \gamma\right] \phi_c \sigma^2. \tag{24}$$

[15] When $\Psi = 1$, the valuation calculations in Hansen et al. (2008) can be used to obtain an exact solution for S. Under the assumptions just made, such calculations show analogously that $\phi_x < 1$ is required to generate a downward sloping equity term structure.

Long-run consumption risk $x_{c,t}$ contributes to the premium on the one-period strip, but only in so far as it has a non-zero correlation with short-run risk $\rho_{c,xc} \neq 0$. The one-period strip will have a zero risk premium in any model where exposure ϕ_c to short-run consumption risk is zero.

These facts imply that to generate $S < 0$ under full information, shocks to the long-run component of dividend growth must be a source of insurance so that long-run consumption shocks generate a negative risk premium. This can be seen by re-writing the market risk premium as a function of the covariance between the log stochastic discount factor, m_{t+1}, and the log excess return on the dividend claim:

$$\begin{aligned} E_t\left[r^{ex}_{d,t+1}\right] + .5V_t(r^{ex}_{d,t+1}) &\approx -\mathrm{Cov}_t\left(r^{ex}_{d,t+1}, m_{t+1}\right) & (25)\\ &= -\mathrm{Cov}_t\left(\phi_c \sigma \varepsilon_{c,t+1}, m_{t+1}\right) & (26)\\ &\quad -\frac{\kappa_d}{1-\rho\kappa_d}\sigma_{xc}\sigma \mathrm{Cov}_t\left(\phi_x \varepsilon_{xd,t+1} - \frac{1}{\Psi}\varepsilon_{xc,t+1}, m_{t+1}\right). & (27) \end{aligned}$$

The term in (26) is the component of the total risk premium attributable to short-run consumption risk. This term always generates a positive risk premium as long as $\phi_c > 0$ because the short-run shock $\varepsilon_{c,t+1}$ is negatively correlated with m_{t+1}. The term in (27) is the component of the total risk premium attributable to long-run consumption risk. This term can, under some plausible parameters, contribute a negative risk premium, implying that the long-run shock is a source of insurance. This term will generate a negative risk premium if ϕ_x and Ψ are sufficiently small because, with $\varepsilon_{xd,t+1}$ and $\varepsilon_{xc,t+1}$ both negatively correlated with m_{t+1}, sufficiently low values for ϕ_x and Ψ imply $\mathrm{Cov}_t\left(\phi_x \varepsilon_{xd,t+1} - \frac{1}{\Psi}\varepsilon_{xc,t+1}, m_{t+1}\right) > 0$ so the second term in (26) is negative. (Note that the outer multiplicative term $\frac{\kappa_d}{1-\rho\kappa_d}\sigma_{xc}\sigma \geq 0$.) Under these parameter values, the impact of a long-run risk shock on expected dividend growth is more than offset by the countervailing effect on the expected risk-free rate, and long-run shocks to consumption growth are a source of insurance, rather than risk.

But the question of whether shocks to consumption growth are a source of insurance rather than risk boils down to the question of whether S is positive or negative:

$$\frac{\kappa_d (1-\rho)}{1-\rho\kappa_d} S = -\frac{\kappa_d}{1-\rho\kappa_d}\sigma_{xc}\sigma \mathrm{Cov}_t\left(\phi_x \varepsilon_{xd,t+1} - \frac{1}{\Psi}\varepsilon_{xc,t+1}, m_{t+1}\right).$$

This shows that $S < 0$ only when $\mathrm{Cov}_t\left(\phi_x \varepsilon_{xd,t+1} - \frac{1}{\Psi}\varepsilon_{xc,t+1}, m_{t+1}\right) > 0$, where long-run consumption risk generates positive covariance with the log pricing kernel and a negative risk premium.

What about $\rho_{xc,xd}$? Equation (22) shows that we can make the spread S negative by lowering $\rho_{xc,xd}$. But, (23) makes clear that $S < 0$ in this model comes at the expense of a lower equity premium. Moreover this problem cannot be remedied by freely setting the correlation ρ_{cx_c} between short-run consumption shocks $\varepsilon_{c,t}$ and long-run shocks $\varepsilon_{xc,t}$. For example, if we restrict $\phi_x \rho_{xc,xd} > 1/\Psi$ to avoid the implications just discussed, then (22) shows that we can obtain a downward sloping term structure by setting $\rho_{cx_c} < 0$. But (23) and (24) show that this will again make the overall equity premium low or negative, since it makes both S and the first term of (24) negative.

In summary, the full information paradigm can generate a negative equity term structure spread S, but only if the model is one of long-run insurance rather than long-run risk. Since in this case the persistent component of consumption growth generates a negative risk premium, the overall market risk premium will typically be small or negative. We explore the magnitudes of these effects for specific parameter values below.

4.2.2 Limited Information

In the BRLI model, there is one source of risk, captured by the ARMA innovation $v_{c,t}^A$. The approximate log linear solution under the BRLI model implies that the slope of the term structure, (21), is

$$S = (\phi_x^A - 1/\Psi)\frac{1}{1-\rho}\left[\kappa_c^A \frac{\gamma - 1/\Psi}{1-\rho\kappa_c} + \frac{\gamma}{K_c^A}\right](K_c^A \sigma_{v_c^A})^2,$$

$$\phi_x^A = \rho_{v_c^A, v_d^A} \frac{K_d^A \sigma_{v_d^A}}{K_c^A \sigma_{v_c^A}}, \tag{28}$$

where $\kappa_c^A \equiv \frac{\overline{P^C/C}}{1+\overline{P^C/C}}$ and $\overline{P^C/C}$ is the mean price-consumption ratio under the BRLI equilibrium, $\rho_{v_c^A, v_d^A}$ is the correlation between the two ARMA innovations v_c^A and v_d^A, and $\sigma_{v_c^A}$ and $\sigma_{v_d^A}$ are their respective standard deviations. Thus ϕ_x^A is the coefficient from a projection of $\widehat{x}_{d,t}^A$ on $\widehat{x}_{c,t}^A$ and is therefore a measure of the exposure of the estimated long-run component in dividend growth to the estimated long-run consumption component (See the Appendix for a derivation). The ϕ_x^A coefficient plays the analogous role to $\phi_x \rho_{xc,xd}$ in (22), which measures the exposure of dividend growth to the persistent component in consumption growth under full information.

In the BRLI model, this exposure is not observed and what is observed–the history of dividend growth Δd_{t+1}–is but a noisy signal of the long-run component, $x_{d,t}$. The estimate of the long-run component $\widehat{x}_{d,t}^A$ is contaminated by the two i.i.d. shocks that affect Δd_{t+1} through the terms $\phi_c \sigma \varepsilon_{c,t+1} + \sigma_d \sigma \varepsilon_{d,t+1}$ in (5). The more volatile are these shocks relative to $x_{d,t}$, the lower is the information content of observable dividend growth for movements in the long-run dividend component $x_{d,t}$. The idiosyncratic shock $\varepsilon_{d,t+1}$ is especially important here, since it is the most volatile. The higher is σ_d, the less correlated dividend growth appears to be with consumption growth and the lower is $\rho_{v_c^A, v_d^A}$ in (28). A low value for $\rho_{v_c^A, v_d^A}$ lowers ϕ_x^A, making dividends appear to have little exposure to long-run consumption risk, and lowering S. Thus shocks to the long-run risk component of dividends are close to being unpriced in the BRLI model, even though they command a substantial risk premium under full information. For this reason, long maturity strips have low risk premia in the BRLI model, generating a low or negative value for S.

Turning our attention to the level of the term structure, we prove in the Appendix that the following relation follows from a log-linear approximation:

$$E_t[r_t^{(1)ex}] + .5V_t(r_t^{(1)ex}) = \left[\kappa_c^A \frac{\gamma - 1/\Psi}{1-\rho\kappa_c} K_c^A + \gamma\right] \underbrace{\rho_{v_c^A, v_d^A} \sigma_{v_c^A} \sigma_{v_d^A}}_{Cov_t^A}. \tag{29}$$

The term in square brackets, which is positive under common parameter configurations, is the price of consumption risk in the BRLI model, and Cov_t^A is the quantity of that risk.

It is instructive to compare (29) with its counterpart under full information (24). There are two ways in which these equations substantively differ. First, the quantity of risk Cov_t^A in (29) differs conceptually from the true covariance between the short-run components $Cov\left(\sigma\varepsilon_{c,t}, \phi_c\sigma\varepsilon_{c,t}\right) = \phi_c\sigma^2$ that appears in (24). But it is straightforward to show that, under a range of parameter values around the benchmark, Cov_t^A is virtually indistinguishable from $\phi_c\sigma^2$. To see why, recall that Cov_t^A is not appreciably affected by shocks to the long-run component of dividend growth $\varepsilon_{xd,t+1}$ in (7) because those shocks are too small to be detected given the larger short-run shocks $\sigma\varepsilon_{c,t+1}$ in (5). And cov_t^A is not greatly affected by the idiosyncratic dividend shocks because those shocks are uncorrelated with consumption. Thus, $Cov_t^A \approx Cov\left(\sigma\varepsilon_{c,t}, \phi_c\sigma\varepsilon_{c,t}\right) = \phi_c\sigma^2$ implying that the *quantity* of short-run consumption risk is virtually identical in the BRLI and LI models. In both cases, it increases monotonically in ϕ_c.

By contrast, the *price* of short-run consumption risk *is* affected by the information structure. The price of risk (in square brackets) differs across these models in the terms that govern the perceived standard deviation of the long-run consumption component (relative to the standard deviation of the short-run component), scaled by the perceived correlation between the short- and long-run consumption shocks. In the full information formula, (24), this term is $\sigma_{xc}\rho_{c,xc}$, where recall that $\sigma_{xc}^2 = Var_t\left(x_{c,t+1}\right)/\sigma^2$. In (29), the term $\sigma_{xc}\rho_{c,xc}$ is replaced by the Kalman gain parameter K_c^A. Note that this parameter, like σ_{xc}, controls the relative volatility of the estimated long-run consumption risk component: $\left(K_c^A\right)^2 = Var_t\left(\widehat{x}_{c,t}^A\right)/\sigma_{v_c^A}^2$. Thus, K_c^A in the BRLI model plays the role of σ_{xc} in full information. While σ_{xc} is multiplied by the correlation $\rho_{c,xc}$ in (24), K_c^A is multiplied by unity in (29), because $Corr_t\left(\widehat{x}_{c,t+1}^A, v_{c,t+1}^A\right) = 1$.

Importantly, under plausible parameter configurations, the parameter K_c^A will be greater than $\sigma_{xc}\rho_{c,xc} > 0$, implying that $E_t\left[r_{t+1}^{(1)ex} + .5V_t\left(r_{t+1}^{(1)ex}\right)\right]$ will be greater under the BRLI model than under full information. This happens because, as we show in the next section, the BRLI decision maker cannot distinguish the short-and long-run components. As a consequence, she erroneously revises upward her estimate of the long-run component $\widehat{x}_{c,t}^A$ in response to an observed increase in consumption driven by a purely short-run (i.i.d.) consumption innovation, thereby making $\widehat{x}_{c,t}^A$ look more volatile (relative to short-run risk) than is the true $x_{c,t+1}$ observable under full information. For this reason, short-run consumption shocks generate a larger negative covariance between the pricing kernel and returns under the BRLI model than under full information, for any given quantity of risk. This feature of the model generates higher risk premia on short maturity strips under the BRLI model than under full information. In the web Appendix we prove that, up to the parameter $\kappa_d^A \equiv \frac{\overline{P^D/D}}{1+\overline{P^D/D}}$ (now a function of the equilibrium $\overline{P^D/D}$ ratio under the BRLI model), the approximate solution (23) for the market risk premium in full information also holds under the BRLI model:

$$E_t(r_{d,t+1}^{ex}) + .5V_t(r_{d,t+1}^{ex}) = E_t\left[r_{t+1}^{(1)ex} + .5V_t\left(r_{t+1}^{(1)ex}\right)\right] + \kappa_d^A\frac{(1-\rho)}{1-\rho\kappa_d^A}S \tag{30}$$

But since $E_t\left[r_{t+1}^{(1)ex} + .5V_t\left(r_{t+1}^{(1)ex}\right)\right]$ will, under plausible parameter configurations, be greater under the BRLI model than under full information, (30) shows that, all else equal, the market risk premium will also be higher under the BRLI model than under full information.

We close this section by noting that the full information long-run risk paradigm cannot generate both a non-trivial downward sloping term structure and sizable equity premium even if the true cash flow process is a pair of $ARMA\,(1,1)$ processes. As we show in the appendix, in this case the only parameterizations of the cash flow process for which this is possible are those in which the model is one of long-run *insurance*, where the persistent component of consumption growth reduces the equity premium, rather than increases it. This is not true of the BRLI model, in which the true cash flow process still implies that the persistent component of consumption growth would contribute positively to the equity premium if the agent could observe it. It is only because the agent conflates the sources of cash-flow variation, coupled with the bounded rationality behavioral assumption, that the term structure can be downward sloping along with a sizable equity premium. In the next section we employ numerical solutions to explore the magnitudes of these effects under plausible parameter values.

4.3 Numerical Solutions

4.3.1 Benchmark Parameter Values

We begin by presenting theoretical results based on a benchmark calibration, and then ask how those results change as we change key parameters. To form a benchmark, we take a special case of the general cash flow model (4)-(8) that is close to what has been used in the long-run risk literature. In this special case, as in (9)-(13), consumption and dividend growth contain a single, common predictable component $x_{d,t} = \phi_x x_{c,t}$ implying $\rho_{xc,xd} = 1$ and $\sigma_{xd} = \phi_x \sigma_{xc}$. Also, $\rho_d = \rho$, $x_{c,0} = x_{d,0} = 0$, and except for $\rho_{xc,xd}$, the off-diagonal elements in (8) are all zero. The benchmark calibration of (4)-(8) with parameters set at monthly frequency is as follows: $\gamma = 10$, $\Psi = 1.5$, $\delta = 0.998985, \mu_c = \mu_d = 0.0015$, $\rho = 0.979$, $\sigma = 0.0078$, $\phi_x = 1$, $\phi_c = 4$, $\sigma_{xc} = 0.044$, $\sigma_{xd} = \phi_x \sigma_{xc}$, $\sigma_d = 4.5$, $\rho_{xc,xd} = 1$ and $\rho_{c,d} = \rho_{c,xc} = \rho_{c,xd} = \rho_{d,xc} = \rho_{d,xd} = 0$. With the exception of ϕ_x and ϕ_c, these values are the same as those in the benchmark specification of Bansal and Yaron (2004). Notice that the innovation variance in $x_{c,t}$ is small relative to the overall volatility of consumption (the standard deviation of ε_{xc} is 0.044 times the standard deviation of ε_c), but the persistence of

$x_{c,t}$ is high, a hallmark of long-run risk models. In Bansal and Yaron (2004), the loadings ϕ_x and ϕ_c are set to 3 and 0, respectively; we instead set the benchmark loadings ϕ_x and ϕ_c to 1 and 4, respectively. These loadings require more discussion.

The parameters ϕ_x and ϕ_c are crucial determinants of the slope and level of the equity term structure, since they govern the exposure of dividends long-run versus short-run consumption risk. Although Bansal and Yaron (2004) set the exposure to short-run risk to zero (implying $\phi_c = 0$), we have seen in (24) that this parametrization implies, counterfactually, that the one-period zero coupon equity strip has no risk premium. Exposure to short-run consumption risk is necessary for the model to match evidence for a non-zero (and sizable) risk premium on the one-period strip. Since the one-period strip is estimated to have a risk premium as much as 5 to 7 percent per annum higher than the market index itself, as a benchmark we set $\phi_c = 4$. Table 1 compares the statistical properties of the implied cash flow processes under the original Bansal and Yaron (2004) calibration with $\phi_c = 0$ (column 1) with this alternative where $\phi_c = 4$ and all other parameters are unchanged (column 2). Although $\phi_c = 4$ is needed to help match the evidence for a sizable premium on the one-period strip, if no other parameter values are changed, column (2) shows that dividend growth is then too autocorrelated and too highly correlated with consumption growth. For this reason, in our benchmark, we lower exposure of dividend growth to long-run risk by setting $\phi_x = 1$ rather than $\phi_x = 3$, largely restoring the model's favorable cash flow properties for the joint behavior of dividend and consumption growth (column (3)). Since $\sigma_{xd} = \phi_x \sigma_{xc}$, under this benchmark calibration $\sigma_{xd} = \sigma_{xc}$. Column 4 of Table 1 shows the cash flow properties of a "refined" calibration that will be discussed below.

It is instructive to begin by comparing the full information and BRLI models with regard to how the cash flow state variables react to primitive shocks under the benchmark calibration. Figure 1 compares the full and BRLI cases by plotting the impulse responses to primitive shocks (in percent deviations from steady state). The responses plotted are for

$x_{c,t}$, as compared to $\widehat{x}^A_{c,t}$ (column 1 of panels), for $x_{d,t}$, as compared to $\widehat{x}^A_{d,t}$ (column 2) and finally for the dividend surprise $(\Delta d_t - E_{t-1}\Delta d_t)$ in the full information and BRLI models (column 3). The first row of each column displays the responses to a one-standard deviation increase in the i.i.d. consumption shock, $\varepsilon_{c,t}$; the second row displays the responses to a one-standard deviation increase in the idiosyncratic dividend shock, $\varepsilon_{d,t}$; and the third row displays responses to a one-standard deviation increase in the innovation to the persistent component of consumption growth, $\varepsilon_{xc,t}$. In the figures, we denote all variables under full information without hats, while variables in the BRLI model are denoted with hats.

Figure 1 shows that a one-standard deviation increase in the i.i.d. consumption shock $\varepsilon_{c,t}$ leads to a sharp, unexpected increase in dividend growth in both the FI and BRLI models (row 1, column 3). With full information, the agent observes the source of the shock and understands that it has no persistence. Accordingly, expectations of future consumption growth and future dividend growth are unchanged, so the impulse responses of $x_{c,t}$ and $x_{d,t}$ are flat at zero. By contrast, in the BRLI model agents cannot directly observe the source of the shock and do not know if it is persistent or transitory. The solution to their filtering problem implies that agents erroneously revise upward their expectation of future consumption growth and, to a lesser extent, future dividend growth, even though in reality the shock has no persistence. Thus, both $\widehat{x}^A_{c,t}$ and $\widehat{x}^A_{d,t}$ rise. In response to a transitory shock, agents in the BRLI model revise their expectations of future consumption and dividend growth *more* than they would under full information.

Now consider the responses to an innovation in the persistent component of consumption, $\varepsilon_{cx,t}$, in the third row of Figure 1. Under full information, investors recognize that this is a shock to the persistent component of consumption and dividend growth and they accordingly revise upward their expectations of future consumption and dividend growth immediately upon observing the shock. Row 3 of Figure 1 shows that a one-standard deviation increase in $\varepsilon_{cx,t}$ leads to a jump upward in $x_{c,t}$ and $x_{d,t}$. By contrast, investors in the BRLI model revise upward their expectation of future consumption and dividend growth only gradually and by much less than they do under full information. The state variable, $\widehat{x}_{c,t}^{A}$ responds sluggishly to the shock and $\widehat{x}_{d,t}^{A}$ barely responds at all. This happens because the persistent component in consumption growth (which drives the persistent component in dividend growth) is too small to be distinguished from the large idiosyncratic dividend shocks. In response to a persistent shock, agents in the BRLI model revise their expectations of future consumption and dividend growth *less* than they would under full information.

Finally, the middle row of Figure 1 shows that a purely idiosyncratic shock to dividend growth, $\varepsilon_{d,t}$, has no affect on expected consumption or dividend growth in full information, and has only a tiny affect on expected dividend growth in the BRLI model. This i.i.d. shock is the most volatile component of dividend growth and in the BRLI model it swamps everything, making dividend growth appear close to i.i.d. We return to these responses to interpret why risk premia differ depending on the information structure.

To illustrate how risk premia depend on the information structure, it is useful to compare economies with different aggregate dividend processes. Specifically, we study how risk premia differ across information specifications when the relative exposure of dividend growth to long-run versus short-run consumption risk differs. This is accomplished by comparing how the loadings governing long-run and short-run risk exposure effect equilibrium outcomes.

Table 2 presents the model's implications for summary statistics on the price-dividend ratio, excess returns, and risk-free rate under limited and full information. The model output is generated by simulating 1000 samples of size 840 months, computing annual returns from monthly data, reporting the average statistics across the 1000 simulations.[16] With the

[16] The average levels of the price-dividend ratios reported below are not directly comparable to their empirical counterparts for actual firms, since unlike real firms, the firms in the model have no debt and do not retain earnings. Dividends in the model are more analogous to free cash flow than to actual dividends, implying that model price-dividend ratios should be lower than measured price-dividend ratios in historical

exception of the parameters ϕ_c and ϕ_x (where results for a range of values are presented), the results in Table 2 are based on the benchmark parameter configuration discussed above.[17] The log return on the dividend claim is denoted $r_{d,t+1} = \ln(R_{d,t+1})$ and the log return on the risk-free rate $r_{f,t+1} \equiv \ln\left(R_{f,t+1}\right)$.

The results in Table 2 show that, with full information, high exposure to long-run consumption risk is required to generate high risk premia. Economies comprised of assets with relatively low exposure ϕ_x to long-run consumption risk and high exposure ϕ_c to short-run consumption risk (e.g., row 2 of Table 2), have lower risk premia and higher price-dividend ratios than do economies comprised of assets with higher ϕ_x and lower ϕ_c (e.g., row 5 of Table 2). In addition, substantial variation in risk premia can only be generated by heterogeneity in the exposure to long-run consumption risk; heterogeneity in short-run risk is inadequate. For example, when $\phi_x = 3$ and ϕ_c is increased from 2.5 to 6, the log risk premium $E(r_d - r_f)$ increases by just one and a third percent, from 7.33% to 9.66% per annum. By contrast if we keep $\phi_c = 2.5$ and increase ϕ_x from 1 to 3, the risk premium rises almost three-fold, from 2.73% to 7.33%.

The results under the proposed bounded rationality limited information model of behavior are much the opposite. Economies comprised of assets with relatively low exposure to long-run consumption risk and high exposure to short-run consumption risk, (e.g., row 2 of Table 2), now have high risk premia and low price-dividend ratios. Under this parametrization, the log risk premium $E(r_d - r_f)$ is over 8 percent per annum in the BRLI model, while it is a smaller 5.24 percent per annum under full information. At the same time, assets with high ϕ_x and low ϕ_c (e.g., row 5 of Table 2), have a log risk premium that is a much smaller 4.9%

data.

[17] The table reports values for ϕ_c ranging from a low of 2.5 to a high of 6. Smaller values for ϕ_c are ruled out in the BRLI model by the requirement that the price-dividend ratio be finite. This is analogous to the requirement in the Gordon growth model that the expected stock return be greater than the expected dividend growth rate to keep the price-dividend ratio finite.

per annum whereas the premium under full information is 7.3% per annum. And, unlike full information, in the BRLI model substantial variation in risk premia can be generated by heterogeneity in the exposure to *short-run* consumption risk. For example, when $\phi_x = 3$ and ϕ_c is increased from 2.5 to 6, the log risk premium increases by over 5 percentage points from 4.91% to 10.73% per annum. On the other hand, fixing ϕ_c and varying ϕ_x generates much less variation in risk premia in the BRLI model.

These findings can be illuminated graphically as in Figure 2, which plots annualized price-dividend ratio policy functions as a function of the ratio of long-run to short-run consumption risk exposure, ϕ_x/ϕ_c. For this figure, the ratio ϕ_x/ϕ_c is varied along the horizontal axis in such as way as to hold the 15-month variance of dividend growth that is attributable to the consumption innovations fixed. The left-most panel plots this ratio in the BRLI model at the steady state value of $\widehat{x}_{c,t}^A$, along with plus and minus two standard deviations around steady state in $\widehat{x}_{c,t}^A$ (holding fixed $\widehat{x}_{d,t}^A$ at its steady-state level). The middle panel plots the price-dividend ratio at the steady state value of $\widehat{x}_{d,t}^A$, along with plus and minus two standard deviations around steady state in $\widehat{x}_{d,t}^A$ (holding fixed $\widehat{x}_{c,t}^A$ at its steady-state level). The right-most panel plots the price-dividend ratio under full information as a function of ϕ_x/ϕ_c, plus and minus two standard deviations around steady state in the single state variable $x_{c,t}$. Note that the plots in Figure 2 are upward sloping in the BRLI model but downward sloping under full information. Since price-dividend ratios are high when risk premia are low, and vice versa, this shows that assets with cash flows that load heavily on the long-run component, $x_{c,t}$, are *more* risky under full information but *less* risky in the BRLI model.

These results can be understood intuitively by noting that the risk premium on these assets is determined primarily by the covariance between the pricing kernel M_t and revisions in expectations (news) about future cash flow growth.[18] As such, cash flow shocks have two

[18] Revisions in expected future returns are relatively unimportant because we have not introduced mechanisms such as changing consumption and dividend volatility for generating time-varying risk premia on the

offsetting effects on the equity premium in the BRLI model as compared to full information. First, when a positive innovation $\varepsilon_{xc,t+1}$ to the persistent component of consumption growth occurs, investors using the BRLI cash flow model assign some weight to the possibility that the change in observed cash flow growth rates is attributable to one of the i.i.d shocks ($\varepsilon_{c,t+1}$ or $\varepsilon_{d,t+1}$). As a consequence, these investors revise upward their expectation of *future* consumption and dividend growth in response to a persistent shock by less than they would if they had full information. Persistent shocks therefore generate a larger (in absolute value) negative correlation between M_t and cash flow news under full information than in the BRLI model. Second, when a positive innovation $\varepsilon_{c,t}$ to the short-run risk component occurs, investors using the BRLI cash flow model assign some weight to the possibility that the shock is persistent (coming from the long-run risk component). As a consequence, these investors revise upward their expectation of future consumption and dividend growth in response to an i.i.d. consumption growth shock more than they would if they had full information. The i.i.d. shocks therefore generate a larger (in absolute value) negative correlation between M_t and cash flow news in the BRLI model than in the FI model.

When ϕ_x is large and ϕ_c relatively small, the first effect dominates the second. In this case, the risk premium in the full information case can be substantial while the premium in the bounded rationality limited information case is small. On the other hand, when ϕ_x is small and ϕ_c relatively large, the second effect dominates. In this case, the risk premium bounded rationality limited information case can be substantial while that under full information is small. Note that when ϕ_c is small, volatility in the ARMA dividend shock $v^A_{d,t+1}$ is dominated by the volatile idiosyncratic cash flow shocks $\varepsilon_{d,t+1}$ that carry no risk premium. This explains why sufficiently high exposure to short-run risk ϕ_c is required to generate a large risk premium in the BRLI model. These results help explain the differing implications for the slope and level of the equity term structure, discussed next.

asset, and because the expected risk-free rate has small volatility.

4.3.2 The Term Structure of Equity

We now turn to the equity term structure. Figure 3 plots summary statistics for expected log excess returns $E\left(r_{t+1}^{(n)} - r_{f,t+1}\right)$ as a function of maturity, n, in months, directly comparing the FI and BRLI models using the benchmark calibration. It is immediately evident that, under full information, the annualized log risk premium increases with maturity (top panel). The log risk premium is 1.3% per annum for equity that pays a dividend one month from now and 2.3% per annum for equity that pays a dividend 15 years from now.

Keeping the same parameter values but imposing the bounded-rationality limited information assumptions, we see that the slope of the term structure is reversed. The annualized log risk premium now declines with maturity. At the benchmark parameter values the spread is modest (a point we return to below), with the log risk premium for equity that pays a dividend one month from now equal to 5.5% per annum compared to 4.5% for equity that pays a dividend 15 years from now (top panel). Nevertheless, short-horizon strips have higher expected returns and lower price-dividend ratios than do long-horizon strips and the term structure slopes down.

The key to the downward sloping term structure in the BRLI model is that the process for dividend growth given by the estimated ARMA cash flow model appears close to i.i.d. (the estimated moving average and autoregressive roots in (16) are close to canceling). Thus, surprises to dividend growth are perceived only to affect dividend growth and returns in the near term, so only assets that pay a dividend in the near future command high risk premia. By contrast, when agents can directly observe $x_{c,t}$, it is understood that innovations to this component can have large, long-term effects on consumption and dividend growth. Accordingly, the long-run appears risky, and assets that pay a dividend in the far future command higher risk premia than those that pay a dividend in the near future. The endogenous relation between equity maturity and risk premia goes the wrong way.

The middle panel of Figure 3 shows that in both models, volatility increases with the horizon. But the bottom panels show that the Sharpe ratios decrease with the horizon under the BRLI model whereas they rise with the horizon under full information. This suggests that the BRLI specification is better able to explain the empirically higher Sharpe ratios of short-duration value stocks as compared to long-horizon growth stocks (Cornell (1999, 2000); Dechow et al. (2004); Da (2005)).

Figure 4 shows that, in the BRLI model, short maturity strips have high CAPM alphas, whereas the long maturity strips have smaller (in absolute value) negative alphas. This feature of the model is consistent with the data and with the prior findings of Fama and French (1992). The bottom panel also shows that, in the BRLI model, long maturity strips—despite their lower expected excess returns—have slightly higher CAPM betas. This also is consistent with the evidence in Cornell (1999, 2000); Dechow et al. (2004); Da (2005); van Binsbergen et al. (2010). By contrast, under full information there is much less variation in the alphas with maturity and the alphas of short maturity strips are lower than those of long maturity strips.[19]

Alternative parameterizations: Full Information We have seen that the full information model generates an upward sloping term structure under the benchmark parametrization. Accordingly, assets that pay a dividend in the far future counterfactually command higher risk premia than those that pay a dividend in the near future. Here we explore the sensitivity of this result to alternative parameter values.

Figure 5 illustrates graphically the role different parameters play in determining both the equity term structure spread S and the total equity premium under full information. Panel

[19] The plots are reminiscent of Hansen et al. (2008). Hansen et. al. report price-dividend zero-coupon equity structures for value and growth firms separately, whereas we plot the zero-coupon-equity curve for aggregate dividends. In this sense, the results in this section are not directly comparable to those in Hansen et al. (2008).

(a) illustrates the role of ϕ_x, by showing the results when this parameter is altered and all others are kept at their benchmark values. When $\phi_x = 1$, as in the benchmark, the term structure slopes up and the log equity premium is 3.82% per annum. When we shut down the exposure of dividend growth to long-run risk entirely, setting $\phi_x = 0$, the term structure now slopes down, $S < 0$, but the log risk premium is only 0.9% per annum. We can make S more negative by making exposure to long-run risk negative, e.g., $\phi_x = -0.5$, but now the equity premium is negative, equal to -0.96%.

Panel (b) of Figure 5 illustrates the role of the correlation between the long-run components in consumption and dividend growth $\varepsilon_{xd,t+1}$ and $\varepsilon_{xc,t+1}$ by altering only $\rho_{xc,xd}$ and keeping all other parameters fixed at their benchmark values. Under the benchmark $\rho_{xc,xd} = 1$, the term structure slopes up and the log equity premium is 3.82% per annum. Panel b shows that we can make the spread S negative by lowering $\rho_{xc,xd}$. When $\rho_{xc,xd} = 0.5$, the spread is slightly negative but almost flat. When $\rho_{xc,xd} = 0$, we get a more sizable negative spread, but now the log risk-premium is low, equal to 0.88% per annum. An empirically plausible value for S would require a steeply negative risk-premium.

The full information specification *can* produce a sizable equity premium with $S < 0$ if exposure of dividend growth to short-run risk is sufficiently high, while exposure to long-run shocks sufficiently low. This is displayed in Panel (c) of Figure 5, where we fix $\phi_x = 0$, so as to completely eliminate the exposure of dividend growth to long-run shocks, and vary exposure to short-run shocks ϕ_c (all other parameters remain at their benchmark levels.) If $\phi_c = 4$, we get a modest downward spread $S = -2\%$ with a small log equity premium of 0.9% per annum. We can increase the log of the market risk premium to 2.42% by increasing ϕ_c to 6. Notice however that, whenever $S < 0$, there is no more long-run risk, only long-run insurance, so more than 100% of the equity premium must be accounted for by short-run risk. For example, when $\phi_c = 6$, the total equity premium is 2.42% but the part attributable to short-run risk is greater than this, equal to 4.38%.

Finally, panel (d) of Figure 5 shows that these challenges cannot be addressed by freely setting the correlation ρ_{c,x_c} between short-run consumption shocks $\varepsilon_{c,t}$ and long-run shocks $\varepsilon_{xc,t}$. Even when $\rho_{c,xc} = -1$, the term structure is essentially flat while the log risk premium on the market is -2.11%.

Alternative Parameterizations: The BRLI Model Figure 6 shows the role of key parameters in determining the slope and level of the term structure in the BRLI model. Panel (a) shows the crucial role of σ_d, the standard deviation of idiosyncratic dividend shocks, in determining the parameter ϕ_x^A in (28). Recall that ϕ_x^A is a measure of the exposure of estimated expected dividend growth, $\widehat{x}_{d,t}^A$, to estimated expected consumption growth, $\widehat{x}_{c,t}^A$, as measured by the coefficient from a projection of the former onto the latter. The panel plots ϕ_x^A as a function of σ_d, for various values of ϕ_c. Also plotted in this panel is the value of ϕ_x, which by definition doesn't vary with σ_d but governs the true exposure to of dividend growth to long-run consumption risk. The higher is σ_d the lower is ϕ_x^A. This represents a crucial difference with full information: the size of the volatility of the idiosyncratic component in dividend growth has no effect on the agent's perception of how exposed dividend growth is to long-run consumption risk under full information (axiomatically since the agent can distinguish the shocks), whereas it can have a large effect in the BRLI model. Panel a also shows that while ϕ_c plays a role in ϕ_x^A for small values of σ_d (with lower values of ϕ_c implying lower ϕ_x^A), in the empirically plausible region of $\sigma_d > 3$, the exposure of dividend growth to short-run consumption shocks is too small relative to the idiosyncratic dividend shocks to have a large effect on ϕ_x^A and therefore only the volatility σ_d of the idiosyncratic shock matters for ϕ_x^A. The bigger is σ_d, the smaller is ϕ_x^A.

Panel (b) of Figure 6 shows that a higher σ_d translates into a lower slope S. Under a range of plausible parameter values for σ_d, the slope S of the term structure is negative. Because the large idiosyncratic shocks to dividend growth cannot be distinguished from either of the other two shocks that drive consumption growth, dividend growth appears close to i.i.d. and shocks to the long-run component of dividend growth are virtually undetectable. Since long-run consumption shocks are undetectable, they are not priced and long maturity strips have lower risk premia than short maturity strips, leading S to be negative. Note that for this result, it is crucial that agents do not use information on consumption growth when modeling the dynamics of dividend growth. If they did, it would be possible to observe whether a given movement in dividend growth is correlated with consumption growth and therefore to distinguish idiosyncratic shocks from systematic ones. Since the i.i.d. idiosyncratic component of dividend growth would then be correctly identified as uncorrelated with consumption growth, it would no longer contaminate the decision maker's estimates of the systematic components of dividend growth that do command a risk premium. Under standard calibrations, the sum of the two systematic components is not close to i.i.d., instead displaying persistence (the moving average and autoregressive roots no longer almost cancel as in the case where all three components are conflated). Long-run shocks would effectively be priced, leading S to be positive. To summarize, it is the large value of σ_d, *coupled with* the bounded rationality behavioral assumption that leads agents to use a univariate specification for dividend growth that generates a negative term structure slope in this model. This result is exhibited below when we compare the implications of the limited information VARMA model–where consumption data are used in a system with dividend data–with those of the BRLI ARMA model.

Panel (c) of Figure 6 shows the impact of ϕ_c on the level of the term structure under the BRLI model. This panel plots $E_t[r_{t+1}^{(1)ex}]$ as a function of ϕ_c, for various values of σ_d. For comparison, it also plots $E_t[r_{t+1}^{(1)ex}]$ as a function of ϕ_c under full information. Two points bear noting. First, σ_d has virtually no affect on the risk premium of the one-period strip (the three lines corresponding to different values of σ_d lie on top of one another). The idiosyncratic shock creates no correlation with consumption so it generates no premium for strips of any maturity. Second, both the FI and BRLI models imply that the risk premium on the one-period strip is increasing in the exposure ϕ_c of dividend growth to short-run consumption risk. The premium in the BRLI model is, however, more steeply increasing in ϕ_c. This occurs for the reasons explained in the discussion of equation (29): ϕ_c controls the quantity of short-run risk in both models but short-run shocks have a higher effective price of risk in the BRLI model than in the FI model. This result implies not only that premia on one-period strips are higher under the BRLI model than under full information, it also implies that the market risk premium can be much higher–see (30). This can be seen in Panel (d) of Figure 6, which shows that the market risk-premium in the BRLI model is sharply increasing in ϕ_c, and sizable (e.g., greater than 5% per annum) at plausible values for ϕ_c such as the benchmark value of $\phi_c = 4$. This contrasts with Panel (c) of Figure 5, which shows that even with $\phi_c = 6$, the full information paradigm generates a risk-premium of just 2.4% when other parameters are set so that S is at most modestly negative (and the model is one of long-run insurance).

Finally, we turn to the magnitude of the zero coupon equity spread. Under the benchmark parameter values used for Figure 3, the absolute value of the spread in risk premia between short and long maturity equity in the BRLI model, while negative, is small, roughly 1% per annum. Figure 7 (left panel) reproduces the zero-coupon equity curve for this benchmark case and reports the log equity premium of 5.4%. Figure 7 also shows that, with a refined calibration, our limited information model of behavior is able to generate a much more negative S, while maintaining the sizable market risk premium. Figure 7 plots the term structure for the calibration given in column (4) of Table 1. This alternative calibration changes the benchmark values by reducing Ψ from 1.5. to 1, reducing σ from 0.78% to 0.76%, increasing ρ from 0.979 to 0.984 and increasing σ_{xc} from 4.40% to 4.84%. With these changes, the summary statistics for dividend growth are plausible, and within range of those recorded for the benchmark case and for the data, as can be seen from Table 1, column (d). But Figure 7 shows that the equity term spread is now much more negative ($S = -3.4\%$) with the log equity premium at 5% per annum (Figure 7). By comparison, under full information this refined calibration generates a spread $S = 0$, while the market risk premium is 2.8%.

We close this section by showing, in the right panel of Figure 7, the term structures under the VARMA limited information model, with the same two parameterizations. Under either parameterization, the slope of the term structure under the VARMA limited information specification is very close to that under full information. As explained above, this occurs because the use of consumption data in a system with dividend data allows the agent to distinguish idiosyncratic shocks from systematic shocks. The i.i.d. idiosyncratic component of dividend growth no longer contaminates the two systematic components, while the latter no longer look close to i.i.d. once combined. Long-run shocks are now effectively priced, leading S to be positive. Thus, the slope under VARMA limited information is either upward sloping or flat for these parameterizations, and the only difference between this case and full information is that the equity premium is slightly higher under VARMA limited information.

5 Conclusion

A recent strand of asset pricing literature emphasizes the potential role of long-run consumption risk for explaining salient asset pricing phenomena. A maintained assumption in the existing theoretical literature is that investors can directly observe the small long-run component of consumption risk and can distinguish its innovations from transitory shocks to consumption and dividend growth, even though in general these components cannot be identified from historical data. In this paper we study how equilibrium asset prices are affected if decision makers in the model must use consumption and dividend data to infer small long-run components in cash flows and consumption. We propose a model of behavior in which a representative decision maker optimizes based on a cash flow model that is both sparse in the sense that it ignores cross-equation restrictions that are difficult if not impossible to infer in finite samples, and fully identified from historical data.

A key result of the present study is that, under many parameter configurations, the bounded rationality limited information model studied here causes market participants to demand a higher premium for engaging in risky assets than would be the case under full information and rational expectations. Specifically, assets that have small exposure to long-run consumption risk but are highly exposed to short-run, even i.i.d., consumption risk can command high risk premia in the bounded rationality limited information case but not under full information. Thus the term structure of equity can be downward sloping in the former case, as in the data, but is upward sloping under full information.

In general, these patterns mean that the bounded rationality limited information specifications we explore are better able than their full information counterparts to reconcile the return properties of zero-coupon equity strips with those of the market index, an average across all strips. In a full information world, long maturity strips can be made less risky than short maturity strips only if the long-run component in consumption growth is a source of insurance rather than risk, implying that the long-run component generates a negative market risk premium. By contrast, the bounded rationality limited information framework we consider here matches empirical evidence for a downward sloping equity term structure simultaneously with a sizable equity market risk premium.

References

Ai, Hengjie, "Information Quality and Long-Run Risk: Asset Pricing Implications," *Journal of Finance*, 2010, *65*, 1333–1367.

Anderson, Evan W., Lars Peter Hansen, and Thomas J. Sargent, "Risk and Robustness in Equilibrium," 1998. Unpublished paper, University of Chicago.

Ang, Andrew and Maxim Ulrich, "Nominal Bonds, Real Bonds, and Equity," 2011. Unpublished paper, Graduate School of Business, Columbia University.

Bansal, Ravi and Amir Yaron, "Risks for the Long-Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, August 2004, *59* (4), 1481–1509.

____, Dana Kiku, and Amir Yaron, "Risks for the Long-Run: Estimation and Inference," 2007. Unpublished paper, Fuqua School, Duke University.

____, ____, and ____, "Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 2012, *1*, 183–221.

____, Robert F. Dittmar, and Christian T. Lundblad, "Consumption, Dividends, and the Cross-Section of Equity Returns," *Journal of Finance*, August 2005, *60* (4), 1639–1672.

____, ____, and Dana Kiku, "Cointegration and Consumption Risk in Equity Returns," *Review of Financial Studies*, 2009, *22*, 1343–1375.

Boguth, Oliver, Murray Carlson, Adlai Fisher, and Mikhail Simutin, "Dividend Strips and the Term Structure of Equity Risk Premia: A Case Study of Limits to Arbitrage," 2011. Unpublished paper, University of British Columbia, Sauder School of Business.

Campbell, John Y., "Consumption-Based Asset Pricing," in George Constantinides, Milton Harris, and Rene Stulz, eds., *Handbook of the Economics of Finance Vol. I-B*, North Holland, Amsterdam: Elsevier Science B.V., 2003, pp. 803–887.

Colacito, Ricardo and Mariano M. Croce, "Risks for the Long Run and the Real Exchange Rate," 2004. Unpublished paper, New York University.

Cornell, Bradford, "Risk, Duration, and Capital Budgeting: New Evidence on Some Old Questions," *The Journal of Business*, 1999, *72*, 183–200.

——, "Equity Duration, Growth Options and Asset Pricing," *The Journal of Portfolio Management*, 2000, *26*, 105–111.

Croce, Mariano M., "Welfare Costs in the Long Run," 2007. Unpublished Paper, UNC Chapel Hilll.

——, "Lon-Run Productivity Risk: A New Hope for Production-Based Asset Pricing?," *Journal of Monetary Economics,* forthcoming, 2014.

Da, Zhi, "Cash Flow, Consumption Risk, and the Cross-Section of Stock Returns," 2005. Unpublished paper, Norwestern University.

Dechow, Patricia M., Richard G. Sloan, and Mark T. Soliman, "Implied Equity Duration: A New Measure of Equity Risk," *Review of Accounting Studies*, 2004, *9*, 197–228.

Epstein, Larry and Stan Zin, "Substitution Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 1989, *57*, 937–968.

—— **and** —— , "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Investigation," *Journal of Political Economy*, 1991, *99*, 555–576.

Fama, Eugene F. and Kenneth R. French, "The Cross-Section of Expected Returns," *Journal of Finance*, 1992, *47*, 427–465.

Gabaix, Xavier, "A Sparsity-Based Model of Bounded Rationality," 2011. NBER Working Paper 16911.

Hansen, Lars Peter and Thomas J. Sargent, "Fragile Beliefs and the Price of Model Uncertainty," 2006. Unpublished Paper, New York University.

—— **, John Heaton, and Nan Li**, "Consumption Strikes Back?: Measuring Long-run Risk," *Journal of Political Economy*, 2008, *116* (2), 260–302.

Johannes, Michael, Lars Lochstoer, and Yiiiiiun Mou, "Learning About Consumption Dynamics," *Journal of Finance,* forthcomng, 2013.

Kaltenbrunner, Georg and Lars Lochstoer, "Long-Run Risk Through Consumption Smoothing," *Review of Financial Studies*, 2010, *23*, 3141–3189.

Kiku, Dana, "Is the Value Premium a Puzzle?," 2005. Unpublished paper, Duke University.

Lettau, Martin and Jessica A. Wachter, "Why is Long-Horizon Equity Less Risky? A Duration Based Explanation of the Value Premium," *Journal of Finance*, February 2007, *LXII* (1), 55–92.

—— **and** —— , "The Term Structures of Equity and Interest Rates," 2009. NBER Working Paper No. 14698.

Li, George, "Information Quality, Learning and Stock Market Returns," *Journal of Financial and Quantitative Analysis*, 2005, *40* (3), 595–620.

Ludvigson, Sydney C., "Advances in Consumption-Based Asset Pricing: Empirical Tests," in George Constantinides, Milton Harris, and Rene Stulz, eds., *Handbook of the Economics of Finance Vol. II,* forthcoming, North Holland, Amsterdam: Elsevier Science B.V., 2012.

Lustig, Hanno and Stijn Van Nieuwerburgh, "Quantitative Asset Pricing Implications of Housing Collateral Constraints," 2006. Unpublished paper, NYU Stern School.

Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-Jorgensen, "Long-run Stockholder Consumption Risk and Asset Returns," *Journal of Finance*, 2009, *64*, 2427–2479.

Parker, Jonathan, "The Consumption Risk of the Stock Market," *Brookings Papers on Economic Activity*, 2001, *2*, 279–348.

____ **and Christian Julliard**, "Consumption Risk and the Cross-Section of Expected Returns," *Journal of Political Economy*, February 2004, *113* (1), 185–222.

van Binsbergen, Jules H., Michael W. Brandt, and Ralph S. J. Koijen, "On the Timing of Pricing of Dividends," 2010. Unpublished paper, Booth School, University of Chicago.

____ **, Wouter H. Hueskes, Ralph S. J. Koijen, and Evert B. Vrugt**, "Equity Yields," 2012. Unpublished paper, Northwestern University.

Veronesi, Pietro, "How Does Information quality Affect Stock Returns?," *Journal of Finance*, 2000, *55* (2), 807–837.

Weil, Philippe, "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 1989, *24* (3), 401–421.

Zhang, Lu, "The Value Premium," *Journal of Finance*, 2005, *60* (1), 67–103.

TABLE 1: CONSUMPTION AND DIVIDEND PROPERTIES

	Data	(1)	(2)	(3)	(4)
Ψ		1.50	1.50	1.50	1.00
ϕ_x		3.00	3.00	1.00	1.00
ϕ_c		0.00	4.00	4.00	4.00
σ		0.78%	0.78%	0.78%	0.76%
σ_{xc}		4.40%	4.40%	4.40%	4.84%
ρ		.979	.979	.979	.984
StD[Δd]	11.49	11.36	14.53	13.64	13.34
ACF[Δd]	0.21	0.34	0.29	0.22	0.23
Corr[Δd, Δc]	0.55	0.27	0.70	0.59	0.58
StD[Δc]	2.93	2.80	2.80	2.80	2.99
ACF[Δc]	0.49	0.43	0.43	0.43	0.50

Notes - This table reports consumption and dividend growth statistics from four monthly calibrations with alternate parameter configurations given in the top portion of the table. Benchmark parameter values are in column (3) $\Psi = 1.5$, $\gamma = 10$, $\delta = 0.999$, $\mu_c = \mu_d = .15\%$, $\rho_c = \rho_d = 0.979$, $\sigma = .78\%$, $\sigma_{xc} = 4.4\%$, $\sigma_d = 4.5$, $\phi_x = 1.0$, $\sigma_{xd} = \phi_x \sigma_{xc}$, $\phi_c = 4$, $\rho_{x_c,x_d} = 1$, $\rho_{c,x_c} = \rho_{c,d} = \rho_{d,xc} = \rho_{d,xd} = \rho_{c,xd} = 0$. For other columns, deviations from this benchmark are reported in the top portion of the table. The calibration in column (1) and the statistics under "Data"are from Bansal and Yaron (2004). The entries for the cash-flow models are obtained by repetitions of small-sample simulations. Simulated monthly data are time aggregated to annual data, as in Bansal and Yaron (2004).

TABLE 2: ASSET PRICING IMPLICATIONS: FI VS. BRLI

Row	Model		$E(P/D)$		$E(r_d - r_f)$		$E(r_f)$		$\sigma(r_d - r_f)$	
	ϕ_x	ϕ_c	FI	BRLI	FI	BRLI	FI	BRLI	FI	BRLI
1	1	2.5	75.42	107.75	2.73	2.93	1.38	0.78	14.02	14.02
2	1	6.0	36.32	17.45	5.24	8.62	1.38	0.78	20.34	19.41
3	2	2.5	28.06	73.48	5.16	3.79	1.38	0.78	15.40	16.75
4	2	6.0	20.38	16.00	7.57	9.50	1.38	0.78	21.23	21.16
5	3	2.5	18.67	55.47	7.33	4.91	1.38	0.78	17.67	20.43
6	3	6.0	15.02	14.45	9.66	10.73	1.38	0.78	22.80	23.68

Notes - This table reports financial statistics of the model with full information (FI) and limited information ARMA(1,1)-based signal extraction (BRLI), for varying degrees of exposure to the long-run and short-run risk components, governed by ϕ_x and ϕ_c, respectively. All other parameters are set as in the benchmark calibration–see notes Table 1. $E(r_d - r_f)$ and $E(r_f)$ denote the annual average of the equity excess return and the risk-free rate, respectively. The standard deviation of the annual equity excess return is denoted by $\sigma(r_d - r_f)$. $E(P/D)$ is the annual price-dividend ratio.

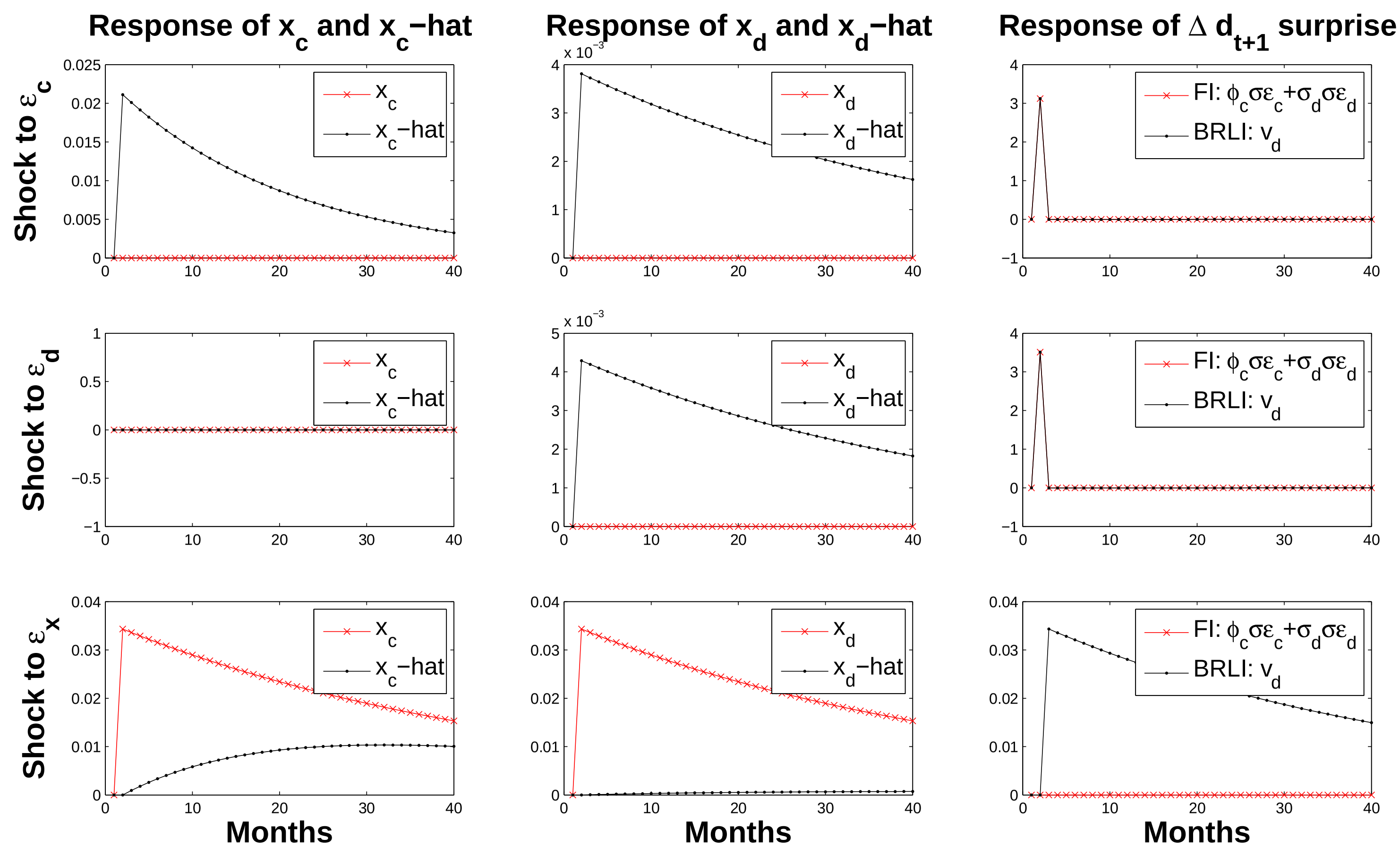


FIGURE 1: RESPONSES OF CASH FLOW FORECASTS AND SURPRISES: FI VS. BRLI

Notes - The figure shows the 40-month impulse response of variables to a one-standard deviation innovation in the shock labeled at each row. The vertical axis represents monthly percent deviations of variables from steady state. Variables denoted with "hat" correspond to those from the BRLI model and are based on ARMA(1,1) estimations for consumption and dividend growth. Variables without a "hat" are from the full information (FI) benchmark. The responses are based on the benchmark monthly calibration: $\mu_c = \mu_d = .15\%$, $\rho_c = \rho_d = 0.979$, $\sigma = .78\%$, $\sigma_{xc} = 4.4\%$, $\sigma_d = 4.5$, $\phi_x = 1.0$, $\sigma_{xd} = \phi_x \sigma_{xc}$, $\phi_c = 4$ and $\rho_{x_c,x_d} = 1$, $\rho_{c,x_c} = \rho_{c,d} = \rho_{d,xc} = \rho_{d,xd} = \rho_{c,xd} = 0$.

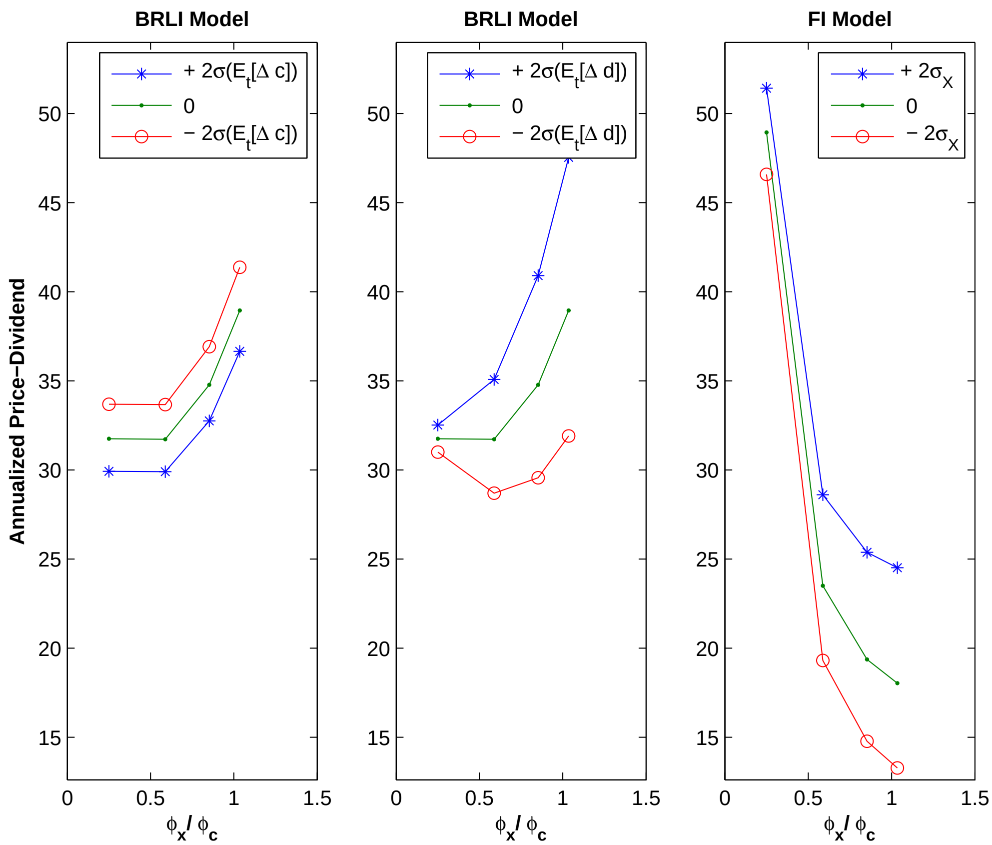


FIGURE 2: PRICE-DIVIDEND RATIOS: FI VS. BRLI MODELS

Notes - This figure displays price-dividend ratios at steady state, and plus/minus two standard deviations of the state variables(s) around steady state, as a function of the relative exposure to long-run risk, governed by ϕ_x, and to short-run risk, governed by ϕ_c. Held fixed is the 2-year variance of dividend growth attributable to the consumption innovations. All parameters other than ϕ_x and ϕ_c are set at benchmark values given in notes to Table 1. The BRLI case refers to the ARMA(1,1)-based signal extraction procedure.

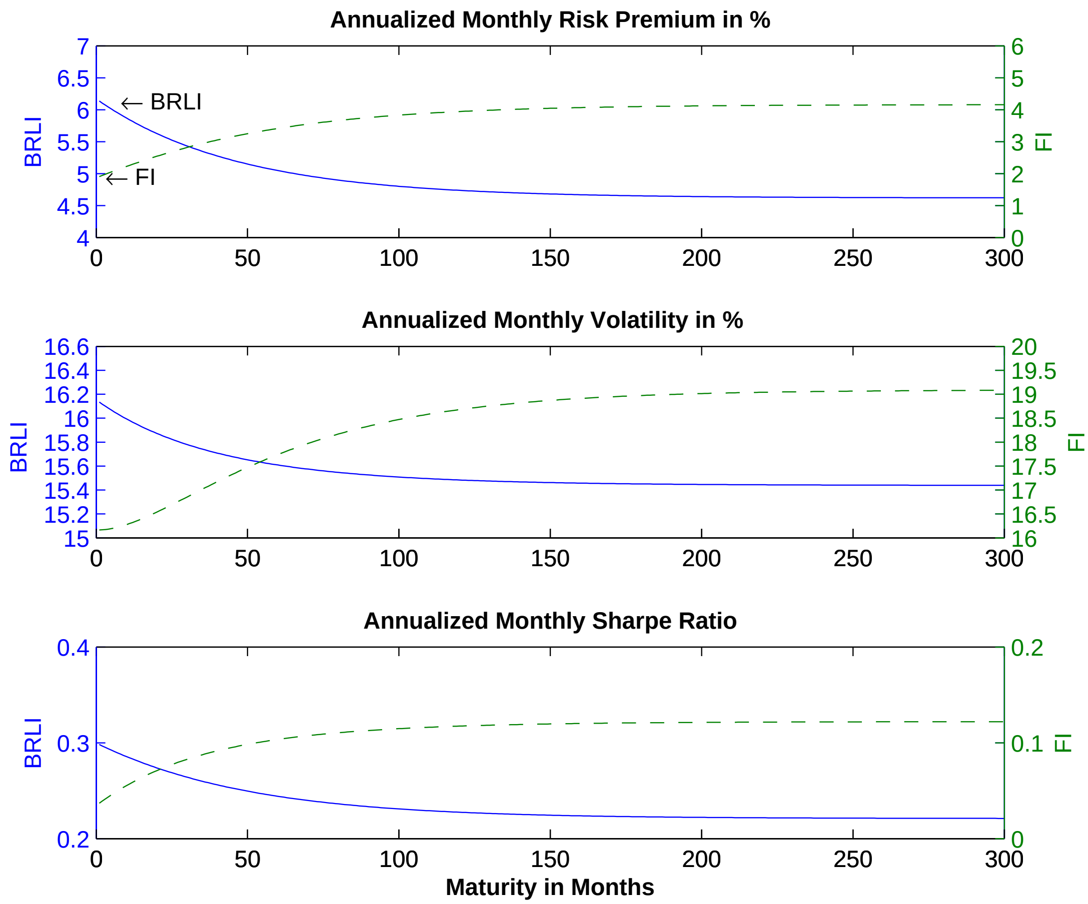


FIGURE 3: ZERO-COUPON EQUITY: FI VS. BRLI

Notes - This figure shows annualized log real risk-premia on zero-coupon equity, $E[r_{t+1}^{(n)} - r_{f,t+1}]$, as a function of maturity, n, in months. The middle panel shows the standard deviation of excess returns on zero-coupon equity. The bottom panel shows the Sharpe ratio. Parameter values are set at benchmark values given in notes to Table 1. The BRLI case refers to the ARMA(1,1)-based signal extraction procedure.

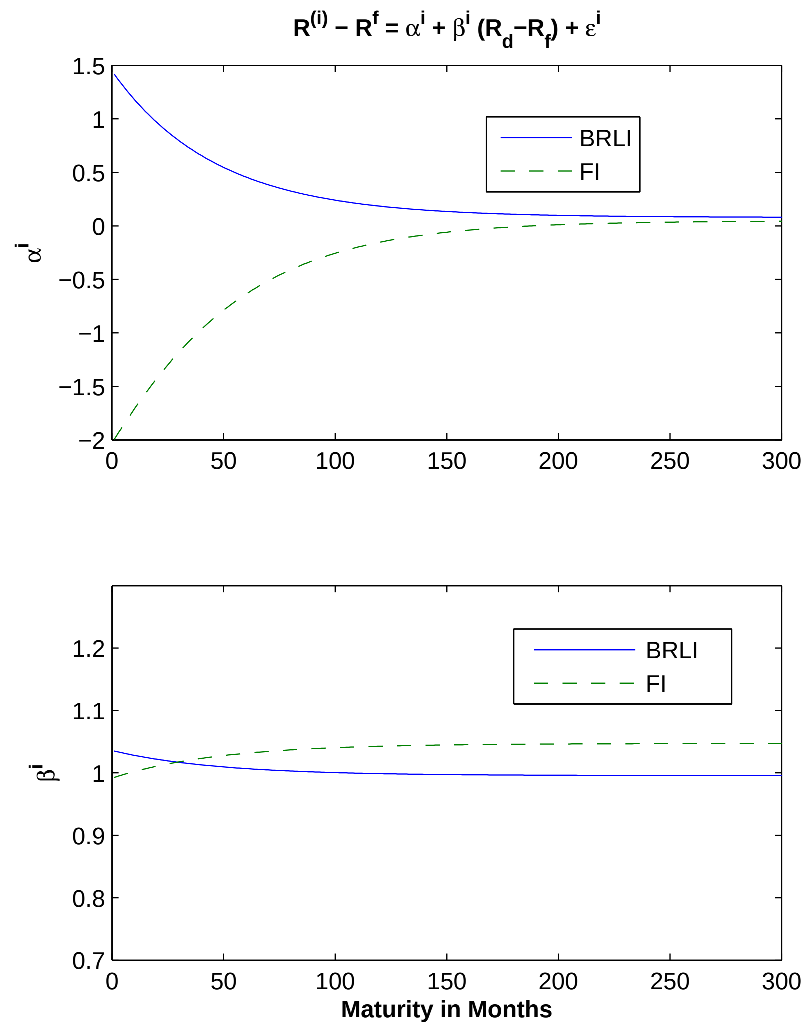


FIGURE 4: CAPM REGRESSIONS FOR ZERO-COUPON EQUITY: FI VS. BRLI

Notes - The top panel shows the intercept from regressions of annual zero-coupon equity excess returns on the excess return of the market, as a function of maturity in months; the bottom panel shows the slope coefficient from the same regression. Parameter values are set at benchmark values given in the notes to Table 1. The BRLI case refers to the ARMA(1,1)-based signal extraction procedure.

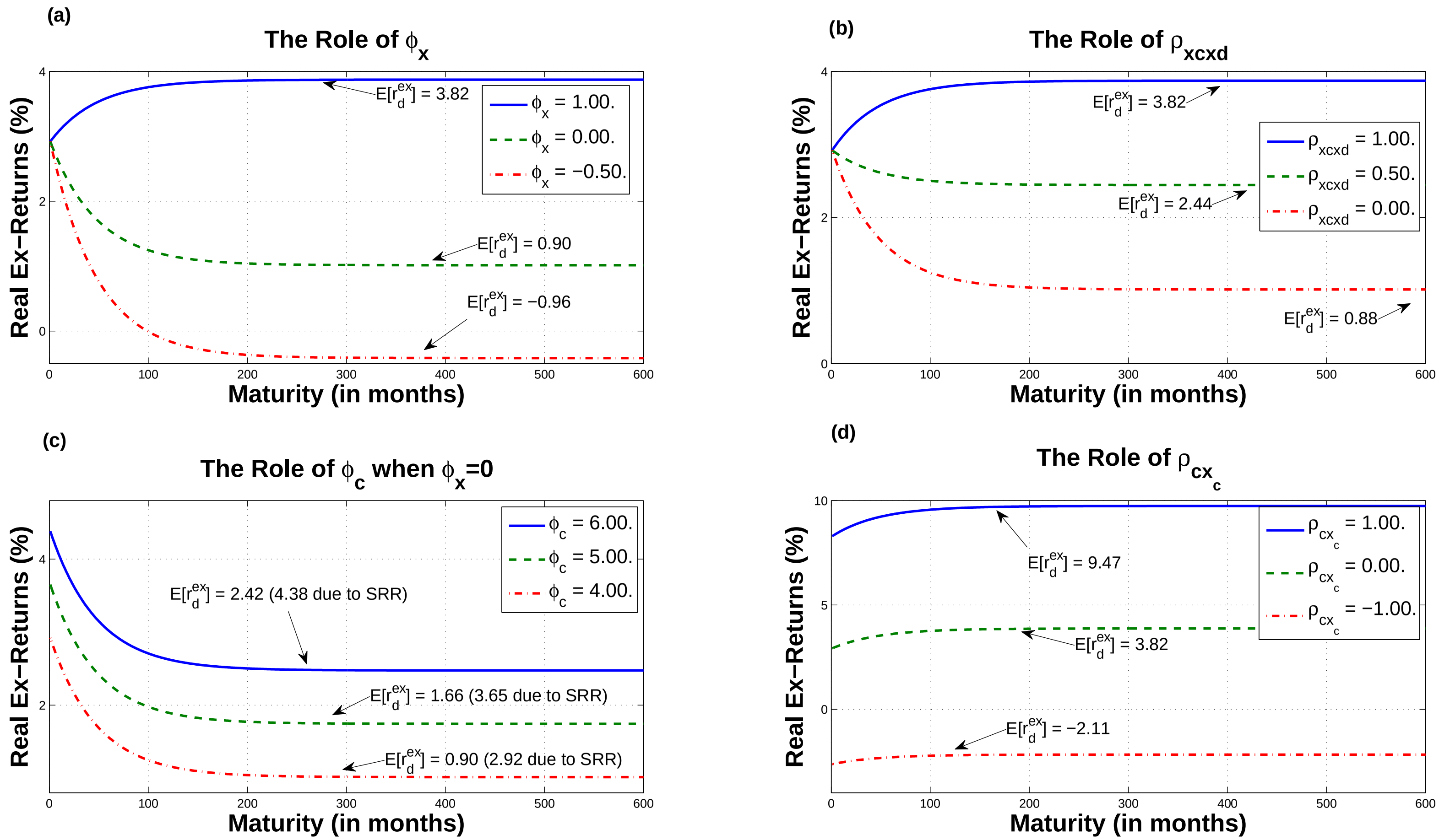


FIGURE 5: DETERMINANTS OF TERM STRUCTURE OF EQUITIES UNDER THE FI MODEL

Notes - This figure shows annualized real risk-premia on zero-coupon equity, $E[r_{t+1}^{(n)} - r_{f,t+1}]$, as a function of maturity n under Full Information (FI). The data generating process is described by the system of equations (9)–(13) and is assumed to be known under FI. In each panel we vary the parameter indicated and keep all the others unchanged at benchmark values given in notes to Table 1. In Panel A, we vary the exposure of dividends to long-run risk, ϕ_x, and assume $\rho_{xcxd} = corr(x_{d,t}, x_{c,t}) = 1$. In Panel B, we allow the correlation between consumption and dividends long-run components, ρ_{xcxd}, to be less than perfect. In Panel C, we assume that dividends have no exposure to consumption long run-risk, $\phi_x = 0$, and run sensitivity analysis with respect to the dividends exposure to short-run consumption shocks, ϕ_c. We also report the portion of market risk-premium attributable to short-run (i.i.d.) consumption risk, $\epsilon_{c,t+1}$. In Panel D, we study the relevance of the correlation of short- and long-run news, $\rho_{cx_c} = corr_t(x_{c,t+1}, \epsilon_{c,t+1})$.

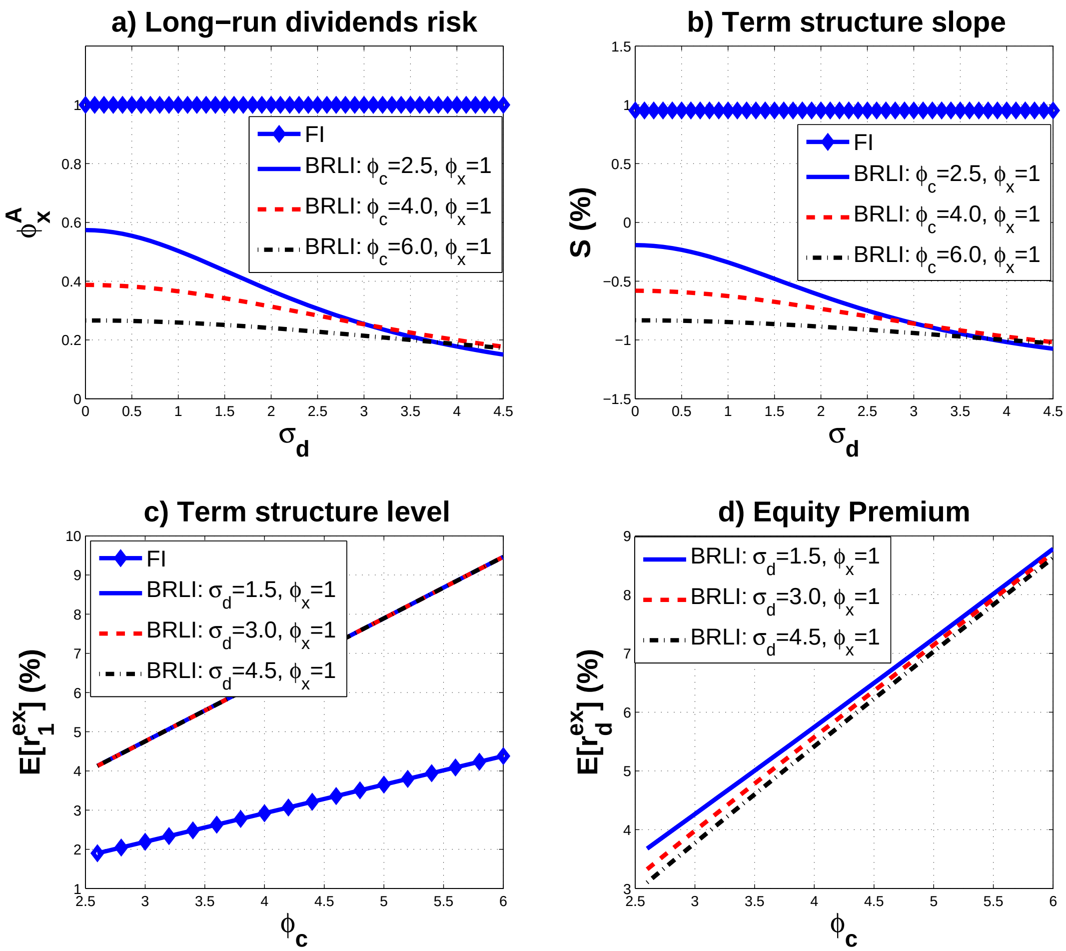
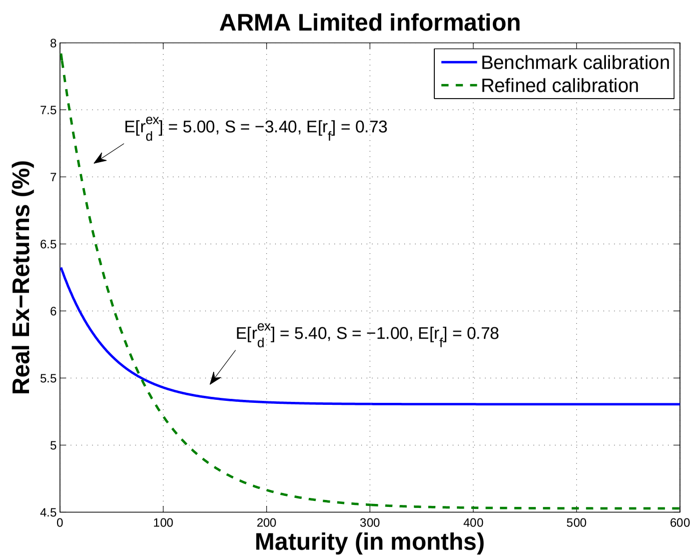
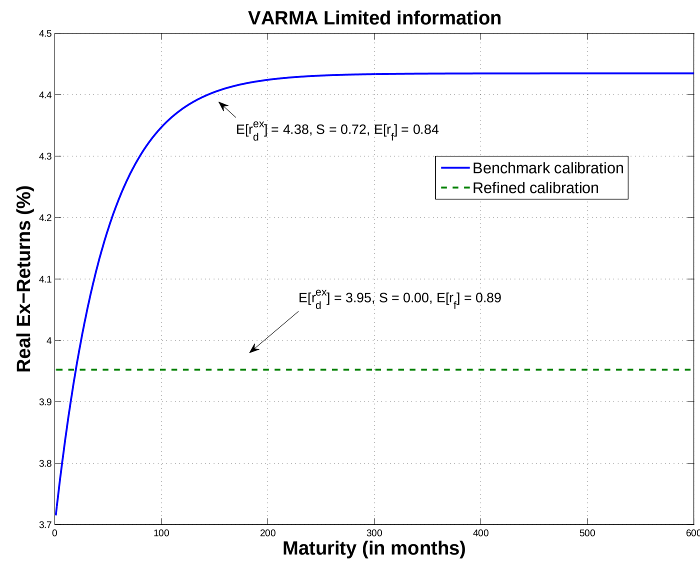


FIGURE 6: TERM STRUCTURE OF EQUITY UNDER THE BRLI MODEL.

Notes - The data generating process is described by the system of equations (9)–(13) and is assumed to be known under full information (FI). All parameters are set at benchmark values given in notes to Table 1. In Panel A, we vary both the exposure of dividends to short-run consumption shocks, ϕ_c, and the volatility of dividends-specific shocks, σ_d. We compare the true dividends long-run risk exposure, ϕ_x, with its counterpart, ϕ_x^A, computed under the BRLI case (see equations (17)–(20) and (29)). In panel B, we depict the term structure slope, $S = E[r_\infty - r_1]$, for the same parameters chosen in Panel A to show the direct correspondence between S and ϕ_x (ϕ_x^A) in the FI (BRLI) case. Panel B shows that σ_d is the key determinant of the term structure slope. In Panel C and D, we show the intercept, $E[r_1^{ex}]$, and the value-weighted average level of the term structure (i.e., the aggregate equity premium, $E[r_d^{ex}]$), respectively. Panels C and D show that under BRLI the level of the term structure is primarily determined by ϕ_c.



(a) ARMA-Based Signal Extraction



(b) VARMA-Based Signal Extraction

FIGURE 7: TERM STRUCTURE OF EQUITY: ARMA VS. VARMA

Notes - This figure shows annualized real risk-premia on zero-coupon equity, $E[r_{t+1}^{(n)} - r_{f,t}]$, as a function of maturity n. $E[r_d^{ex}]$, $E[r_{f,t}]$, and S denote the aggregate equity risk premium, the average risk-free rate, and the equity term structure spread, respectively. The true data generating process is given by the system of equations (9)–(13). In Panel A, equilibrium asset prices are obtained using an ARMA-based signal extraction procedure (BRLI), defined by the system of equations (17)–(20). In Panel B, asset prices are computed using the VARMA-based signal extraction procedure described in equation (14). In both panels, the "Benchmark calibration" is given in the notes to Table 1. The "Refined calibration" refers to the calibration in column (4) of Table 1, and it features a slightly lower IES and a more persistent long-run risk process.