Investor Information, Long-Run Risk, and the Duration of Risky Cash Flows

Mariano M. Croce* NYU

Martin Lettau† NYU, CEPR and NBER

Sydney C. Ludvigson‡ NYU and NBER

Comments Welcome

First draft: August 15, 2005

This draft: October 11, 2006

*Department of Economics, New York University, 269 Mercer Street, 7th Floor, New York, NY 10003; Email: massimiliano.croce@nyu.edu; http://homepages.nyu.edu/~mmc287

†Department of Finance, Stern School of Business, New York University, 44 West Fourth Street, New York, NY 10012-1126; Email: mlettau@stern.nyu.edu; Tel: (212) 998-0378; http://www.stern.nyu.edu/~mlettau

‡Department of Economics, New York University, 269 Mercer Street, 7th Floor, New York, NY 10003; Email: sydney.ludvigson@nyu.edu; Tel: (212) 998-8927; Fax: (212) 995-4186; http://www.econ.nyu.edu/user/ludvigsons/

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 Foundation and the CV Starr Center at NYU. The authors thank Dave Backus, John Y. Campbell, Timothy Cogley, Michael Gallmeyer, Lars Hansen, John Heaton, Thomas Sargent, Stijn Van Nieuwerburgh and seminar participants at the 2006 Society for Economic Dynamics conference, the summer 2006 NBER Asset Pricing meeting, NYU, Texas A&M, UCLA, and UNC Chapell Hill for helpful comments. Any errors or omissions are the responsibility of the authors.

Investor Information, Long-Run Risk, and the Duration of Risky Cash Flows

Abstract

We study the role of information in asset pricing models with long-run cash flow risk. To illustrate the importance of the information structure, we show how the implications of the long-run risk paradigm for the cross-sectional properties of stock returns and cash flow duration are affected by information. When investors can fully distinguish short- and long-run consumption risk components of dividend growth innovations (*full information*), only exposure to long-run consumption risk generates significant risk premia, implying that high-return value stocks are long-duration assets, contrary to the historical data. By contrast, when investors observe the change in consumption and dividends each period but not the individual components of that change (*limited information*), exposure to short-run risk can generate large risk premia, so that high-return value stocks are short-duration assets while low-return growth stocks are long-duration assets, as in the data. We also show that, in order to explain empirical finding that long-horizon equity is less risky than short-horizon equity, the properties of the cash flow model and the values of primitive preference parameters must be quite different from those emphasized in the existing long-run risk literature.

JEL: G10, G12

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 perhaps best 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 maintained assumption in this literature is that investors can directly observe the small long-run component and distinguish its innovations from transitory shocks to consumption and dividend growth. We refer to this assumption as the *full information* specification. While this is a natural starting place and an important case to understand, in this paper we consider two alternative *limited information* specifications in which market participants are faced with distinct signal extraction problems. In each case, they can observe the change in consumption and dividends each period, but they cannot observe the individual components of that change. The two cases differ in the how the information is used, and therefore in the form of signal extraction problem solved.

Information about long-run cash flow risk is likely to be limited. In finite samples it is difficult or impossible to distinguish statistically between a purely i.i.d. process and one that incorporates a small persistent component. Hansen et al. (2005) use samples of the size currently available and show that the long-run riskiness of cash flows is hard to measure econometrically; they argue that such statistical challenges are likely to plague market participants as well as econometricians. Perhaps most important, for specifications of the dividend process that have been studied in the long-run risk literature, the distinct roles of persistent and transitory shocks cannot be separately identified from the history of consumption and dividend data. 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 covary with one another, even though the data

[1] A growing body of theoretical and empirical work has been devoted to studying the role of long-run risk in consumption and dividend growth for explaining asset pricing behavior. See Parker (2001); Parker and Julliard (2004); Colacito and Croce (2004); Bansal, Dittmar and Kiku (2005); Hansen, Heaton and Li (2005); Kiku (2005); Malloy, Moskowitz and Vissing-Jorgensen (2005); Bansal, Dittmar and Lundblad (2006) Hansen and Sargent (2006).

give us no guide for observing these components separately.

As an illustration of the potential importance of the information structure, we study the implications of the long-run risk paradigm for the cross-sectional properties of stock returns and cash flow duration. Empirical evidence indicates that assets with low ratios of price to measures of fundamental value (value stocks) have higher average returns than assets with high ratios of price to fundamental value (growth stocks) (Graham and Dodd (1934); Fama and French (1992)). The long-run risk explanation implies that assets with high average returns command a high risk premium because they are more exposed to long-run cash flow risk. Accordingly, this line of thought implies that value stocks must be more exposed to long-run cash flow risk than are growth stocks. At the same time, however, a second strand of empirical evidence finds that the cash flow duration of value stocks is considerably shorter than that of growth stocks (Cornell (1999, 2000); Dechow, Sloan and Soliman (2004); Da (2005)). Shorter duration means that the timing of value stocks' cash flow fluctuations is weighted more toward the near future than toward the far future, whereas the opposite is true for growth stocks. The duration perspective of equity seems to suggest that value stocks are *less* exposed to long-run cash flow risk than are growth stocks. An unanswered question is whether the long-run risk perspective of value and growth assets can be reconciled with the seemingly contradictory cash flow duration evidence. We argue here that what may seem to be small changes in the information structure can have important implications for such questions.

For example, it may seem apparent that in models where long-run consumption risk plays a central role in the determination of risk premia, long duration equity should be riskier than short duration equity, thereby contradicting empirical evidence on the return and duration properties of value and growth stocks. We show that, while this is true under full information, it is not necessarily true under limited information. When information is limited, the long-run risk paradigm can be made consistent with evidence that long-horizon equity commands a lower risk premium than short-horizon equity.

We study a model in which the dividend growth rates of individual assets are differentially exposed to two systematic risk components driven by aggregate consumption growth, in addition to a purely idiosyncratic component uncorrelated with aggregate consumption. One is a small but highly persistent (long-run) component as in Bansal and Yaron (2004), while the second is a transitory (short-run) i.i.d. component with much larger variance. 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). With recursive utility, investors are not indifferent to the intertemporal composition of risk, implying that the relative exposure to short- versus long-run risks has a non-trivial influence on risk premia.

In order to isolate the endogenous relation between cash-flow duration and risk premia in models with long-run consumption risk, we model firms as differing only in the timing of their cash flows. This may be accomplished by recognizing that an equity claim is a portfolio of zero-coupon dividend claims with different maturities. It follows that long duration assets can be modeled as equity with a greater share of long-horizon zero-coupon aggregate dividend claims than short duration assets.

We find that, in long-run risk models with full information, assets that have low price-dividend ratios and high risk premia (value stocks) will endogenously be long-duration assets, while those with high price-dividend ratios and low risk premia (growth stocks) will be short-duration assets, contrary to the historical data. Here the endogenous relation between cash flow duration and risk premia goes the wrong way. By contrast, under limited information, value stocks with low price-dividend ratios and high risk premia can be (endogenously) short-duration assets, while growth stocks with high price-dividend ratios and low risk premia are long-duration assets, in line with the data. These results show that limited information can be an important source of additional risk, and it can completely reverse the type of asset that commands a high risk premium.

The intuition for this result is straightforward. When investors can observe the long-run component in cash flows–in which a small shock today can have 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-duration assets must in equilibrium command high risk premia. But under limited information, the opposite can occur. Assets with high exposure to short-run (i.i.d.) consumption shocks may command high risk premia because investors' optimal forecasts of the long-run component assign some weight to the possibility that shocks to the i.i.d. component will be persistent. At the same time, assets with low exposure to short-run consumption shocks but high exposure to the small long-run component may command small risk premia because fluctuations in those assets appear dominated by the large idiosyncratic cash flow innovations that carry no risk premium.

A counter-intuitive aspect of these results is that limited information can, under some parameter configurations, lead market participants to command a higher risk premium than they would under full information. For all of these results, the presence of a small long-run risk component is central to delivering high risk premia, just as it is in the full information specifications. But the limited information specifications generate a richer set of results, in which the relative exposure of cash flows to shocks with different degrees of persistence, and investors' perceptions of these shocks as seen through an optimal filtering lens, matters as much for risk premia as an asset's exposure to long-run consumption risk.

The main purpose of this paper is to make these points qualitatively. Nevertheless, we show that some calibrations of the limited information models we explore can match the properties of cross-sectional data quantitatively. In general, risk aversion must be sufficiently high and/or the volatility of the long-run risk component cannot be too small. In these cases, the limited information specifications are not only consistent with the cash flow duration properties of value and growth stocks and a sizable value premium, they also explain the higher empirical Sharpe ratios of value stocks and the failure of the capital asset pricing model (CAPM) of Sharpe (1964) and Lintner (1965) to account for the value premium. In addition, the limited information model is consistent with the ability of high-minus-low factor (HML) of Fama and French (1993) to explain the value premium.

The rest of this paper is organized as follows. The next section discusses related literature not discussed above. Section 3 presents the asset pricing model, the model for cash flows, and the informational assumptions. We consider two cases of limited information to illustrate how differences in the amount of information as well as in the way it is processed can affect asset prices. One case is based on the estimation of separate univariate processes for consumption and dividend growth, while the other is based on system estimation for both series. Section 4 shows how these two specifications of limited information influence equilibrium asset returns and compares them to a full information benchmark. A theme of this section is that the duration evidence for value and growth assets can be used to distinguish among long-run risk models that differ according to their information structures, cash flow properties, and primitive preference parameter values. Section 5 concludes.

2 Related Literature

In terms of motivation, our paper is most closely related to recent work by Hansen and Sargent (2006). Like us, Hansen and Sargent are concerned about the agent's ability to observe the long-run risk component in aggregate cash flows. The agents in their model form decision rules that are robust to misspecification of their approximating model of cash-flows. They assume that a representative consumer assigns positive probability to both an i.i.d. model for consumption growth and a model in which consumption growth contains a small predictable component, and learns which model is a better description of the data over time. They show that uncertainty over which model of consumption is correct can add a volatile and time-varying component to the stochastic discount factor, helping to explain both the high mean and time variation in the Sharpe ratio of aggregate stock market returns.

Our paper, while similar in motivation to Hansen and Sargent (2006), differs along a number of dimensions. The most important is that Hansen and Sargent focus on the role of learning with robust preferences, while we focus on the agent's information structure. Thus, in order to isolate the effects on asset returns arising purely from the information structure, we assume that agents must solve a signal extraction problem, but we do not impose a preference for robustness in their utility maximization. We consider our investigation to be complementary to that of Hansen and Sargent and discuss in the conclusion how concerns for robustness could be included in the analysis. In addition, Hansen and Sargent study their model's implications for the unobservable return on a claim to aggregate consumption, whereas we investigate a levered equity claim similar to that studied in Bansal and Yaron (2004). This distinction is important because the parameters of the levered equity claim process we study are in general unidentified, implying that even an agent armed with an infinite sample of consumption and dividend data can never learn the correct cash flow model and must always perform a signal extraction problem. It is this signal extraction problem in the presence of long-run risk that our analysis is designed to focus on. Finally, Hansen and Sargent study the time-series properties of the aggregate consumption claim, whereas we model multiple risky assets that differ according to the timing of their cash flows. This allows us to study the cross-sectional properties of the model and its relation to cash flow duration.

Our paper is also related to a recent literature that tries to reconcile the cross-sectional properties of equity returns simultaneously with the cash flow duration properties of value and growth assets. Lettau and Wachter (2006) use techniques from the affine term structure literature to develop a dynamic risk-based model that captures the value premium, the cash flow duration properties of value and growth portfolios, and the poor performance of the CAPM. However, Lettau and Wachter forgo modeling preferences and instead directly specify the stochastic discount factor. An essential element of their results is that the pricing kernel must contain state variables that can be at most weakly correlated with aggregate fundamentals. (Lettau and Wachter set this correlation to zero in their benchmark model.)

By contrast, models that specify preferences as a function of aggregate fundamentals often have difficulty matching the cross-sectional properties of stock returns. For example, the habit model of Campbell and Cochrane (1999) has received significant attention for its ability to explain the time-series properties of aggregate stock market returns. But Lettau and Wachter (2006) and Wachter (2006) show that the Campbell and Cochrane model implies that assets with greater risk premia are long duration assets, rather than short duration assets as in the data for value and growth portfolios. Santos and Veronesi (2005) modify the Campbell and Cochrane model by adding cash flow risk for multiple risky securities and successfully generate a value premium for short-horizon assets. However, they also find that the cross-sectional dispersion in cash flow risk required to explain the magnitude of the premium is implausibly high. 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. Other researchers have studied the cross-sectional properties of stock returns in production-based asset pricing models. 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.

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. (2005) and Malloy et al. (2005)).

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 intertemporal elasticity of substitution Ψ.

Let $P_{j,t}^{D}$ denote the ex-dividend price of a claim to an asset that pays a dividend stream $\{D_{j,t}\}_{t=1}^{\infty}$ measured at the end of time t, and 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

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

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

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 $R_{t+1}^{f} \equiv \left(E_t\left[M_{t+1}\right]\right)^{-1}$.

3.1 The Cash Flow Model and Informational Assumptions

Equities are modeled as claims to a dividend process, specified below. We first describe the form of the stochastic process for aggregate dividend growth, and explain later how this form can be adapted to model individual asset's cash flows. Let D_t denote the aggregate dividend at time t, and let P_t^D denote the ex-dividend price of a claim to the asset that pays the stream $\{D_t\}_{t=1}^{\infty}$. We use lower case letters denote log variables, e.g., $\log\left(C_t\right) \equiv c_t$.

We seek a model for equity cash flows that allows dividend growth rates to be potentially exposed to both transitory and persistent sources of consumption risk that drive M_{t+1}, as well as to purely idiosyncratic shocks that command no risk premium. Denote the conditional means of consumption and dividend growth as $x_{c,t}$ and $x_{d,t}$, respectively. To model the persistent fluctuations in consumption risk, we follow Bansal and Yaron (2004) and assume that consumption and dividend growth rates contain a single, common predictable component with an autoregressive structure. In addition, we assume here that dividend growth rates may also be exposed to transitory (i.i.d.) consumption risk. These assumptions give rise to the following dynamic system:

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

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

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

$$\varepsilon_{c,t+1}, \varepsilon_{d,t+1}, \varepsilon_{xc,t} \sim N.i.i.d\,(0,1)\,. \tag{7}$$

Note that the conditional mean of dividend growth is proportional to the conditional mean of consumption growth $x_{d,t} = \phi_x x_{c,t}$, a specification that follows much of the long-run risk literature. We refer to the system (4)-(6), with the correlation structure (7), as the *true data generating process.*

3.1.1 Full Information

In the existing literature on long-run consumption risk (e.g., Bansal and Yaron (2004)), it is commonplace to assume that agents can directly observe the cash flow processes, including the latent conditional means $x_{c,t}$ and $x_{d,t}$. We refer to this as the *full information* assumption. The term in equation (4) labeled "LR risk" captures the small long-run risk component emphasized 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 can observe $x_{c,t}$. In this paper we also allow dividend growth to be exposed to transitory consumption shocks, by introducing the component $\sigma\varepsilon_{c,t+1}$ in (5). We refer to this component as the short-run risk component, labeled "SR risk," since its correlation with the stochastic discount factor contributes to the systematic riskiness of the dividend claim, but its purely i.i.d. nature makes that risk short-lived. Because the innovation $\varepsilon_{d,t+1}$ is uncorrelated with consumption growth, it does not contribute to systematic risk. The loadings ϕ_x and ϕ_c govern the exposure of dividend growth to long-run and short-run consumption risk, respectively.

Full information is a strong assumption, since the conditional means $x_{c,t}$ and $x_{d,t}$ are latent and (as discussed below), the system (4)-(7) cannot in general be observed from historical data on consumption and dividends. The full information assumption therefore implies that market participants have more information than do econometricians with historical data on consumption and dividends.[2]

[2] Notice that, in the model, asset prices should not contain additional information about $x_{c,t}$. In the model, prices are determined endogenously from exogenously given consumption and dividend processes. Thus, if agents have access only to historical dividend and consumption data, endogenous prices will reflect only that information and not additional information about $x_{c,t}$.

3.1.2 Limited Information

Under limited information, investors observe historical consumption and dividend data, but they do not directly observe the latent variables $x_{c,t}$ and $x_{d,t}$ or the innovations of the true data generating process (4)-(7). Armed with historical data on dividends and consumption, investors could in principal use Maximum Likelihood and the Kalman filter to estimate a general dynamic system in which innovations in the long-run components of consumption and dividend growth $x_{c,t}$ and $x_{d,t}$ have arbitrary correlations with each other and with the i.i.d. innovations of these series. As shown in the Appendix, however, in the absence of apriori restrictions on the parameters, the general dynamic system 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. Nevertheless, agents with access to historical data on consumption and dividend growth can directly estimate Wold representations for these series, as long as they follow covariance-stationary processes. In this paper, we study two variants of what we refer to as *limited information* in which investors estimate either univariate or multivariate Wold representations for consumption and dividend growth.

The first variant is based on a multivariate Wold representation for the dynamic system (4)-(6), which can be written as a $VARMA(1,1)$ process:

$$\begin{bmatrix} \Delta c_{t+1} \\ \Delta d_{t+1} \end{bmatrix} = \begin{bmatrix} \mu_c (1-\rho) \\ \mu_d (1-\rho) \end{bmatrix} + \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \Delta c_t \\ \Delta d_t \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v^V_{c,t+1} \\ v^V_{d,t+1} \end{bmatrix} - \begin{bmatrix} b_{cc} & b_{cd} \\ b_{dc} & b_{dd} \end{bmatrix} \begin{bmatrix} v^V_{c,t} \\ v^V_{d,t} \end{bmatrix}. \tag{8}$$

The i.i.d. innovations $v^V_{c,t+1}$ and $v^V_{d,t+1}$ will in general be correlated, and are composites of the underlying innovations in (4)-(6). In addition, the parameters b_{cc},...,b_{dd} and the variance-covariance matrix of $v^V_{c,t+1}$ and $v^V_{d,t+1}$ are complicated nonlinear functions of the parameters of the true data generating process μ_c, μ_d, ρ, ϕ_x,ϕ_c, and the elements of $\mathbf{\Omega}$.

The second variant of limited information presumes that agents estimate separate univariate Wold representations for Δc_t, and Δd_t. Given the true data generating process, these can be written as a pair of $ARMA(1,1)$ processes:[3]

$$\Delta c_{t+1} = \mu_c (1-\rho) + \rho \Delta c_t + v^A_{c,t+1} - b_c v^A_{c,t} \quad (9)$$

$$\Delta d_{t+1} = \mu_d (1-\rho) + \rho \Delta d_t + v^A_{d,t+1} - b_d v^A_{d,t}. \quad (10)$$

As for the $VARMA$ representation, the $ARMA$ parameters are functions of the primitive parameters of the dynamic system (4)-(7) and the innovations $v^A_{c,t+1}$ and $v^A_{d,t+1}$, which are in general correlated, are composites of the underlying innovations in (4)-(6).

Specifications in which agents estimate the $VARMA$ representation are an intermediate case between two extremes: the full information specification in which agents directly observe $x_{c,t}$ and $x_{d,t}$, and limited information specifications in which agents estimate univariate $ARMA$ processes for consumption and dividend growth. We refer to the estimation of (8) as *system signal extraction*, or simply *system limited information*, and estimation of (9)-(10) separately as *univariate signal extraction*, or simply *univariate limited information*. Both cases of limited information are of interest, for several reasons.

First, the same information is employed in the system signal extraction case as in univariate signal extraction, but the information is used differently. As such, comparing these cases allows us to study how the way information is processed affects equilibrium asset returns. Second, both forms of limited information are potentially plausible descriptions of investor behavior. Single equation methods are generally more robust to model misspecification than are system estimation methods and consumption is often thought to be less well measured than are asset market data. Therefore in practice investors may trade off potential efficiency gains for robustness and measurement concerns and estimate the single equation representations. Third, the univariate signal extraction case is especially useful for building intuition about how limited information can affect asset returns. For this reason, we often discuss this case in detail before presenting results from the system signal extraction problem.

In either case, the nature of the signal extraction problem can be made explicit by noting that the Wold representations above can be written as "innovations representations," familiar from Kalman filter derivations. Let the (2×1) vector $\widehat{\mathbf{x}}^V_t = \left(\widehat{x}^V_{c,t}, \widehat{x}^V_{d,t}\right)'$ denote the optimal linear forecasts of $x_{c,t}$ and $x_{d,t}$ based on the history of both Δc_{t+1} and Δd_{t+1}. The

[3]Anderson, Hansen and Sargent (1998) study risk premia for a claim to aggregate consumption in a continuous time, robust-control asset pricing model in which consumption growth follows an $ARMA(1,1)$, in effect giving the agent the same information structure for consumption growth as in (9). 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.

$VARMA(1,1)$ process above may be recast in terms of a vector innovations representation in which Δc_t and Δd_t are functions of $\widehat{\mathbf{x}}_t^V$ (see the Appendix). Analogously, let the scalar variable $\widehat{x}_{c,t}^A$ denote the optimal linear forecast of $x_{c,t}$ based on the history of Δc_{t+1}, and let the scalar variable $\widehat{x}_{d,t}^A$ denote the optimal linear forecast of $x_{d,t}$ based on the history of Δd_{t+1}. The pair of $ARMA(1,1)$ processes above may be recast in terms of univariate innovations representations for Δc_t and Δd_t that will be functions of $\widehat{x}_{c,t}^A$ and $\widehat{x}_{d,t}^A$, respectively. As the Appendix shows, the optimal forecasts $\widehat{\mathbf{x}}_t^V$, $\widehat{x}_{c,t}^A$ and $\widehat{x}_{d,t}^A$ are functions of the observable VARMA and ARMA parameters and innovations. Thus, in either form of limited information, observations on Δc_{t+1} and Δd_{t+1} provide noisy signals of the latent variables $x_{c,t}$ and $x_{d,t}$.

3.1.3 Policy Function Solution and State Variables

For the full information specification, $x_{c,t}$ is observable and summarizes the information upon which conditional expectations are based. Since $x_{d,t} = \phi_x x_{c,t}$ it does not constitute an additional state variable. 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 in (4)-(6). This delivers a policy function for the price-consumption and price-dividend ratios as a function of a single state variable $x_{c,t}$. In the limited information specifications, equilibrium price-consumption and price-dividend ratios are calculated assuming market participants observe only the composite shock processes given in either (8) or (9) and (10), even though the data are actually generated by the dynamic system (4)-(6) with distinct short- and long-run components. For the system signal extraction case, the policy functions for both the price-consumption and price-dividend ratios are a function of the two-dimensional state vector $\widehat{\mathbf{x}}_t^V$, while in the univariate signal extraction case the policy function for the price-consumption ratio is a function of the single state variable $\widehat{x}_{c,t}^A$, and 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 in (4)-(7), and use solutions to the policy functions to generate equilibrium paths for asset prices.[4] The process is iterated

[4]A minor complication is that the policy functions for the limited information specifications are a function of the current innovation in the composite processes that appear in (9) and (10), whereas the actual innovations are generated from (4)-(6). However, the moving average representations are invertible, and their innovations can be recovered from the sums $\sum_i \mathbf{b}^i \left(\Delta \mathbf{y}_{t-i} - \mathbf{F}\Delta \mathbf{y}_{t-i-1} - \boldsymbol{\mu}\right)$ in the system signal ex-

forward to obtain simulated histories for asset returns. The Appendix explains how we solve for these functional equations numerically on a grid of values for the state variables.

3.2 How Does The Information Structure Affect Equilibrium Outcomes?

The information structure affects equilibrium asset prices because it determines the set of state variables upon which expectations are based. This mechanism can be illustrated by comparing how the respective state variables under full and limited information specifications react to primitive shocks. Figure 1 compares the full information and univariate signal extraction cases by plotting impulse responses to primitive shocks (in percent deviations from steady state) of $x_{c,t}$, as compared to $\widehat{x}^{A}_{c,t}$, $x_{d,t}$, as compared to $\widehat{x}^{A}_{d,t}$, and the dividend surprise $(\Delta d_t - E_{t-1}\Delta d_t)$ under full and limited information. The first row 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, $\varepsilon_{xc,t}$. Figure 2 compares the full information and system signal extraction cases by plotting impulse response functions of $x_{c,t}$, as compared to $\widehat{x}^{V}_{c,t}$, $x_{d,t}$, as compared to $\widehat{x}^{V}_{d,t}$, and the dividend surprise $(\Delta d_t - E_{t-1}\Delta d_t)$ under full and system limited information. In the figures, we denote all variables under full information without hats, and variables under limited information with hats.

The results in Figures 1 and 2 are based on the following calibration of parameters set at monthly frequency: $\mu_d = \mu_d = 0.0015$, $\rho = 0.979$, $\sigma = 0.0078$, $\sigma_{xc} = 0.044$. These parameter values are the same as those in the benchmark specification of Bansal and Yaron (2004), making our results comparable to those in the existing literature on long-run risk. 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. We deviate slightly from the calibration in Bansal and Yaron (2004) by setting σ_d equal to 6 rather than 4.5, in order to better match the correlation between consumption and dividends observed in the data.[5] The loadings ϕ_x and ϕ_c are set to 1 and 6, respectively, a calibration that we show below generates significant differences in risk premia between full

traction case, and from $\sum_i b^i_c (\Delta c_{t-i} - \rho\Delta c_{t-i-1} - \mu_c)$ and $\sum_i b^i_d (\Delta d_{t-i} - \rho\Delta d_{t-i-1} - \mu_d)$, respectively, in the univariate signal extraction case.

[5] Dividend growth is more volatile and less persistent than aggregate consumption growth (Cochrane (1994)). Nevertheless, the results reported in this section are not sensitive to the precise value of σ_d.

and limited information. The qualitative results discussed in this section are not influenced by the precise values of ϕ_x and ϕ_c.

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 under both full and univariate limited information (row 1, column 3). Under 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 in response to an innovation in $\varepsilon_{c,t}$, so the impulse responses of $x_{c,t}$ and $x_{d,t}$ are zero. By contrast, under limited information, agents cannot directly observe the source of the shock and do not know if it is persistent or transitory. The solution to the optimal filtering problem therefore implies that agents 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, but the former rises by much more (note the scales). In response to a transitory shock, agents with limited information 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 with limited information 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, $\widehat{x}^A_{c,t}$ responds sluggishly to the shock and $\widehat{x}^A_{d,t}$ barely responds at all. The error between both $x_{c,t}$ and $\widehat{x}^A_{c,t}$ and $x_{d,t}$ and $\widehat{x}^A_{d,t}$ dies out slowly over time. In response to a persistent shock, agents with limited information 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 under univariate limited information.

How do these results change under system signal extraction? As in the univariate signal extraction model, under system limited information agents cannot disentangle the i.i.d. shock from the persistent shock. As a consequence, agents revise upward their expectation of future consumption growth in response to an $\varepsilon_{c,t}$ shock even though in reality the shock has no persistence (row 1, Figure 2). But there is an important difference from the univariate signal extraction case: here agents also significantly revise upward their expectation of future dividend growth. Thus, in response to a one standard deviation shock to $\varepsilon_{c,t}$, $\widehat{x}^V_{d,t}$ responds by much greater magnitude than does $\widehat{x}^A_{d,t}$. This occurs because, unlike $\widehat{x}^A_{d,t}$ and $\widehat{x}^A_{c,t}$, $\widehat{x}^V_{d,t}$ and $\widehat{x}^V_{c,t}$ are perfectly correlated.

Now consider the responses to a one standard deviation increase in the idiosyncratic dividend shock, $\varepsilon_{d,t}$. Under full information this shock is correctly perceived to be both transitory and idiosyncratic; it therefore has no effect on either $x_{c,t}$ or $x_{d,t}$. By contrast, under system limited information, an increase in $\varepsilon_{d,t}$ causes a *decline* in both $\widehat{x}^V_{c,t}$ and $\widehat{x}^V_{d,t}$. This occurs because, under limited information, a shock to $\varepsilon_{d,t}$ cannot be distinguished from a shock to $\varepsilon_{c,t}$. Since investors observe an increase in dividend growth accompanied by no change in consumption, the solution to the optimal system filtering problem assigns some weight to the possibility that there have been exactly offsetting movements in $x_{c,t-1}$ and $\varepsilon_{c,t}$. This has a persistent negative affect on $\widehat{x}^V_{c,t}$ (and therefore on $\widehat{x}^V_{d,t}$), displayed in row 2 of Figure 2.

Finally, the third row of Figure 2 shows the responses to an innovation in the persistent component of consumption, $\varepsilon_{cx,t}$. As above, a one standard deviation increase in this shock leads to a jump upward in $x_{c,t}$ and $x_{d,t}$, while $\widehat{x}^V_{c,t}$ and $\widehat{x}^V_{d,t}$ respond only gradually. In the next section, we return to these responses as a way to build intuition for why risk premia differ depending on the information structure.

4 Theoretical Results

4.1 Long-Run Versus Short-Run Consumption Risk Exposure

To understand how the information structure affects equilibrium asset prices, it is instructive to begin by comparing economies with different aggregate dividend processes. Specifically, we study how risk premia differ when the relative exposure of dividend growth to long-run versus short-run consumption risk differs. Comparisons made by varying the loadings ϕ_x and ϕ_c should be thought of as comparisons among separate economies with different aggregate dividend processes, rather than comparisons among multiple risky assets in a single economy.

We begin by investigating 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, and reporting the average statistics across the 1000 simulations.[6] With the exception of the parameters ϕ_c and ϕ_x (where results for a range of values are presented), Table 1 presents results using the parameter configuration for cash flows discussed above. In addition, the preference parameters are set as follows: $\delta = 0.998985, \Psi = 1.5$, $\gamma = 10$, as in the benchmark calibration of Bansal and Yaron (2004). We denote the log return on the dividend claim $r_{t+1} = \ln(R_{t+1})$ and the log return on the risk-free rate $r_{f,t+1} \equiv \ln\left(R^f_{t+1}\right)$. Table 1 presents statistics for full information (FI), system limited information (LI-V), and univariate limited information (LI-A). We discuss comparisons between full information and each form of limited information in separate subsections next.

4.1.1 Univariate Limited Information v.s. Full Information

The results in Table 1 show that large differences in risk premia are possible between full and limited information. Consider first the results for full information, given under the column labeled "FI." In this case, 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 1), have low risk premia and high price-dividend ratios, whereas economies comprised of assets with high ϕ_x and low ϕ_c (e.g., row 5 of Table 1), have high risk premia and low price-dividend ratios. 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 2.2 to 6, the log risk premium $E(r_i - r_f)$ increases by just one and a quarter percent, from 5.20% to 6.63% per annum.

The results under univariate limited information are much the opposite. The columns labeled "LI-A" show results for the univariate signal extraction case, or ARMA filtering. 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 1), have high risk

[6] 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 data.

premia and low price-dividend ratios. Under this parameterization, the log risk premium $E(r_i - r_f)$ is almost 8 percent per annum under limited information, while it is only 2.45 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 1), the log risk premium under limited information is only 1.26% per annum whereas it is 5.2% per annum under full information.[7] Last, notice that, unlike full information, 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 2.2 to 6, the log risk premium increases by over 7 percentage points from 1.26% to 8.42% per annum. On the other hand, fixing ϕ_c and varying ϕ_x generates little variation in risk premia under limited information.

These findings are illustrated graphically in Figure 3, which plots annualized price-dividend ratios 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 fixed the 15-month variance of dividend growth that is attributable to the consumption innovations. The left-most panel plots this ratio under limited information at the steady state value of $\widehat{x}^A_{c,t}$, along with plus and minus two standard deviations around steady state in $\widehat{x}^A_{c,t}$ (holding fixed $\widehat{x}^A_{d,t}$ at its steady-state level). The middle panel plots the price-dividend ratio under limited information at the steady state value of $\widehat{x}^A_{d,t}$, along with plus and minus two standard deviations around steady state in $\widehat{x}^A_{d,t}$ (holding fixed $\widehat{x}^A_{c,t}$ 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 $x_{c,t}$.

The plots in Figure 3 are upward sloping under limited information 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 under limited information.

These results can be understood by noting that the risk premium on any asset in this economy is primarily determined by the covariance between the pricing kernel M_t and revisions in expectations (news) about future cash flow growth.[8] As such, cash flow shocks

[7] The table reports values for ϕ_c as low as 2.2. Smaller values for ϕ_c are ruled out in the limited information calibration studied here 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.

[8] Revisions in expected future returns are relatively unimportant because we have not introduced mecha-

have two offsetting effects on the equity premium in full information as compared to limited information. First, when a positive innovation ε_{xc} to the persistent component of consumption growth occurs, investors with limited information assign some weight to the possibility that the shock is transitory (coming from ε_c or ε_d). As a consequence, investors with limited information revise upward their expectation of future consumption and dividend growth by less than they would under full information. This generates a larger (in absolute value) negative correlation between M_t and cash flow news under full information than under limited information. Second, when a positive innovation ε_c to the short-run risk component occurs, investors with limited information assign some weight to the possibility that the shock is persistent (coming from the long-run risk component). As a consequence, investors with limited information revise upward their expectation of future consumption and dividend growth more than they would under full information. This generates a larger (in absolute value) negative correlation between M_t and cash flow news under limited information than under full information.

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 limited information case is quite small. On the other hand, when when ϕ_x is small and ϕ_c relatively large, the second effect dominates the first. In this case, the risk premium in the limited information case can be substantial while the premium in the full information case is quite small. Notice that, when ϕ_c is small and the long-run risk component has small variance, the ARMA dividend shock $v^A_{d,t+1}$ is largely dominated by the volatile idiosyncratic cash flow shocks $\varepsilon_{d,t+1}$ that carry no risk premium. Thus, under univariate limited information, sufficiently high exposure to short-run risk is required to generate large a large risk premium.

This intuition can be illustrated by examining impulse response functions. Figure 4 plots impulse responses of the stochastic discount factor and return to the dividend claim. The SDF under full information is denoted M_t, and under limited information, $\widehat{M_t}$. Similarly, we denote the log return to the dividend claim under full information, $r_{m,t}$, and the same return under limited information, $\widehat{r}_{m,t}$. (Recall that both the stochastic discount factor and the market return depend on the perceived data generating process for Δc_t and Δd_t, so these will differ across full and limited information.) The responses are based on the same

nisms such as changing consumption and dividend volatility for generating time-varying risk premia on the asset.

calibration used to produce the impulse responses of the state variables in Figures 2 and 3 ($\phi_x = 1$, $\phi_c = 6$), which delivers a high equity premium under limited information (7.73%) but a low equity premium under full information (2.45%).

The first row of Figure 4 shows why an innovation in the i.i.d. consumption shock $\varepsilon_{c,t}$ leads to higher risk premia under limited information than under full information. Recall that, under limited information, investors respond to such a shock by revising upward their expectation of future consumption growth substantially, while an investor with full information makes no such revision in expected future consumption growth (Figure 1). As shown in Figure 4 (row 1), this generates a larger decline in the SDF in response to an i.i.d. consumption shock under limited information than under full information, and hence a larger (negative) correlation with the return. By contrast, an innovation in the persistent component of consumption, $\varepsilon_{cx,t}$, leads to higher risk premia under full information than under limited information. This occurs because investors with limited information revise their expectations of future consumption growth by less than they would under full information (row 3, Figure 1). As a consequence, there is both a much larger decline in the SDF and a much larger increase in the return under full information than under limited information in response to a persistent consumption shock. Finally, the middle row of Figure 4 shows that a purely idiosyncratic shock to dividend growth, $\varepsilon_{d,t}$, has a negligible impact on the stochastic discount factor in either full or limited information, and so generates a negligible risk premium.

Of course, the total risk premium is influenced by all three shocks. The reason that the total risk premium is higher under limited information than under full information in this calibration is that the effect of persistent shocks is dominated by those of the i.i.d. shocks, due to the much smaller loading on the persistent component than on the i.i.d. component in the dividend process ($\phi_x = 1$, $\phi_c = 6$).

We close this subsection by briefly making one observation about the cash flow betas studied in Bansal et al. (2006). Bansal et. al. point out that regressions of dividend growth on 4 and 8 quarter trailing moving averages of consumption growth, where the slope coefficient in this regression is called the 'cash flow beta,' show that value stocks have higher cash flow betas than growth stocks.[9] It is clear that heterogeneity in ϕ_x, governing exposure to long-run consumption risk, can generate heterogeneity in cash flow betas with respect

[9] One caveat with this observation is that the cash flow betas are measured with considerable error, and therefore are not statistically distinguishable from one another.

to moving averages of consumption growth over longer horizons. A brief section in the Appendix shows that—when consumption and dividend data are time-aggregated, as in the historical data—heterogeneity in ϕ_c, governing exposure to i.i.d. consumption risk, can also generate heterogeneity in cash flow betas with respect to moving averages of consumption growth over 4 or 8 quarter horizons.

4.1.2 System Limited Information v.s. Full Information

The results in columns labeled "LI-V" of Table 1 show that limited information based on system signal extraction is a hybrid of full information and limited information based on univariate signal extraction. Like the univariate limited information case, the limited information specifications of this section deliver higher risk premia than under full information whenever the asset's exposure to long-run consumption risk is low but its exposure to short-run consumption risk is high. But the system limited information specifications also display a likeness to the full information benchmark: they deliver higher risk premia when exposure to long-run consumption risk is high and exposure to short-run risk is low. The combined result is that risk premia under system limited information are now higher than under full information for every combination of ϕ_x and ϕ_c. Under the calibration of Table 1, the magnitude of this difference is about 0.7 percent per annum. It should be noted, however, that these differences in risk premia can be made larger if risk aversion is raised. An example below using zero-coupon equity illustrates.

As before, the intuition for these results can be developed graphically by examining impulse response functions. We saw in Figure 2 that, under system limited information, investors significantly revise upward their expectation of both future consumption growth and future dividend growth in response to an i.i.d. consumption shock $\varepsilon_{c,t}$. (Such a shock generates no revision in expectations under full information, since it is correctly perceived to have no persistence.) As shown in Figure 5, this causes the SDF to decrease by more and the return to the dividend claim to increase by more under system limited information than under full information in response to an $\varepsilon_{c,t}$ shock (row 1). This generates a greater negative correlation between the pricing kernel and returns and hence greater risk premia under limited information than under full information.

The middle row of Figure 5 shows the responses of the SDF and returns to a one standard deviation increase in the idiosyncratic dividend shock, $\varepsilon_{d,t}$. This shock leads to a positive surprise in dividend growth and therefore returns, but under full information the shock generates no risk premium because it is correctly perceived to be purely idiosyncratic and uncorrelated with the stochastic discount factor (the response of the SDF is zero). By contrast, under system limited information, an increase in $\varepsilon_{d,t}$ causes a *decline* in both $\widehat{x}_{c,t}^V$ and $\widehat{x}_{d,t}^V$ (Figure 2). The decline in $\widehat{x}_{c,t}^V$ leads to a rise in the SDF, generating a positive correlation between $\widehat{M}_t$ and returns, which by itself contributes to a negative risk premium on the asset. The total risk premium on the asset is positive, however, because the insurance-like affects of an $\varepsilon_{d,t}$ shock are more than offset by the risk-like affects of an $\varepsilon_{c,t}$ shock. These offsetting influences help explain why the magnitude of the spread in risk premia between limited and full information is often lower than in the previous subsection.

The final row of Figure 5 shows the responses to an innovation in the persistent component of consumption, $\varepsilon_{cx,t}$. Recall that a one standard deviation increase in this shock leads to an immediate rise in $x_{c,t}$ and $x_{d,t}$, while $\widehat{x}_{c,t}^V$ and $\widehat{x}_{d,t}^V$ respond only gradually (Figure 3). In Figure 5 we see that this generates a larger (negative) contemporaneous correlation between the stochastic discount factor and the return under full information than under system limited information (row 3), contributing to a larger risk premium under full information. The total risk premium shown in Table 1 is still higher under limited information because, under all the calibrations shown in the table, this effect is dominated by that of the i.i.d. consumption shock.

Finally, we note that the volatility of the risk-free rate is low in all three cases: the annualized standard deviation of the risk-free rate is 1.18% in full information, 0.74% under univariate limited information, and 0.75% under system limited information.

4.2 Implications For Equity Duration

4.2.1 Zero-Coupon Equity

To study the implications of the information structure for the endogenous relation between risk premia and cash flow duration, we investigate the properties of zero-coupon equity. The idea here is that an equity claim can be represented as a portfolio of zero-coupon dividend claims with different maturities (e.g., Lettau and Wachter (2006)). Let $P_{n,t}$ denote the price of an asset at time t that pays the aggregate dividend n periods from now, and $R_{n,t}$ the one-period return on zero-coupon equity that pays the aggregate dividend in n periods:

$$R_{n,t+1} = \frac{P_{n-1,t+1}}{P_{n,t}}.$$

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

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

where the process for cash flows that generates the data D_t is given by (4)-(6). Denote $r_{n,t+1} = \ln\left(R_{n,t+1}\right)$. The appendix provides detailed information on how the recursion above is solved numerically. Since the aggregate market is the claim to all future dividends, the market price-dividend ratio $P_t^D/D_t = \sum_{n=1}^{\infty} P_{n,t}/D_t$. Plotting $E\left(r_{n,t+1} - r_{t+1}^f\right)$ against n produces a yield curve, or term structure, of zero-coupon dividend claims.

4.2.2 Univariate Limited Information v.s. Full Information

In this section, we compare the term structure of zero-coupon equity under univariate limited information with that of the full information benchmark. To do so, we consider a slightly altered set of parameter values. To generate a large spread in risk premia between short- and long-duration assets requires some combination of higher risk aversion and/or greater volatility in the long-run risk component than what has been considered so far. We modify the previous parameter values as follows: $\gamma = 16.5$, $\Psi = 1.3$, $\sigma_{xc} = 0.10$, $\delta = 0.997$, $\rho = 0.983$; $\sigma = 0.0057$, $\sigma_d = 5.9$.[10] Figure 6 plots summary statistics for log excess returns $r_{n,t+1} - r_{f,t+1}$ as a function of maturity, n, under this parameter configuration. The analogous plots for the parameter configuration studied in Table 1 are qualitatively the same, but the spread in risk-premia between long- and short-duration equity is lower. The aggregate dividend claim is assumed to follow the process (4)-(6) with $\phi_x = 1$ and $\phi_c = 3$. The plots reveal how the information structure impacts the endogenous relation between cash flow duration and risk premia.

Figure 6 shows that, under limited information, the annualized log risk premium declines with maturity (top panel). The log risk premium is 14% per annum for equity that pays a

[10] Without any further adjustments, the increase in σ_{xc} and ρ makes annualized consumption growth too volatile, so we reduce σ from 0.0078 to 0.0057. This calibration delivers a standard deviation for annualized consumption growth of 3.6%, a value within one standard deviation of the point estimate reported in Bansal and Yaron (2004). The parameters δ and σ_d are also slightly modified in order to maintain plausible implications for the risk-free rate and the relative volatility of dividend growth.

dividend one month from now and 5% per annum for equity that pays a dividend 15 years from now (top panel). Thus, under limited information, short-duration assets, those with more weight in low-maturity equity, will endogenously have higher expected returns and lower price-dividend ratios than long-duration assets with more weight in distant-maturity equity. A downward sloping equity yield curve is needed for the model to match empirical evidence that long-horizon equity is less risky than short-horizon equity.

By contrast, under full information, the annualized log risk premium increases with maturity. The log risk premium is 1.8% per annum for equity that pays a dividend one month from now and 5.5% per annum for equity that pays a dividend 15 years from now.

The key to the downward sloping zero-coupon equity curve under limited information, displayed in the top panel of Figure 6, is that the process for dividend growth under univariate signal extraction appears close to i.i.d. (the estimated moving average and autoregressive roots in (9) and (10) are close to canceling). Thus, shocks are perceived only to affect dividend growth and returns in the near term, implying that only assets that pay a dividend in the near future command high risk premia. By contrast, under full information, when agents can directly observe $x_{c,t}$, it is understood that shocks can have a large, long-term affect 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 cash flow duration and risk premia goes the wrong way.

The middle panel of Figure 6 shows that in both limited and full information, volatility increases with the horizon. But the bottom panels show that the Sharpe ratios decrease with the horizon under limited information whereas they rise with the horizon under full information. This suggests that the limited information specification is better able to explain the empirically higher Sharpe ratios of short-duration value stocks as compared to long-horizon growth stocks.

Figure 7 shows that, under limited information, the shortest-duration equity have high CAPM alphas (as high as 8% for equity that pays a dividend in one month), whereas the longest-maturity equity have smaller (in absolute value) negative alphas (-2% for equity that pays a dividend 15 years from now). This feature of limited information is consistent with the data (see Table 2, discussed below) and with the prior findings of Fama and French (1992). The bottom panel also shows that, under limited information, long-duration equity—despite its having lower expected excess returns than short-duration equity—has slightly higher CAPM betas, as in the data. By contrast, under full information, there is much less variation in the alphas with maturity and the variation goes the wrong way: alphas of short-duration assets are lower than those of long-duration assets.

The plots just discussed are based on particular values of the loadings ϕ_c and ϕ_x for the market portfolio. Results (not reported) indicate that these findings hold for a wide range of parameters in both the cash flow process and the utility function. Regardless of the values of ϕ_c and ϕ_x, the term structure of equity is always downward sloping under univariate limited information, and always upward sloping under full information. Changing the relative loadings ϕ_c and ϕ_x merely changes the slope of the term structure, it does not change the sign of the slope. This is true as long as parameter values are set so that greater exposure to $x_{c,t}$ makes the market portfolio riskier rather than providing insurance. In a long-run "insurance" model, the full information term structure slopes down, but overall risk premia are very low or even negative. These latter results will be important when we discuss the system signal extraction case below.

The results above may be related to similar zero-coupon equity plots in the literature. Hansen et al. (2005) present zero-coupon equity plots for price-dividend ratios $P_{n,t}/D_t$ rather than mean excess returns, as in Figure 6. Since high price-dividend ratios correspond to low mean excess returns, the plots presented in Hansen et al. (2005) are mirror-images of those above. Their plots are based on the same Epstein-Zin-Weil model of preferences used here, but the results are formed from historical data and somewhat different parameter values. Hansen et. al. report price-dividend term zero-coupon equity structures for value and growth firms separately, whereas we plot the zero-coupon-equity curve for aggregate dividends. (The dividend payments of value and growth firms are modeled below as time-varying shares in a sequence of aggregate dividend claims, $\{D_t\}_{t=0}^{\infty}$, with different maturities.) In this sense, the results in this section are not directly comparable to those in Hansen et al. (2005). But it is notable that Hansen et. al. find that the price-dividend decomposition for the growth portfolio eventually exceeds those of the value portfolio, as required by the data, only at sufficiently high levels of risk aversion. This finding is echoed in the results reported here: it is only with sufficiently high risk aversion and/or sufficiently volatile innovations to the long-run expected growth rate of consumption that we find a significant spread in average returns between short-duration value stocks and long-duration growth stocks.

Portfolios of Firms It is possible to study the quantitative aspects of the link between duration and risk premia presented above by modeling firms explicitly. Here we do so by forming portfolios of individual firms that differ only in the timing of their cash flows. Long-duration growth firms are modeled as equity with relatively more weight placed on long-horizon dividend claims, while short-duration value firms are modeled as equity with relatively more weight placed on short-horizon dividend claims. This methodology for describing the cash flows of individual securities was first employed in a continuous time setting by Menzly, Santos and Veronesi (2004), Santos and Veronesi (2004), and Santos and Veronesi (2005), and in a discrete setting by Lynch (2003) and Lettau and Wachter (2006). Here we follow the discrete time methodology described in Lettau and Wachter (2006). We outline only the main aspects of this approach and refer the reader to that article for further detail.

This admittedly stylized model of dividend payments is not meant to be fully realistic. For example, a simplifying assumption is that a firm's dividends depend only on the aggregate dividend shock, whereas in reality they contain a firm-specific component. What the methodology does do, however, is to highlight an important source of observed heterogeneity in individual cash flows, driven by differences in duration.[11] By proceeding in this way, we are able to avoid additional complexity and maintain focus on the endogenous link between risk premia and cash flow duration in models with long-run consumption risk.

Consider a sequence of $i = 1, ..., N$ firms. (Hereafter we refer to these simply as 'firms' for brevity, even though they are better thought of as portfolios of firms at the same stage in their life cycle.) The ith firm pays a share, $s_{i,t+1}$, of the aggregate dividend D_{t+1} at time $t+1$. The share process is deterministic, with $\underline{s}$ the lowest share of a firm in the economy. Firms experience a life-cycle in which this share grows deterministically at a rate g_s until reaching a peak $s_{i,N/2+1} = (1 + g_s)^{N/2} \underline{s}$ when it shrinks deterministically at rate g_s until reaching $s_{i,N+1} = \underline{s}$. The cycle then repeats. Thus, firms are identical except that their life-cycles are out-of-phase, i.e.., firm 1 starts at $\underline{s}$, firm 2 at $(1 + g_s) \underline{s}$, and so on. The parameter g_s is set to 1.67% per month, or 20% per year, as in Lettau and Wachter (2006). Shares are such that $s_{i,t} \geq 0$ and $\sum_{i=1}^{N} s_{i,t} = 1$ for all t.

Since each firm pays a dividend $s_{i,t+1}D_{t+1}$, $s_{i,t+2}D_{t+2}, ...$, no arbitrage implies that the

[11] Da (2005) empirically measures equity duration in the manner modeled here, namely as the deviation of an asset's current share in aggregate dividends from its steady-state value. Consistent with previous findings, (which were based on somewhat different methodologies for measuring equity duration), he finds that value stocks have much shorter cash flow duration than growth stocks.

ex-dividend price of firm i at time t is given by

$$P_{i,t} = \sum_{n=1}^{\infty} s_{i,t+n} P_{n,t}.$$

When $s_{i,t+1}$ is low, dividend payments are low today but will be high in the future when n is large; these are long-duration assets with greater weight placed on distant-maturity dividend claims. When $s_{i,t+1}$ is high, dividend payments are high today but will be low in the future; these are short-duration assets with greater weight placed on short-maturity dividend claims. From the downward sloping term structure plots presented above, we already know that, under limited information, firms with high price-dividend ratios and low risk premia will be those that pay a small share of the aggregate dividend today, but a greater share farther into the future. Such "growth" assets will endogenously have both high price-dividend ratios (accompanied by relatively low risk premia) and long duration in their cash flows. Conversely, firms with low price-dividend ratios and high risk premia will be those that pay a larger share of the aggregate dividend today, but a small share farther into the future. Such "value" assets will endogenously have both low price-dividend ratios (accompanied by relatively high risk premia) and short duration in their cash flows.

To consider the quantitative properties of models with limited information, we sort firms into portfolios on the basis of price-dividend ratio. Portfolio returns are created by simulating a time-series for aggregate dividends and prices and, using the share process described above, forming 10 equally-weighted portfolios of the N firms by sorting firms into deciles based on their price-dividend ratios.[12] The portfolios are rebalanced every simulation year. This procedure creates portfolios of firms that display heterogeneity not only in their price-dividend ratios, but also (endogenously) in the duration of their cash flows. Table 2 reports the statistical properties of these portfolios. For comparison, Table 2 also provides updated evidence on the value premium in U.S. data. The table shows summary statistics from U.S. data for portfolios of firms sorted into deciles on the basis of book-to-market ratio, with decile 1 containing firms in the lowest 10 percent according book-to-market ratio, and decile 10 containing firms in the highest 10 percent according to book-to-market ratio. The monthly data are from the Center for Research in Securities Prices, and span the period 1947-2004.[13]

Statistics are presented for expected excess returns, Sharpe ratios, and CAPM regressions,

[12] We set the number of firms to be 1020, implying a 1020 month, or 85 year life-cycle for a firm.

[13] We thank Kenneth French for compiling the portfolios from these data and making them available on his web page.

based on a single long simulation of the true data generating process in (4)-(6). For the model-based results, we refer to the portfolio in the highest price-dividend decile as the growth portfolio, denoted G in the table, and the portfolio in the lowest price-dividend decile as the value portfolio, denoted V in the table. We present these statistics only for the limited information specifications. As the upward-sloping zero-coupon equity term structure demonstrates, specifications with full information generate a value premium by making long-duration assets more risky than short-duration assets. Since the aim is to generate a value premium that implies long-duration assets are *less* risky than short-duration assets, we do not pursue that avenue here.

Although this simplified model of firm cash flows omits some important aspects of reality, it is notable that this simple framework comes close to matching the magnitudes of financial statistics observed in the data. For example, Table 2 shows that the limited information specification is capable of generating a sizeable value premium. The mean excess return on the extreme growth portfolio is 6.39%, while that of the extreme value portfolio is 11.47%, leaving a spread between the two of 5.08% . These numbers are close to those reported in U.S. data, for the mean excess return in the lowest book equity-to-market capitalization quintile (B/M quintile) compared to the highest B/M quintile. The specification also predicts that Sharpe ratios rise when moving from growth to value portfolios, as in the data. In fact, the calibration here comes remarkably close to matching the data: the Sharpe ratio of the growth portfolio is 0.38, while that of the value portfolio is 0.70. In the data, the lowest B/M quintile has a Sharpe ratio of 0.38 and the highest has a Sharpe ratio of 0.64.

The limited information model produces little spread in the CAPM betas across portfolios. The value portfolio has a beta that is slightly less than that of the growth portfolio, a pattern found the classic results of Fama and French (1992) and in the updated data reported in Table 2. The pattern of alphas is also consistent with the data. In Table 2, model-based alphas rise from -1.2% for the extreme growth portfolio to 4.22% for the extreme value portfolio. By comparison, in the post-war data the lowest B/M quintile has an alpha of -1.68% and the highest has an alpha of 4.19%.

Finally, Table 3 shows the results of adding the HML (high-minus-low) factor of Fama and French (1993) as an additional regressor in CAPM time-series regressions of the excess portfolio returns onto the excess market return. HML is constructed as the return on a portfolio short in the extreme growth decile and long in the extreme value decile. Consistent with the empirical findings of Fama and French (1993), the model implies that adding HML

as an additional factor significantly reduces the magnitude of the CAPM alphas in all decile portfolios.

4.2.3 System Limited Information v.s. Full Information

We now study the implications of the system signal extraction problem for the cash flow duration properties of value and growth assets. The essential difference from univariate signal extraction is that the off-diagonal elements on the **b** matrix in (8) are non zero, implying that information on consumption is used directly in estimating the stochastic process for dividend growth.

Figure 8 plots the annualized log risk premium, $E\left(r_{n,t+1} - r_{f,t+1}\right)$, as a function of maturity, n, under the same parameter configuration used to produce the results in Figure 6 for univariate limited information. Figure 8 shows that, while univariate limited information generates a downward sloping term structure, both system limited information and full information generate upward sloping term structures under this parameterization. This occurs because the dividend growth process is perceived to be much more persistent under system limited information than under univariate limited information. Accordingly, assets that pay a dividend in the far future command higher risk premia than those that pay a dividend in the near future.

Although the slope of the term structure under system limited information is virtually identical to that under full information, the level of the term structure is not. The limited information curve in Figure 8 is higher than the full information curve at all maturities, by a little over 2 percentage points under this calibration. This reflects the finding, discussed above, that risk premia are higher under system limited information than under full information, for a range of parameter values.

To understand both the slope and the level of the term structure, it is instructive to consider the role played by key model parameter values. Two parameters are especially important for governing the slope of the term structure: the exposure ϕ_x of dividend growth to the persistent component of consumption growth, and the intertemporal elasticity of substitution, Ψ. The lower is ϕ_x, the less persistent is dividend growth and the less upward sloping the curve. Indeed, a sufficiently small value for ϕ_x can flip the slope of the zero-coupon equity curve to downward sloping. This occurs for reasons already discussed: when the dividend growth process has little persistence, only shocks to dividend growth in the near term generate significant revisions in expected future consumption and cash flow growth, and hence command a significant risk premium; short-duration assets are riskier than long-duration assets.

The intertemporal elasticity of substitution affects the slope of the term structure by affecting expected future returns, rather than expected future dividend growth. The lower is Ψ, the more expected returns increase in response to any given increase in expected consumption growth. Thus, a positive innovation in expected consumption growth does two things. First, it leads to an increase in expected future returns, 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 imply a positive contemporaneous correlation between the pricing kernel and returns, making the overall risk premium on the asset low or even negative. 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 IES, the lower are risk premia on long-duration assets relative to short-duration assets, and the less upward sloping the zero-coupon-equity curve.

These properties of the model suggest one way that parameter values could be changed in order to make the term structure of equity downward sloping: reduce ϕ_x and Ψ. The role of these parameters on the slope of the zero-coupon-equity term structure can be illustrated by the approximate log-linear solution of the model, similar to Campbell (2003). Let $V_t(\cdot)$ denote the conditional variance of the generic argument "$\cdot$". Define the slope of the log equity term structure (adjusted for Jensen's inequality terms) as

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

where the superscript "ex" denotes the excess return over the log risk-free rate. In the full information case it can be shown that, to a first-order approximation,

$$S = \frac{\phi_x - 1/\Psi}{1-\rho}\sigma_{xc}\sigma^2\left[\gamma\rho_{cx_c} + \kappa_c\left(\frac{\gamma - 1/\Psi}{1-\rho\kappa_c}\right)\sigma_{x_c}\right] \tag{11}$$

where $\kappa_c \equiv \frac{\overline{P^C/C}}{1+\overline{P^C/C}}$ is a positive linearization constant less than unity, and ρ_{cx_c} denotes the unconditional correlation between the i.i.d. consumption shock, $\varepsilon_{c,t}$, and the shock to long-run expected consumption growth, $\varepsilon_{xc,t}$. For the rest of this discussion, we maintain the assumption that $\gamma > 1/\Psi$. If we also assume for the moment that $\rho_{cx_c} \geq 0$, then the term in the square brackets is positive, and it is possible to generate a downward sloping term structure of equity (a negative spread, $S < 0$) by setting $\phi_x < 1/\Psi$. When $\Psi = 1$, the valuation calculations in Hansen et al. (2005) 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.

Under full information, however, this strategy presents an important difficulty. When $\phi_x < 1/\Psi$, the model becomes one of long-run *insurance* rather than long-run *risk*. That is, innovations in $x_{c,t}$ (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. This occurs because such a parameterization has the undesirable property that an increase in the long-run expected consumption growth rate leads to a decline in the market price-dividend ratio. This property is immediately evident from the approximate formula for the log price-dividend ratio of the market return under full information:

$$p_t^D - d_t = \overline{pd} + \frac{\phi_x - 1/\Psi}{1 - \rho\kappa_d} x_{c,t},$$

where $\overline{pd}$ is a constant and $\kappa_d \equiv \frac{\overline{P^D/D}}{1+\overline{P^D/D}}$ is a positive linearization constant less than one. Note that the coefficient on $x_{c,t}$ is negative whenever $\phi_x < 1/\Psi$. The system limited information model shares this problem.

In addition, in a long-run "insurance" model, the full information term structure slopes down, but the overall equity premium for the market is low or negative. This can be understood by examining the loglinear approximation of the market equity premium under full information, given by

$$E_t(r_{d,t+1}^{ex}) + .5V_t(r_{d,t+1}^{ex}) = \gamma\phi_c\sigma^2 + \kappa_d \frac{(1-\rho)}{1-\rho\kappa_d} S + \kappa_c \frac{\gamma - 1/\Psi}{1-\rho\kappa_c} \sigma^2 \left[\phi_c \rho_{cx_c} + \rho_{dx_c} \sigma_{x_c} \sigma_{x_d}\right], \quad (12)$$

where ρ_{dx_c} denotes the unconditional correlation between the idiosyncratic dividend shock, $\varepsilon_{d,t}$, and the shock to long-run expected consumption growth, $\varepsilon_{xc,t}$. Since $\kappa_d \frac{(1-\rho)}{1-\rho\kappa_d} > 0$, if the slope S of the equity term structure is negative, it is difficult to generate a sizable equity premium.

Figure 9 illustrates this point by plotting the term structure of equity for a calibration in which $\phi_x = 0.76$ (instead of $\phi_x = 1$ as in Figure 8), and in which the IES is $\Psi = 1$ (instead of $\Psi = 1.3$); hence $\phi_x < 1/\Psi$ and $S < 0$. If no other parameter values are changed, such a calibration produces a downward sloping term structure of equity, but the spread in risk premia between short- and long-horizon equity is small. One remedy is to adjust risk aversion upward and then insure that exposure ϕ_c to short-run risk is sufficiently high to help increase the market equity premium (the level of the term structure). The results in Figure 9 are displayed for $\gamma = 50$ and $\phi_c = 3.6$, with all other parameters set as in the calibration of Figure 8. As Figure 9 shows, under this calibration both the system signal extraction model of limited information and the full information model produce a downward sloping zero-coupon equity curve. However, the market risk premium is now negative under full information. By contrast, the system limited information specification produces a downward sloping zero-coupon equity curve with both a reasonable spread in risk premia and a reasonable market risk premium. Under the system version of limited information, the term structure is higher than under full information, so it is possible to make the term structure downward sloping without requiring the overall market risk premium to be implausibly low or negative.

The difficulty posed by the full information specification in generating a downward sloping term structure for equity simultaneously with a high equity premium cannot be easily remedied by freely setting the correlation ρ_{cx_c} between current consumption shocks and shocks to the long-run expected consumption growth. For example, if we restrict $\phi_x > 1/\Psi$ to avoid the implications just discussed, then (11) implies that we can obtain a downward sloping term structure by setting $\rho_{cx_c} < 0$. Unfortunately, (12) shows that this again makes the overall equity premium low or negative, since it makes both S and the third term of (12) negative.

In summary, in a long-run "insurance" model, the full information term structure slopes down, but the overall risk premium for the market is low or negative. Under the system version of limited information, the same downward slope can be achieved with a positive market risk premium of reasonable magnitude, but like the full information specification, this requires parameter values for which the price-dividend ratio responds negatively to increases in the long-run expected growth rate of the economy. Still, the results in this section reveal some important commonalities between the univariate and system signal extraction models. To match empirical evidence that long-horizon equity is less risky than short-horizon equity, simultaneously with a sizable market risk premium, an asset's exposure to long-run risk must not be too large, while its exposure to short-run risk must be sufficiently high. These results contrast with those for the full information models and the specifications emphasized in Bansal and Yaron (2004), designed to match only the empirical properties of the market portfolio. Those models have emphasized modest risk aversion accompanied by sufficiently high exposure of aggregate dividend growth to long-run consumption risk, with no exposure to short-run consumption risk.

5 Conclusion

A recent strand of asset pricing literature emphasizes the potential role of long-run consumption risk for explaining salient asset pricing phenomena. Because such long-run components are small and difficult to identify from the data, econometricians face concrete statistical hurdles in observing these components directly. Yet a maintained assumption in the existing theoretical literature is that investors can directly observe such small long-run components and can distinguish their innovations from transitory shocks to consumption and dividend growth. In this paper we have studied how equilibrium asset prices may be affected if market participants–like econometricians–must use consumption and dividend data to infer small long-run components in cash flows and consumption.

We find that the asset pricing implications of long-run risk models can be quite sensitive to the information investors have about the long-run. To illustrate the importance of the information structure, we study the cash flow duration perspective of value and growth assets. A key result of this study is that, under many parameter configurations, limited information causes market participants to demand a higher premium for engaging in risky assets than would be the case under full information. 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 under limited information but not under full information.

These findings may partly explain why there is considerable statistical uncertainty over the extent to which value and growth stocks are differentially exposed to long-run consumption risk (Hansen et al. (2005); Bansal et al. (2006)). Even in a model where long-run risk plays a central role in determining risk premia, once limited information is introduced, stocks that have high average returns need not be those that are more highly exposed to long-run consumption risk. In general, these patterns mean that the limited information specifications we explore are better able than their full information counterparts to reconcile the return properties of value and growth assets with their quite different cash flow duration properties.

In a full information world where market participants can fully discern the distinct roles of persistent and transitory cash flow shocks, long-duration assets can be made less risky than short-duration assets under parameter configurations for which the model is one of long-run insurance rather than long-run risk. As such, a downward sloping term structure for zero-coupon equity comes at the expense of a negative market risk premium. By contrast, under limited information, a downward slope in the term structure of equity can be achieved with a positive market risk premium of reasonable magnitude, as long as risk aversion and exposure to short-run consumption risk is sufficiently high.

Much of the existing long-run risk literature focuses on explaining the behavior of the aggregate market return and/or the return properties of value and growth stocks, but little attention has been given to how equilibrium returns are related to equity duration. Here we find that, in order to explain the empirical finding that long-horizon equity is less risky than short-horizon equity–in addition to explaining the aggregate market return–the properties of the cash flow model, the information structure, and the values of primitive preference parameters must be quite different from those emphasized in the existing long-run risk literature. In particular, information must be limited, risk aversion must be higher and the IES lower than previously considered, and exposure of assets' cash flows to long-run consumption risk must be sufficiently low relative to short-run consumption risk.

There are at least two ways in which this research could be extended. First, in order to focus on the role of information and its relation to the cross-section of average returns, we have not incorporated additional sources of time-varying risk that may also be unobservable, such as changing volatilities of cash flows. As such, our model of the excess return on the market does not display significant predictability, implying that the volatility of the market price-dividend ratio is lower than in the data. If the volatility of cash flows changes in a way that is not observed, the agent must solve a complex nonlinear filtering problem. We are currently studying this problem. Second, the specification of cash flows and informational assumptions pursued here is but one of many that could be fruitfully studied in future work. In addition, investors may have preferences that differ from those presumed here, and the form of these preferences may interact with informational barriers in interesting ways. For example, informational barriers may be compounded by uncertainty over the cash flow model itself, possibly leading investors to have a preference for robustness, as in the work of Anderson et al. (1998), Anderson, Hansen and Sargent (2003) and Hansen and Sargent (2006). Exploring these extensions presents a challenge for future work.

6 Appendix

6.1 Limited Information Identification, Innovations Representation

6.1.1 Identification

We assume that the general data generating process that agents with limited information would like to estimate takes the form

$$\Delta c_{t+1} = \mu_c + \underbrace{x_{c,t}}_{\text{LR risk}} + \sigma \underbrace{\varepsilon_{c,t+1}}_{\text{SR risk}} \tag{13}$$

$$\Delta d_{t+1} = \mu_d + x_{d,t} + \phi_c \sigma \varepsilon_{c,t+1} + \sigma_d \sigma \varepsilon_{d,t+1} \tag{14}$$

$$x_{c,t} = \rho x_{c,t-1} + \sigma_{xc} \sigma \varepsilon_{xc,t} \tag{15}$$

$$x_{d,t} = \rho_d x_{d,t-1} + \sigma_{xd} \sigma \varepsilon_{xd,t} \tag{16}$$

$$(\varepsilon_{c,t+1}, \varepsilon_{d,t+1}, \varepsilon_{xc,t}, \varepsilon_{xd,t}) \sim N.i.i.d\,(\mathbf{0}, \mathbf{\Omega}) \tag{17}$$

Note that the equation (14) can be re-written as

$$\Delta d_{t+1} = \mu_d + \phi_x x_t + \widetilde{\varepsilon}_{d,t+1}.$$

Market participants could in principal obtain a consistent estimate of these parameters simultaneously with estimates of $x_{c,t}$ and $x_{d,t}$, by writing the dynamic system above in state space form and applying maximum likelihood to the history of consumption and dividend data. Agents could use the Kalman filter to form an estimate of the unobservable conditional means $x_{c,t}$ and $x_{d,t}$, by sequentially updating a linear projection on the basis of consumption and dividend data observed through date t.

This system is not identified, however. The system (13)-(17) has 14 unknown parameters (including ten unknown parameters in $\mathbf{\Omega}$). Estimation of (8) identifies 11 parameters, three short of what's needed for exact identification. That is, given a sufficiently long sample of data on consumption and dividend growth, the parameters of the dynamic system (13)-(17) can be observed in certain combinations as the estimates ρ, b_{cc},..., b_{dd} and the variance-covariance matrix of $v_{c,t+1}$ and $v_{d,t+1}$, but this information is not enough to separately identify the parameters of (13)-(17). The true data generating process (4)-(6) is a special case of this system that imposes the restrictions $x_{d,t} = \phi_x x_{c,t}$, requiring $\rho_d = \rho$, $\sigma_{xd} = \phi_x \sigma_{xc}$, $\varepsilon_{xd,t} = \varepsilon_{xc,t}$, $x_0 = x_{d0} = 0$, and the shocks to (13)-(15) to be uncorrelated.

6.1.2 Innovations Representation

The $VARMA(1,1)$ process may be recast in terms of the following system innovations representation:

$$\Delta\mathbf{y}_{t+1} = \boldsymbol{\mu} + \widehat{\mathbf{x}}_t^V + \mathbf{v}_{t+1}^V \tag{18}$$

$$\widehat{\mathbf{x}}_{t+1}^V = \mathbf{F}\widehat{\mathbf{x}}_t + \mathbf{K}\mathbf{v}_{t+1}^V, \tag{19}$$

where $\widehat{\mathbf{x}}_t^V$ denotes the optimal linear forecasts of $\mathbf{x}_t \equiv (x_{c,t}, x_{d,t})'$, the conditional mean of consumption and dividend growth, respectively, based on the history of both series, i.e., $\widehat{\mathbf{x}}_t^V \equiv \widehat{E}(\mathbf{x}_t|\mathbf{z}^t)$, $\widehat{\mathbf{x}}_t^V \equiv (\widehat{x}_{c,t}^M, \widehat{x}_{d,t}^M)'$, $\mathbf{z}^t \equiv (\Delta c_t, \Delta c_{t-1}, ..., \Delta c_1, \Delta d_t, \Delta d_{t-1}, ..., \Delta d_1)'$, so $\widehat{E}(\mathbf{x}_t|\mathbf{z}^t)$ denotes the linear projection of $\mathbf{x}_t$ on $\mathbf{z}^t$ and a constant. In addition, $\Delta\mathbf{y}_{t+1} \equiv (\Delta c_{t+1}, \Delta d_{t+1})'$, $\boldsymbol{\mu} \equiv (\mu_c, \mu_d)'$, $\mathbf{v}_{t+1}^A \equiv (v_{c,t+1}^V, v_{d,t+1}^V)'$, and $\mathbf{K} = \mathbf{F} - \mathbf{b}$ with

$$\mathbf{F} \equiv \begin{bmatrix} \rho & 0 \\ 0 & \rho \end{bmatrix}, \qquad \mathbf{b} \equiv \begin{bmatrix} b_{cc} & b_{cd} \\ b_{dc} & b_{dd} \end{bmatrix}.$$

The innovations representation is often obtained using the Kalman filter. Application of the Kalman filter to the dynamic system (13)-(15) typically requires the assumption that the innovations in the state equation (here $\varepsilon_{xc,t}$ and $\varepsilon_{xd,t}$) are uncorrelated at all lags with the innovations in the observation equations (here $\varepsilon_{d,t}$ and $\varepsilon_{c,t}$). In this case the system is identified, but a signal extraction problem must still be solved to obtain estimates of the latent variables $x_{c,t+1}$ and $x_{d,t+1}$. If the steady state Kalman filter is applied to the system (13)-(15), it yields the innovations representation in above. The parameter $\mathbf{K}$ in (19) is the steady state Kalman gain matrix associated with the state space representation of the dynamic system (13)-(16).

Similarly, the $ARMA(1,1)$ processes may be recast in terms of the following pair of innovations representations:

$$\Delta c_{t+1} = \mu_c + \widehat{x}_{c,t}^A + v_{c,t+1}^A \tag{20}$$

$$\widehat{x}_{c,t+1} = \rho\widehat{x}_{c,t}^A + Kv_{c,t+1}^A \tag{21}$$

$$\Delta d_{t+1} = \mu_d + \widehat{x}_{d,t}^A + v_{d,t+1}^A \tag{22}$$

$$\widehat{x}_{d,t+1} = \rho\widehat{x}_{d,t}^A + K^d v_{d,t+1}^A, \tag{23}$$

where $K \equiv \rho - b_c$ and $K^d \equiv \rho - b_d$. Here, $\widehat{x}_{c,t}^A$ and $\widehat{x}_{d,t}^A$ denote optimal linear forecasts based on the history of consumption and dividend data separately, i.e., $\widehat{x}_{c,t}^A \equiv \widehat{E}(x_{c,t}|\mathbf{z}_c^t)$, and $\widehat{x}_{d,t}^A \equiv \widehat{E}(x_{d,t}|\mathbf{z}_d^t)$, where $\mathbf{z}_c^t \equiv (\Delta c_t, \Delta c_{t-1}, ..., \Delta c_1)'$ and $\mathbf{z}_d^t \equiv (\Delta d_t, \Delta d_{t-1}, ..., \Delta d_1)'$.

The optimal forecasts are functions of the observable VARMA and ARMA parameters and innovations:

$$\begin{aligned}
\widehat{\mathbf{x}}_t^M &= -\mathbf{F}\boldsymbol{\mu} + \mathbf{F}\Delta\mathbf{y}_t - \mathbf{b}\mathbf{v}_t^M \\
\widehat{x}_{c,t}^A &= -\rho\mu_c + \rho\Delta c_t - b_c v_{c,t}^A \\
\widehat{x}_{d,t}^A &= -\rho\mu_d + \rho\Delta d_t - b_d v_{d,t}^A.
\end{aligned}$$

6.2 Numerical Solution

We describe our numerical solution procedure for the full information specifications and the univariate signal extraction case. A description of the system signal extraction case is directly analogous and is omitted for brevity.

6.2.1 Full Information

Under Full Information, there is a single state variable, x_t. We discretize and bound its support by forming a grid of K points $\{x_1, x_2, ... \ x_K\}$ on the interval [-5V(x) +5V(x)]. We choose K to be odd so that the unconditional mean of the state x is the middle point of our grid.

We discretize also the distribution of a standardized normal random variable by forming a grid of equidistant points $\{\epsilon_1, \epsilon_2, ... \ \epsilon_I\}$ over the interval [-5 +5], imposing:

$$p_i = \frac{e^{-\epsilon_i^2/2}}{\sum_1^I e^{-\epsilon_i^2/2}}, \ i = 1, 2, ...I$$

Again, we choose I to be odd so that $\epsilon_{(I-1)/2+1} = 0$.

Rewrite the Euler equations for the price-consumption ratio as:

$$\begin{aligned}
W_c(x_k) &= \left(\sum_{i=1}^{I} \sum_{j=1}^{I} \delta^\theta e^{(1-\gamma)(\mu + x_k + \sigma\epsilon_i)} [1 + W_c(x'_{j|k})]^\theta p_i p_j \right)^{\frac{1}{\theta}} \qquad (24) \\
x'_{j|k} &= \rho x_k + \sigma\varphi_x \epsilon_j \\
k &= 1, 2, ..., K,
\end{aligned}$$

where $W_c(x_k)$ is the price-consumption ratio as a function of x in state k. The functional in (24) can be solved by noting that its right hand side is a contraction and treating $W_c(x)$ as the fixed point of $W_{c,n+1}(x) = T(W_{c,n}(x))$.

Approximate $W_{c,n}$ by a third order polynomial in x, and impose:

$$W_{c,n}(x'_{j|k}) = [1 \; x'_{j|k} \; (x'_{j|k})^2 \; (x'_{j|k})^3][\beta_{1,n} \; \beta_{2,n} \; \beta_{3,n} \; \beta_{4,n}]'$$

where the operator is initialized with an initial guess on the parameters β_0. Compute $W_{c,1}(x_k)$ for every $x_k \in \{x_1, x_2, \ldots \, x_K\}$, and stack the resulting values in the vector $\overrightarrow{W}_{c,1} \in R^K$. Using least squares the guesses are updated: $\beta_1 = (\Upsilon'\Upsilon)^{-1}\Upsilon'\overrightarrow{W}_{c,1}$, where:

$$\Upsilon = \begin{bmatrix} 1 & x_1 & (x_1)^2 & (x_1)^3 \\ 1 & x_2 & (x_2)^2 & (x_2)^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_k & (x_k)^2 & (x_k)^3 \\ & & & \end{bmatrix}$$

We repeat these steps until convergence (tolerance level = .1e-5).
Once $W_c(x) = [1 \; x \; x^2 \; x^3]\beta$ has been found, the stochastic discount factor has the following expression:

$$M_{k,i,j} = \delta^\theta e^{-\gamma(\mu + x_k + \sigma\epsilon_i)} \left(\frac{1 + W_c(\rho x_k + \sigma\varphi_x \epsilon_j)}{W_c(x_k)} \right)^{\theta - 1}$$

price-dividend ratios are found in a similar way by iterating until convergence the following recursion:

$$\begin{aligned} W_{d,n+1}(x_k) &= \sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{l=1}^{I} \delta^\theta e^{-\gamma(\mu + x_k + \sigma\epsilon_{c,i})} \left(\frac{1 + W_c(x'_{j|k})}{W_c(x_k)} \right)^{\theta-1} \times \\ & \quad \times [1 + W_{d,n}(x'_{j|k})] e^{(\mu + \phi_x x_k + \phi_c \sigma \epsilon_{c,i} + \sigma\varphi_d \epsilon_{d,l})} p_i p_j p_l \qquad (25) \\ W_{d,n}(x'_{j|k}) &= [1 \; x'_{j|k} \; (x'_{j|k})^2 \; (x'_{j|k})^3]\beta_{d,n} \end{aligned}$$

The coefficients of the polynomial expansion for the price-dividends are updated by the following OLS formula: $\beta_{d,n+1} = (\Upsilon'\Upsilon)^{-1}\Upsilon'\overrightarrow{W}_{d,n+1}$.
For $n \to \infty$, $\beta_{d,n+1} \to \beta_d = (\Upsilon'\Upsilon)^{-1}\Upsilon'\overrightarrow{W}_d$.

To solve for zero coupon equity price-dividend Ratios note the following equivalence holds:

$$\begin{aligned} W_{d,t} &= \sum_{n=1}^{\infty} W_{d,t}^{n} \\ &\text{where} \\ W_{d,t}^{0} &\equiv 1 \\ W_{d,t}^{n} &= E_t\left[e^{m_{t+1}+\Delta d_{t+1}} W_{d,t+1}^{n-1}\right],\ n=1,2,... \end{aligned} \tag{26}$$

Implement the following recursion across maturities:

$$\begin{aligned} W_d^n(x_k) &= \sum_{i=1}^{I}\sum_{j=1}^{I}\sum_{l=1}^{I} \delta^{\theta} e^{-\gamma(\mu+x_k+\sigma\epsilon_i)} \left(\frac{1+W_c(x'_{j|k})}{W_c(x_k)}\right)^{\theta-1} \times \\ &\quad \times [W_d^{n-1}(x'_{j|k})] e^{(\mu+\phi_x x_k+\phi_c\sigma\epsilon_i+\sigma\varphi_d\epsilon_l)} p_i p_j p_l \end{aligned} \tag{27}$$

where

$$\begin{aligned} k &= 1,2,...,K \\ W_d^{n-1}(x'_{j|k}) &= [1\ x'_{j|k}\ (x'_{j|k})^2\ (x'_{j|k})^3][\beta_1^{n-1}\ \beta_2^{n-1}\ \beta_3^{n-1}\ \beta_4^{n-1}]' \\ \beta^{n-1} &= (\Upsilon'\Upsilon)^{-1}\Upsilon'\overrightarrow{W}_d^{n-1}\ n=2,3,.... \\ \beta^0 &\equiv [1\ 0\ 0\ 0]' \end{aligned}$$

$$\lim_{n\to\infty}\sum_{j=1}^{n}\beta^{n-1} = \beta_d$$

This amounts to a sequence of quadrature problems that have to be solved recursively since the price of the asset with maturity n depends on the price of the asset with maturity $n-1$.

6.2.2 Limited Information

In Limited Information, the Price-Consumption Ratio and the stochastic discount factor depend just on one relevant state: $\widehat{x}$, here denoted $\widehat{\Delta c}$. We discretize and bound its support by forming a grid of K points $\{\widehat{\Delta c}_1, \widehat{\Delta c}_2, ... \widehat{\Delta c}_K\}$ on the interval [-5$V(\widehat{\Delta c})$ +5$V(\widehat{\Delta c})$]. We choose K to be odd so that the unconditional mean of the state $\widehat{\Delta c}$ is the middle point of our grid, $\widehat{\Delta c}_t \perp v_{c,t+1}$.

The Euler equation for the Price-Consumption ratio is:

$$W_c(\widehat{\Delta c}_k) = \left(\sum_{j=1}^{I} \delta^{\theta} e^{(1-\gamma)(\mu + \widehat{\Delta c}_k + \sigma_{v_c}\epsilon_j)} [1 + W_c(\widehat{\Delta c}'_{j|k})]^{\theta} p_j \right)^{\frac{1}{\theta}} \tag{28}$$

where

$$\widehat{\Delta c}'_{j|k} = \rho \widehat{\Delta c}_k + (\rho - b_c)\sigma_{v_c}\epsilon_j$$

solved by iterating until convergence the following recursion:

$$\begin{aligned} W_{c,n}(\widehat{\Delta c}_k) &= \left(\sum_{j=1}^{I} \delta^{\theta} e^{(1-\gamma)(\mu + \widehat{\Delta c}_k + \sigma_{v_c}\epsilon_j)} [1 + W_{c,n-1}(\widehat{\Delta c}'_{j|k})]^{\theta} p_j \right)^{\frac{1}{\theta}} \\ n &= 1, 2, ... \end{aligned}$$

where the function is interpolated by a third order polynomial in $\widehat{\Delta c}$ such that:

$$\begin{aligned} W_{c,n-1}(x'_{j|k}) &= [1 \ \widehat{\Delta c}'_{j|k} \ (\widehat{\Delta c}'_{j|k})^2 \ (\widehat{\Delta c}'_{j|k})^3][\beta_{1,n-1} \ \beta_{2,n-1} \ \beta_{3,n-1} \ \beta_{4,n-1}]' \\ \beta_n &= (\Phi'\Phi)^{-1}\Phi'\overrightarrow{W}_{c,n} \text{ n = 1, 2, 3,...} \\ &\text{where} \\ \Phi &= \begin{bmatrix} 1 & \widehat{\Delta c}_1 & (\widehat{\Delta c}_1)^2 & (\widehat{\Delta c}_1)^3 \\ 1 & \widehat{\Delta c}_2 & (\widehat{\Delta c}_2)^2 & (\widehat{\Delta c}_2)^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \widehat{\Delta c}_k & (\widehat{\Delta c}_k)^2 & (\widehat{\Delta c}_k)^3 \\ & & & \end{bmatrix} \\ \beta_0 &: \text{ initial guess} \end{aligned}$$

The price-dividend ratio is a function of the state variable $\widehat{x}_d \equiv \widehat{\Delta d}$ and the shock v_d:

$$\begin{bmatrix} v_{c,t+1} \\ v_{d,t+1} \end{bmatrix} \sim i.i.d.N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{v_c} & \sigma_{v_c,c_d} \\ \sigma_{v_c,v_d} & \sigma_{v_d^2} \end{bmatrix} \right)$$

and

$$\begin{bmatrix} \widehat{\Delta c} \\ \widehat{\Delta d} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\widehat{\Delta c}} & \sigma_{\widehat{\Delta c},\widehat{\Delta d}} \\ \sigma_{\widehat{\Delta c},\widehat{\Delta d}} & \sigma^2_{\widehat{\Delta d}} \end{bmatrix} \right)$$

- A grid of combinations $(\widehat{\Delta d}_{g|k}, \widehat{\Delta c}_k)$ is stacked in a matrix S with dimension $(K \times G) \times 2$:

$$S = \begin{bmatrix} \widehat{\Delta c}_1 & \widehat{\Delta d}_{1|1} \\ \widehat{\Delta c}_1 & \widehat{\Delta d}_{2|1} \\ \vdots & \vdots \\ \widehat{\Delta c}_1 & \widehat{\Delta d}_{g|1} \\ \widehat{\Delta c}_2 & \widehat{\Delta d}_{1|2} \\ \vdots & \vdots \\ \widehat{\Delta c}_K & \widehat{\Delta d}_{g|K} \end{bmatrix}$$

The recursion used to find the price-dividend ratio is given by:

$$\begin{aligned} W_{d,n}(\widehat{\Delta c}_s, \widehat{\Delta d}_s) &= \sum_{j=1}^{I}\sum_{i=1}^{I} \delta^{\theta} e^{-\gamma(\mu + \widehat{\Delta c}_s + \sigma_{v_c}\epsilon_j)} \left(\frac{1 + V_c(\widehat{\Delta c}'_{j|s})}{V_c(\widehat{\Delta c}_s)} \right)^{\theta-1} \times \\ &\quad \times [1 + W_{d,n-1}(\widehat{\Delta c}'_{j|s}, \widehat{\Delta d}'_{i|s})] e^{\mu + \widehat{\Delta d}_s + \sigma_{v_d}\epsilon_i} p_{ij} \\ (\widehat{\Delta c}_s, \widehat{\Delta d}_s) &= [S_{s,1} S_{s,2}] \\ s &= 1, 2, ..., K \times G \end{aligned}$$

The price-dividend ratio is interpolated as above by a quadratic polynomial in the two states:

$$\begin{aligned} W_{d,n-1}(\widehat{\Delta c}_s, \widehat{\Delta d}_s) &= [1 \ \widehat{\Delta c}'_{j|k} \ \widehat{\Delta d}'_{i|k} \ (\widehat{\Delta c}'_{j|k})^2 \ (\widehat{\Delta d}'_{i|k})^2 \ \widehat{\Delta c}'_{j|k} \widehat{\Delta d}'_{i|k}] \times \\ &\quad \times [\beta^d_{1,n-1} \ \beta^d_{2,n-1} \ \beta^d_{3,n-1} \ \beta^d_{4,n-1} \ \beta^d_{5,n-1} \ \beta^d_{6,n-1}]' \\ \beta^d_n &= (\Phi^{d'} \Phi^d)^{-1} \Phi^{d'} \overrightarrow{W}_{d,n} \\ n &= 1, 2, 3, ... \end{aligned}$$

where

$$\begin{aligned} \Phi^d &= \begin{bmatrix} 1 & S_{1,1} & S_{1,2} & S^2_{1,1} & S^2_{1,2} & S_{1,1}S_{1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & S_{G\times K,1} & S_{G\times K,2} & S^2_{G\times K,1} & S^2_{G\times K,2} & S_{G\times K,1}S_{G\times K,2} \end{bmatrix} \\ \beta_0 &: \text{initial guess} \end{aligned}$$

Foe zero coupon equity price-dividends, we implement the following recursion:

$$\begin{aligned} W_d^n(\widehat{\Delta c}_s, \widehat{\Delta d}_s) &= \sum_{j=1}^{I}\sum_{i=1}^{I} \delta^\theta e^{-\gamma(\mu+\widehat{\Delta c}_s+\sigma_{v_c}\epsilon_j)} \left(\frac{1+V_c(\widehat{\Delta c}'_{j|s})}{V_c(\widehat{\Delta c}_s)} \right)^{\theta-1} \times \qquad (29) \\ & \times W_d^{n-1}(\widehat{\Delta c}'_{j|s}, \widehat{\Delta d}'_{i|s}) e^{\mu+\widehat{\Delta d}_s+\sigma_{v_d}\epsilon_i} p_{ij} \\ W_d^{n-1}(\widehat{\Delta c}'_{j|s}, \widehat{\Delta d}'_{i|s}) &= [1\ \widehat{\Delta c}'_{j|k}\ \widehat{\Delta d}'_{i|k}\ (\widehat{\Delta c}'_{j|k})^2\ (\widehat{\Delta d}'_{i|k})^2\ \widehat{\Delta c}'_{j|k}\widehat{\Delta d}'_{i|k}] \times \\ & \times [\beta_1^{n-1}\ \beta_2^{n-1}\ \beta_3^{n-1}\ \beta_4^{n-1}\ \beta_5^{n-1}\ \beta_6^{n-1}]' \\ \beta_d^n &= (\Phi^{d'}\Phi^d)^{-1}\Phi^{d'}\overrightarrow{W}_d^n \\ n &= 1, 2, 3, ... \\ \beta_d^0 &= [1\ 0\ 0\ 0\ 0\ 0]. \end{aligned}$$

6.3 Cash Flow Betas

Table A.1 shows the output from regressions of dividend growth on 4 and 8 quarter trailing averages of consumption growth, using simulated data for cash flow models of the form (13)-(15). The slope coefficients in these regressions are denoted φ, and are reported for four models that vary only by the short-run risk exposure parameter ϕ_c. The model is

$$\Delta d_{t+1} = \alpha + \varphi \left(\frac{1}{K} \sum_{i=1}^{K} \Delta c_{t+1-i} \right) + \varepsilon_{t+1}.$$

The model is simulated at a monthly frequency, consumption and dividend data are time-aggregated to quarterly frequency, and regressions run on quarterly data, as in Bansal et al. (2006). The results for one parameter configuration are displayed in Table A.1, but findings for other parameter configurations studied in the main text are similar. The Table shows that heterogeneity in exposure to short-run consumption risk can generate heterogeneity in cash flow betas φ, when the cash flow betas are constructed from $K = 4$ and $K = 8$ quarter trailing moving averages of consumption growth. This occurs only when the data are time-averaged; regressions on monthly data produce no such discernible spread in cash flow betas across assets that differ solely by ϕ_c. The reason is that time-averaging introduces additional serial correlation into the growth rates of consumption and dividends. The overlapping nature of the time-aggregate data therefore generates a correlation between dividend growth and lagged consumption growth that rises with the sensitivity of dividend growth to consumption risk that is i.i.d. at the monthly frequency (but not at the time-aggregate quarterly frequency). The longer the horizon K, the smaller is this affect.

References

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

___, ___, and ___, "A Quartet of Semigroups for Model Specification, Robustness, Prices of Risk, and Model Detection," *Journal of the European Economic Association*, 2003, *1* (1), 68–123.

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.

___, Robert F. Dittmar, and Christian T. Lundblad, "Consumption, Dividends, and the Cross-Section of Equity Returns," *Journal of Finance,* forthcoming, 2006.

___, ___, and Dana Kiku, "Long Run Risks and Equity Returns," 2005. Unpublished paper, Fuqua School, Duke University.

Campbell, John Y., "Consumption-Based Asset Pricing," in George Constantinides, Milton Harris, and Rene Stulz, eds., *Handbook of the Economics of Finance* forthcoming, Amsterdam: North-Holland, 2003.

___ and John H. Cochrane, "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 1999, *107*, 205–251.

Cochrane, John H., "Permanent and Transitory Components of GDP and Stock Prices," *Quarterly Journal of Economics*, 1994, *109*, 241–265.

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.

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.

___ **and** ___ , "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 1993, *33*, 3–56.

Graham, Benjamin and David L. Dodd, *Security Analysis*, New York, NY: McGraw Hill, 1934.

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," 2005. Unpublished Paper, University of Chicago.

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* forthcoming, 2006.

Lintner, J., "Security Prices, Risk and Maximal Gains from Diversification," *Journal of Finance*, 1965, *20*, 587–615.

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

Lynch, Anthony, "Portfolio Choice With Many Risky Assets, Market Clearing, and Cash Flow Predictability," 2003. Unpublished manuscript, New York University, Stern School of Business.

Malloy, Christopher J., Tobias J. Moskowitz, and Annette Vissing-Jorgensen, "Long-run Stockholder Consumption Risk and Asset Returns," 2005. Unpublished manuscript, University of Chicago, GSB.

Menzly, Lior, Tano Santos, and Pietro Veronesi, "Understanding Predictability," *Journal of Political Economy*, February 2004, *112* (1), 1–47.

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.

Santos, Jesus and Pietro Veronesi, "Conditional Betas," 2004. Unpublished paper, University of Chicago.

____ **and** ____ , "Cash-Flow Risk, Discount Risk, and the Value Premium," 2005. Unpublished paper, University of Chicago.

Sharpe, W., "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 1964, *19*, 425–444.

Wachter, Jessica, "A Consumption Based Model of the Term Structure of Interest Rates," *Journal of Financial Economics*, 2006, *79*, 365–399.

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

Asset Pricing Implications: Full Information vs. Limited Information

Row	Model		$E(P/D)$			$E(r_i - r_f)$			$E(r_f)$			$\sigma(r_i)$		
	ϕ_x	ϕ_c	FI	LI-A	LI-V	FI	LI-A	LI-V	FI	LI-A	LI-V	FI	LI-A	LI-V
1	1	2.2	166	300	164	1.06	1.20	1.63	1.37	0.95	0.81	17.29	17.43	17.42
2	1	6	46	14	45	2.45	7.73	3.04	1.37	0.95	0.81	22.79	22.44	22.87
3	2	2.2	35	300	34	3.31	1.26	4.00	1.37	0.95	0.81	18.40	19.95	18.92
4	2	6	23	14	22	4.90	8.12	5.62	1.37	0.95	0.81	23.67	23.92	24.02
5	3	2.2	22	238	21	5.20	1.26	6.01	1.37	0.95	0.81	20.24	23.29	21.19
6	3	6	17	13	16	6.63	8.42	7.46	1.37	0.95	0.81	25.02	26.09	25.83

Notes: This table reports financial statistics of the model with full information (FI) and limited information based on system (LI-V) and univariate (LI-A) signal extraction, for varying degrees of exposure to the long-run and short-run risk components, governed by ϕ_x and ϕ_c, respectively. The other parameters are set to $\gamma = 10$, $\psi = 1.5$, $\delta = 0.998985$, $\mu = 0.0015$, $\rho = 0.979$, $\sigma = 0.0078$, $\sigma_{xc} = 0.044$, $\sigma_d = 6$. $E(r_i - r_f)$ denotes the annual log risk-premium, in percent; $E(r_f)$ denotes the annual log risk-free rate, in percent, and $\sigma(r_i)$ and $\sigma(r_f)$ denote the standard deviations of the annual equity return and risk-free rate, respectively. $E(P/D)$ is the annual price-dividend ratio. Statistics are averages from 1000 simulated samples of 840 monthly observations.

Table 2

Limited Information Implications of Value and Growth Portfolios: Univariate Signal Extraction

		G			Growth to Value						V	V-G
Portfolio		1	2	3	4	5	6	7	8	9	10	10-1
$E\left(R^i - R^f\right)$	Model	6.39	6.41	6.45	6.55	6.76	7.16	7.89	9.01	10.28	11.47	5.08
	Data	6.50	7.56	7.47	7.60	7.48	9.07	9.15	8.98	10.73	11.92	5.42
Sharpe Ratio	Model	0.38	0.38	0.38	0.39	0.40	0.43	0.48	0.55	0.63	0.70	0.33
	Data	0.38	0.49	0.49	0.49	0.64	0.63	0.61	0.72	0.67	0.64	0.26
CAPM: $R_t^i - R_t^f = \alpha_i + \beta_i \left(R_t^m - R_t^f\right) + \varepsilon_{it}$												
α_i	Model	-1.21	-1.18	-1.13	-1.03	-0.80	-0.36	0.44	1.65	2.99	4.22	5.42
	Data	-1.68	-0.05	0.08	0.24	2.47	2.31	2.41	4.10	3.71	4.19	5.87
β_i	Model	1.02	1.02	1.02	1.01	1.01	1.01	1.00	0.99	0.98	0.97	-0.05
	Data	1.10	1.02	1.01	0.96	0.89	0.90	0.86	0.87	0.92	1.00	-0.10

Notes: Results are presented for limited information specifications based on univariate signal extraction. The results in rows labled "Model," are produced as follows. In each simulation year, firms are sorted into deciles based on the price-dividend ratio. Returns are calculated over the subsequent year. Intercepts and slope coefficients are from OLS time-series regressions of excess portfolio returns on the excess market return. Parameter values are set as follows: $\gamma = 16.5$, $\psi = 1.3$, $\delta = 0.997$, $\mu = 0.0015$, $\rho = 0.983$, $\sigma = 0.0057$, $\sigma_{xc} = 0.1$, $\sigma_d = 5.9$ and the market portfolio has $\phi_x = 1$ and $\phi_c = 3$. Results in rows labled "Data" are produced as follows. Portfolios are formed by sorting firms into deciles on the boot-to-market ratio (B/M). Moments are annualized in percentages (multiplied by 1200 in the case of means and $12/\sqrt{12}$ in the case of Sharpe ratios). Intercepts and slope coefficients are calculated from OLS time-series regressions of excess portfolio returns on the excess return on the CRSP value-weighted index. Intercepts are annualized in percentages (multiplied by 1200). The return data are monthly and span the period 1947-2004.

Table 3
Limited Information Implications of Value and Growth Portfolios: Univariate Signal Extraction

CAPM & HML: $R_t^i - R_t^f = \alpha_i + \beta_i \left(R_t^m - R_t^f\right) + \gamma_i HML_t + \varepsilon_{it}$											
	G				Growth to Value					V	V-G
Portfolio	1	2	3	4	5	6	7	8	9	10	10-1
α_i	0.41	0.41	0.40	0.38	0.33	0.24	0.10	-0.04	0.06	0.41	0.00
β_i	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.00
γ_i	-0.30	-0.29	-0.28	-0.26	-0.21	-0.11	0.06	0.31	0.54	0.70	1.00

Notes: Results are presented for limited information specifications based on univariate signal extraction. In each simulation year, firms are sorted into deciles based on the price-dividend ratio. Returns are calculated over the subsequent year. Intercepts and slope coefficients are from OLS time-series regressions of excess portfolio returns on the excess market return together with HML. Parameter values are set as follows: Parameter values are set as follows: $\gamma = 16.5$, $\psi = 1.3$, $\delta = 0.997$, $\mu = 0.0015$, $\rho = 0.983$, $\sigma = 0.0057$, $\sigma_{xc} = 0.1$, $\sigma_d = 5.9$ and the market portfolio has $\phi_x = 1$ and $\phi_c = 3$.

Table A1

Cash Flow Betas

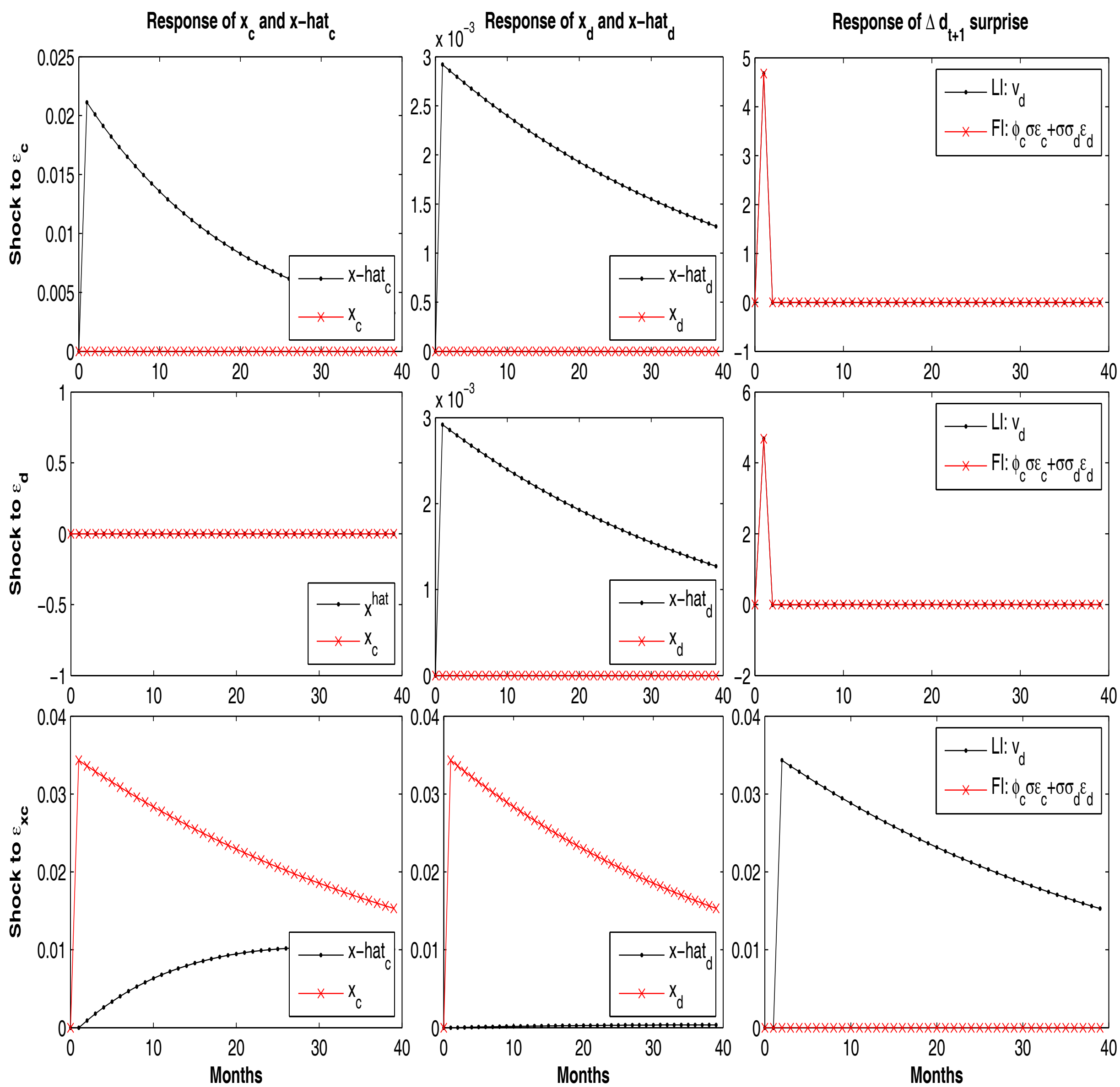
Regression: $\Delta d_{t+1} = \alpha + \varphi \left(\frac{1}{K} \sum_{i=1}^{K} \Delta c_{t+1-i} \right) + \varepsilon_{t+1}$

		$K=4$		$K=8$	
		φ	t-stat	φ	t-stat
$\phi_x = 3$	$\phi_c = 0.5$	0.96	1.69	1.23	1.71
$\phi_x = 3$	$\phi_c = 3$	1.19	1.94	1.37	1.76
$\phi_x = 3$	$\phi_c = 6$	1.45	1.95	1.52	1.61
$\phi_x = 3$	$\phi_c = 10$	1.80	1.81	1.73	1.36

Notes: This table displays regression coefficients and t-statistics from regressions of quarterly dividend growth on to smoothed consumption growth. The quarterly data are time-aggregated from monthly data. The reported statistics are averages from 1000 simulations of length 1000 months (250 quarters). The other parameters are set to $\mu = 0.0015$, $\rho = 0.979$, $\sigma = 0.0078$, $\sigma_{xc} = 0.044$, $\sigma_d = 6$.

Figure 1

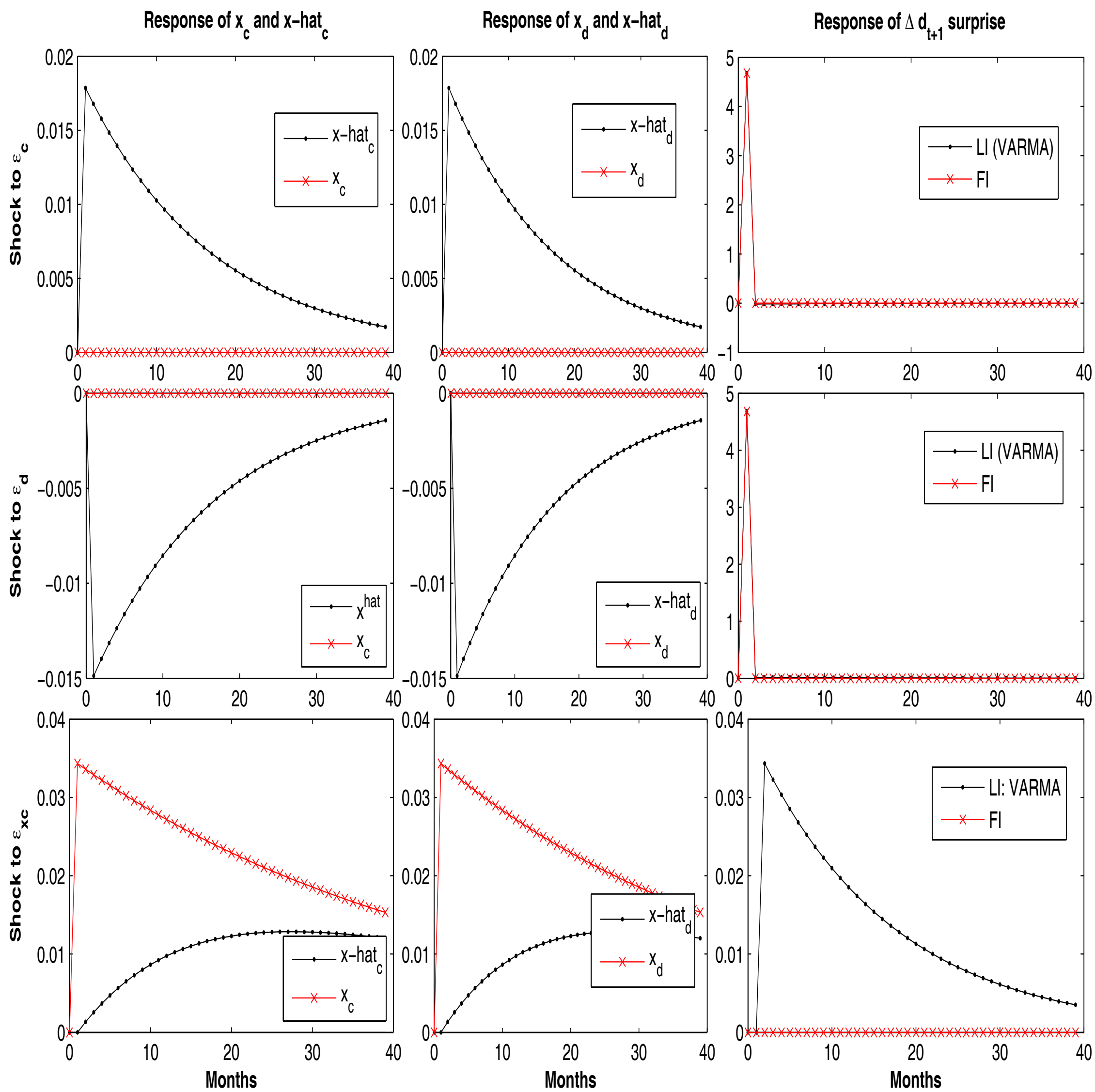
Impulse Responses of Cash Flow Forecasts and Surprises: Univariate Signal Extraction v.s. Full Information



Notes: The figure shows the 40-month 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 univariate signal extraction limited information case, based on ARMA(1,1) estimation. Variables without a "hat" are from the full information benchmark. The responses are based on the calibration $\delta = 0.998985$, $\mu_d = \mu_c = 0.0015$, $\rho = 0.979$, $\sigma = 0.0078$, $\sigma_{xc} = 0.044$, $\Psi = 1.5$, $\gamma = 10$, $\sigma_d = 6$, $\phi_x = 1$, $\phi_c = 6$.

Figure 2

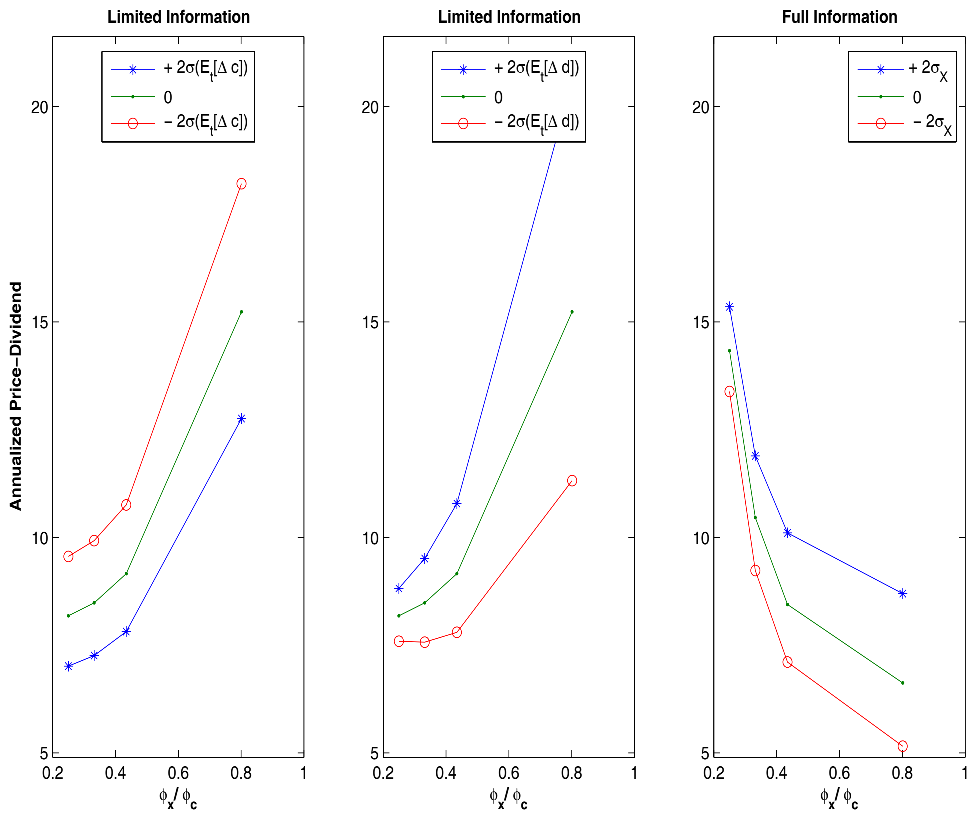
Impulse Responses of Cash Flow Forecasts and Surprises: System Signal Extraction v.s. Full Information



Notes: The figure shows the 40-month 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 system signal extraction limited information specification, based on *VARMA(1,1)* estimation. Variables without a "hat" are from the full information benchmark. The responses are based on the calibration $\delta = 0.998985$, $\mu_d = \mu_c = 0.0015$, $\rho = 0.979$, $\sigma = 0.0078$, $\sigma_{xc} = 0.044$, $\Psi = 1.5$, $\gamma = 10$, $\sigma_d = 6$, $\phi_x = 1$, $\phi_c = 6$.

Figure 3

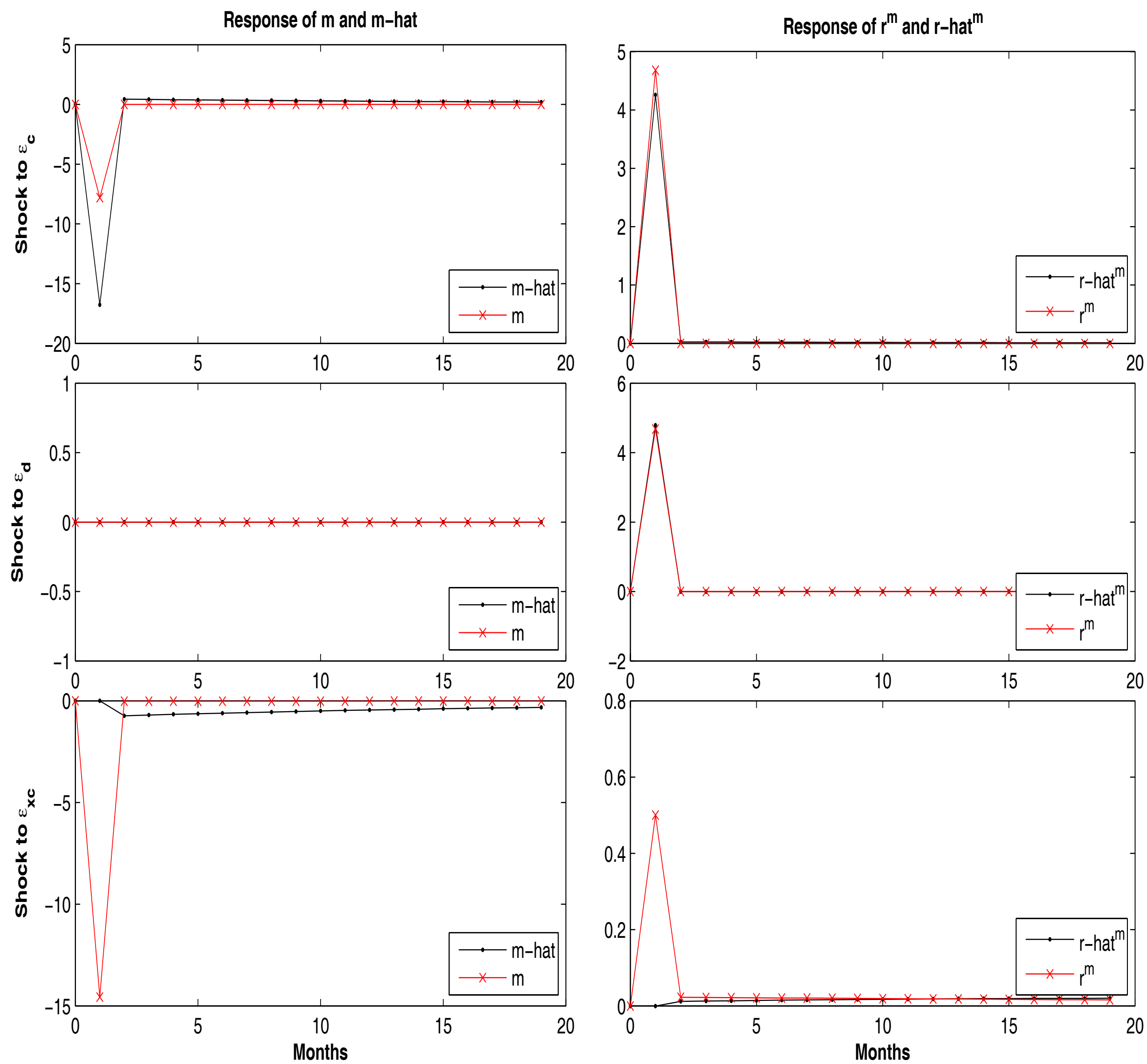
Price-Dividend Ratios: Univariate Signal Extraction v.s. Full Information



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 five-quarter variance of dividend growth attributable to the consumption innovations. The limited information model corresponds to the univariate signal extraction specifications discussed in the text. Parameter values are set as follows: γ=16.5, Ψ=1.3, δ=0.997, $\mu_c = \mu_d$=0.0015, ρ=0.983, σ=0.0057, σ_{xc}=0.1, σ_d=5.9 and the market portfolio has ϕ_x=1 and ϕ_c=3.

Figure 4

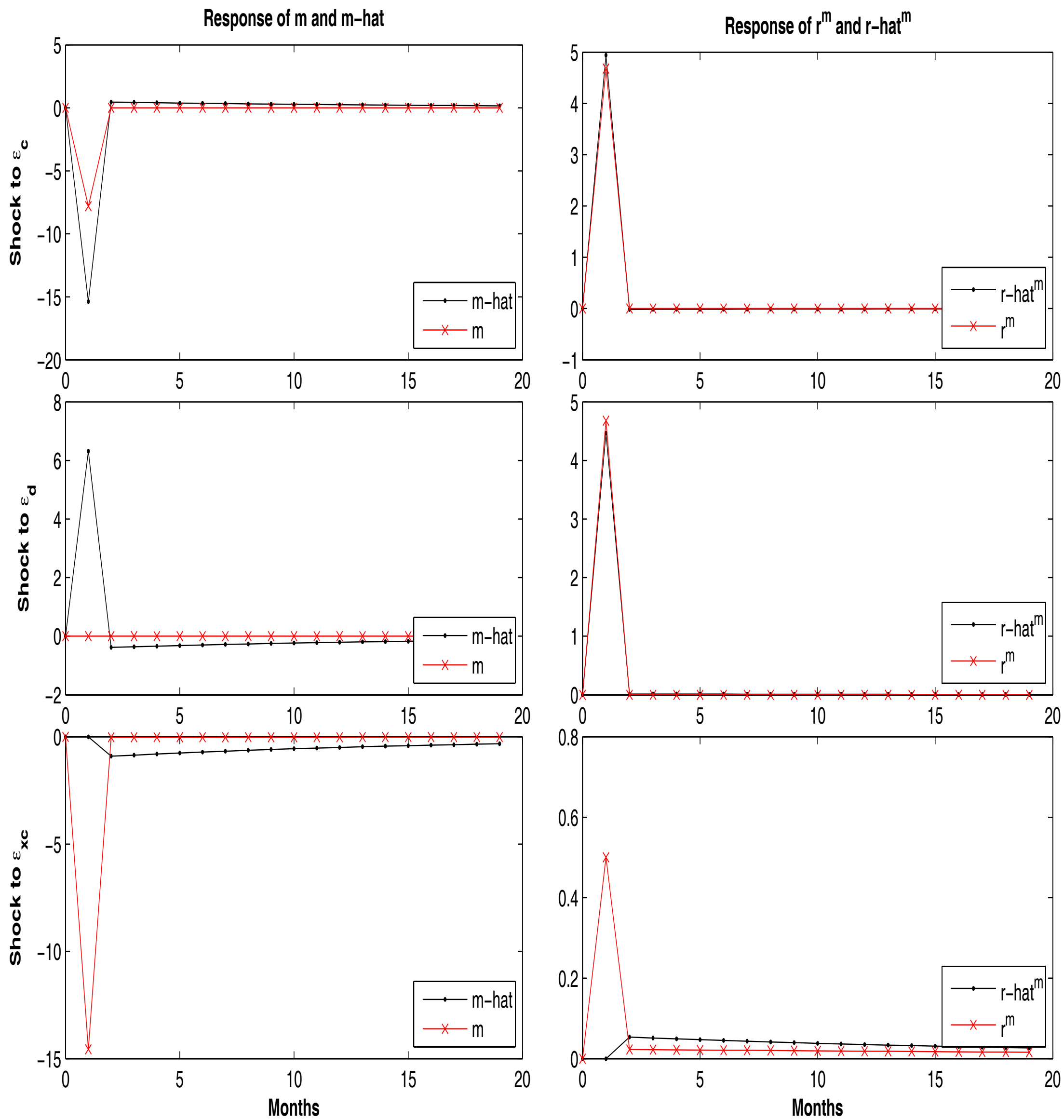
Impulse Responses of SDF and Returns: Univariate Signal Extraction v.s. Full Information



Notes: The figure shows the 40-month 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 univariate signal extraction limited information case, based on *ARMA(1,1)* estimation. Variables without a "hat" are from the full information benchmark. The variable r^m denotes the return on the dividend claim; *m* denotes the stochastic discount factor. The responses are based on the calibration $\delta = 0.998985$, $\mu_d = \mu_c = 0.0015$, $\rho = 0.979$, $\sigma = 0.0078$, $\sigma_{xc} = 0.044$, $\Psi = 1.5$, $\gamma = 10$, $\sigma_d = 6$, $\phi_x = 1$, $\phi_c = 6$.

Figure 5

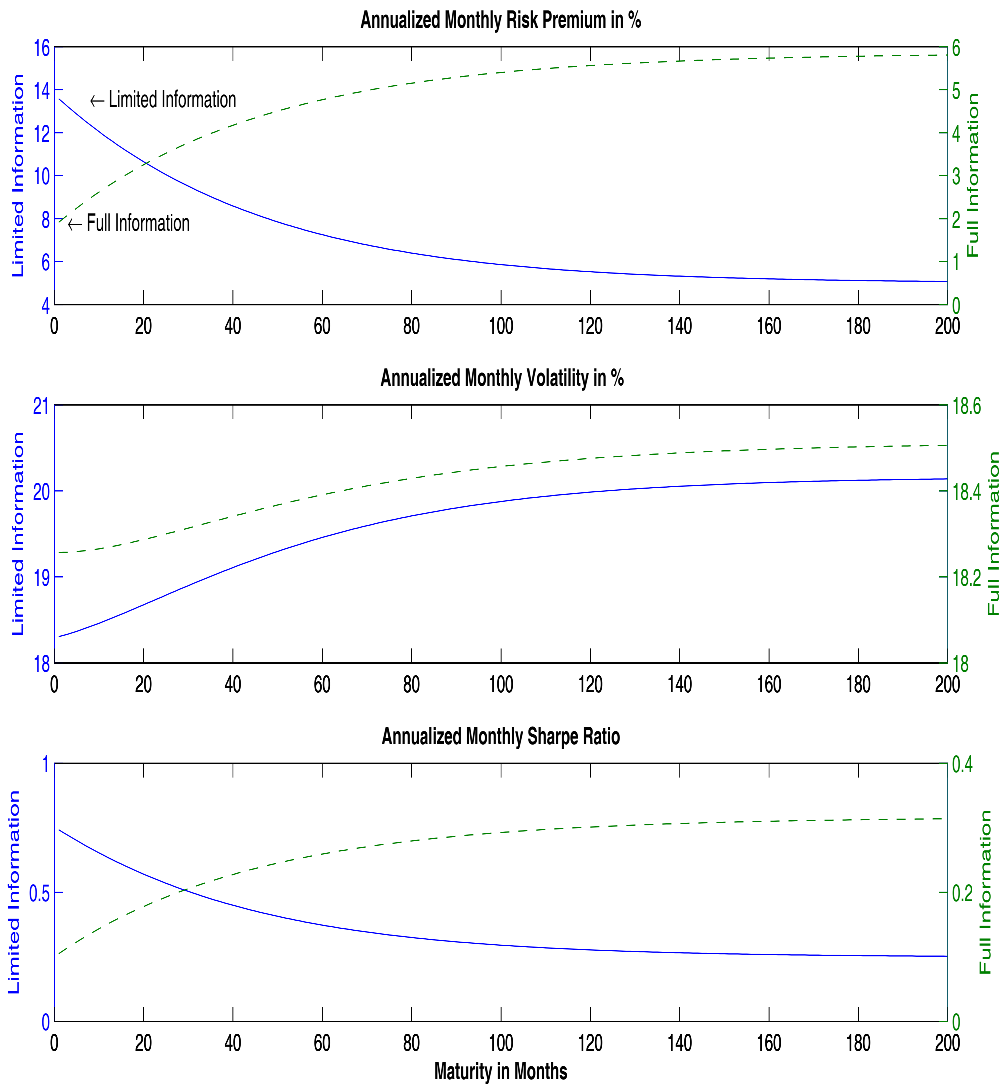
Impulse Responses of SDF and Returns: System Signal Extraction v.s. Full Information



Notes: The figure shows the 40-month 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 system signal extraction limited information case, based on *VARMA(1,1)* estimation. Variables without a "hat" are from the full information benchmark. The variable r^m denotes the return on the dividend claim; *m* denotes the stochastic discount factor. The responses are based on the calibration $\delta = 0.998985$, $\mu_d = \mu_c = 0.0015$, $\rho = 0.979$, $\sigma = 0.0078$, $\sigma_{xc} = 0.044$, $\Psi = 1.5$, $\gamma = 10$, $\sigma_d = 6$, $\phi_x = 1$, $\phi_c = 6$.

Figure 6

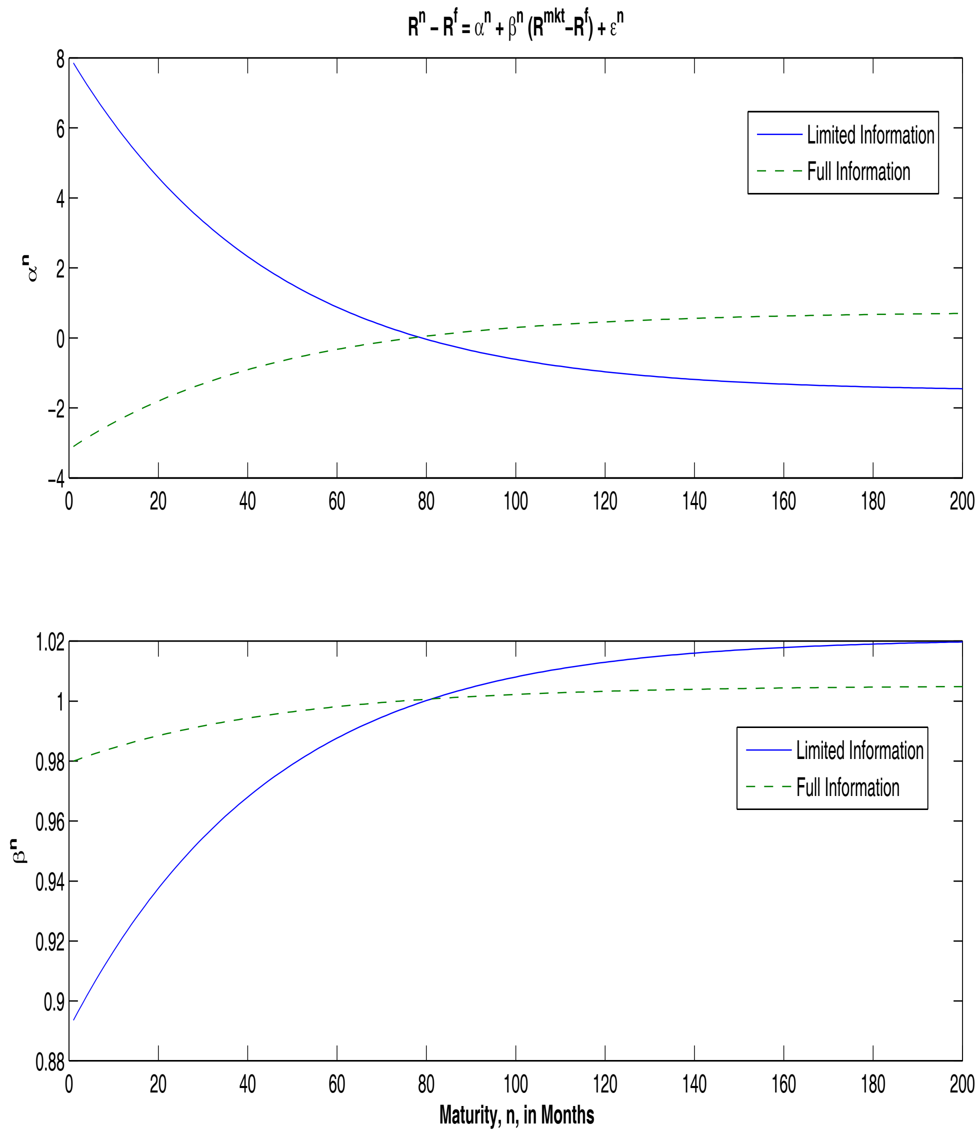
Zero-Coupon Equity: Univariate Signal Extraction v.s. Full Information



Notes: The top panel shows log risk-premia on zero-coupon equity $E\left(r_{n,t+1} - r^{f}{}_{t+1}\right)$ 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. Returns are simulated at a monthly frequency and aggregated to annual frequency. The limited information model corresponds to the univariate signal extraction specifications discussed in the text. Parameter values are set as follows: γ=16.5, Ψ=1.3, δ=0.997, $\mu_c = \mu_d$=0.0015, ρ=0.983, σ=0.0057, σ_{xc}=0.1, σ_d=5.9 and the market portfolio has ϕ_x=1 and ϕ_c=3.

Figure 7

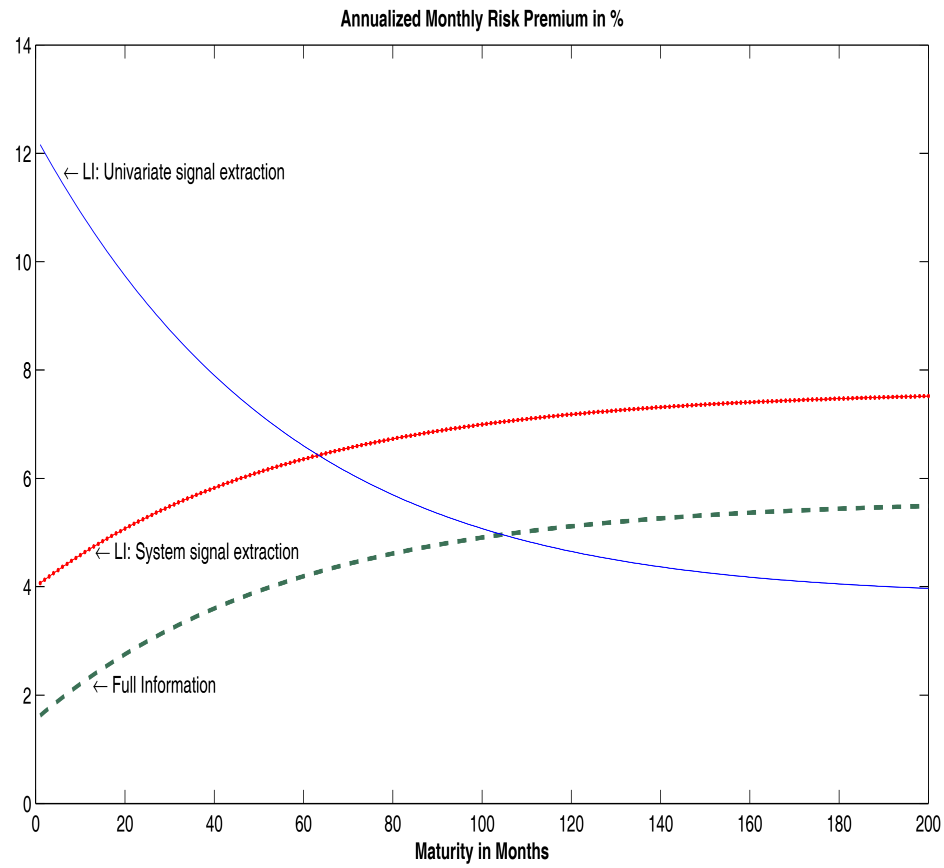
CAPM Regressions for Zero-Coupon Equity: Univariate Signal Extraction v.s. Full Information



Notes: The top panel shows the intercept from regressions of 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. Returns are simulated at a monthly frequency and aggregated to annual frequency. The limited information model corresponds to the univariate signal extraction specifications discussed in the text. Parameter values are set as follows: γ=16.5, Ψ=1.3, δ=0.997, $\mu_c = \mu_d$=0.0015, ρ=0.983, σ=0.0057, σ_{xc}=0.1, σ_d=5.9 and the market portfolio has ϕ_x=1 and ϕ_c=3.

Figure 8

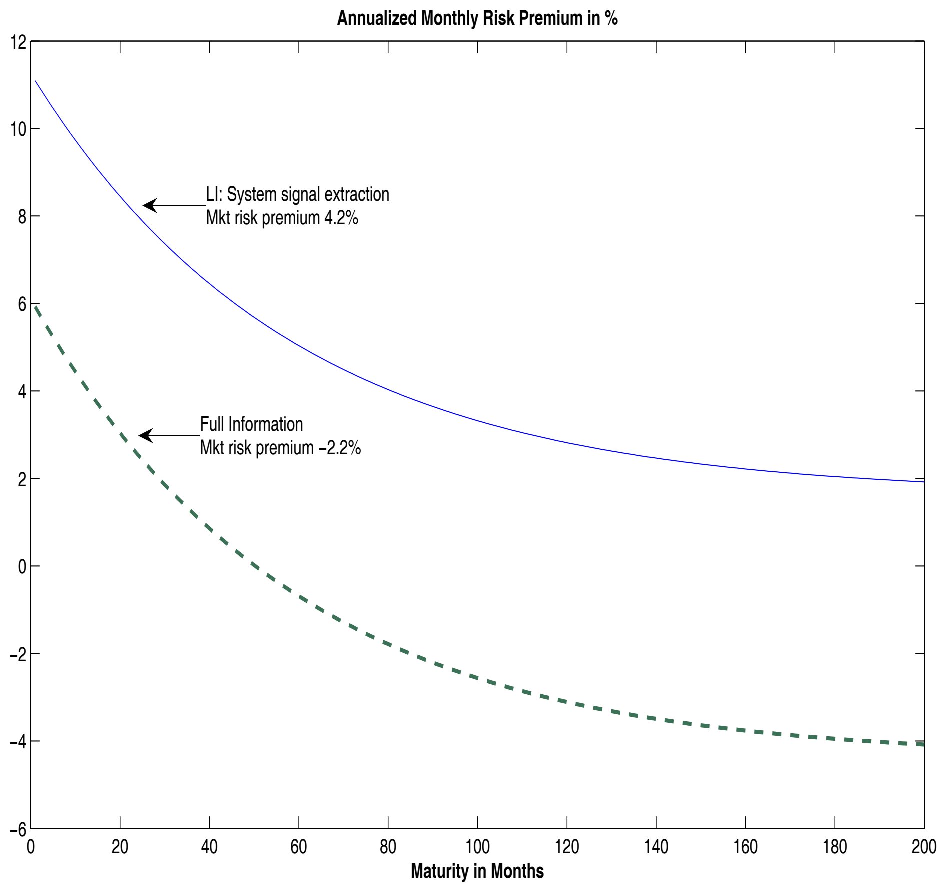
Zero-Coupon Equity: Two Types of Limited Information v.s. Full Information



Notes: The figure shows log risk-premia on zero-coupon equity $E\left(r_{n,t+1} - r^{f}{}_{t+1}\right)$ as a function of maturity, n, in months. Returns are simulated at a monthly frequency and aggregated to annual frequency. Parameter values are set as follows: γ=16.5, Ψ=1.3, δ=0.997, $\mu_c = \mu_d$=0.0015, ρ=0.983, σ=0.0057, σ_{xc}=0.1, σ_d=5.9 and the market portfolio has ϕ_x=1 and ϕ_c=3.

Figure 9

Zero-Coupon Equity: System Signal Extraction and Long-Run Insurance Full Information Model



Notes: The figure shows log risk-premia on zero-coupon equity $E\left(r_{n,t+1} - r^{f}{}_{t+1}\right)$ as a function of maturity, *n*, in months. Returns are simulated at a monthly frequency and aggregated to annual frequency. Parameter values are set as follows: γ=50, Ψ=1, $\delta = 0.99327$, $\mu_c = \mu_d$=0.0015, ρ=0.983, σ=0.0057, σ_{xc}=0.1, σ_d=5.9 and the market portfolio has $\phi_x = 0.78$ and $\phi_c = 3.6$.