Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach

Frank Schorfheide* — University of Pennsylvania, CEPR and NBER
Dongho Song — Boston College
Amir Yaron — University of Pennsylvania, NBER

This Version: October 4, 2017

Abstract

We document that consumption growth rates are far from *iid* and have a highly persistent component. First, we estimate univariate and multivariate models of cash-flow (consumption, output, dividends) growth that feature measurement errors, time-varying volatilities, and mixed-frequency observations. Monthly consumption data are important for identifying the stochastic volatility process; yet the data are contaminated, which makes the inclusion of measurement errors essential for identifying the predictable component. Second, we develop a novel state-space model for cash flows and asset prices that imposes the pricing restrictions of a representative-agent endowment economy with recursive preferences. To estimate this model we use a particle MCMC approach that exploits the conditional linear structure of the approximate equilibrium. Once asset return data are included in the estimation, we find even stronger evidence for the persistent component and are able to identify three volatility processes: the one for the predictable cash-flow component is crucial for asset pricing, whereas the other two are important for tracking the data. Our model generates asset prices that are largely consistent with the data in terms of sample moments and predictability features. The state-space approach allows us to track over time the evolution of the predictable component, the volatility processes, the decomposition of the equity premium into risk factors, and the variance decomposition of asset prices.

*Correspondence: Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297. Email: schorf@ssc.upenn.edu (Frank Schorfheide). Department of Economics, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467. Email: dongho.song@bc.edu (Dongho Song). The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6367. Email: yaron@wharton.upenn.edu (Amir Yaron). We thank Gianluca Violante (co-editor), four anonymous referees, Bent J. Christensen, Ian Dew-Becker, Frank Diebold, Emily Fox, Roberto Gomez Cram, Lars Hansen, Arthur Lewbel, Lars Lochstoer, Ivan Shaliastovich, Neil Shephard, Minchul Shin, and participants at many seminars and conferences for helpful comments and discussions. Schorfheide gratefully acknowledges financial support from the National Science Foundation under Grant SES 1424843. Yaron thanks the Rodney White and Jacob Levy centers for financial support.

1 Introduction

The dynamics of aggregate consumption play a key role in business cycle models, tests of the permanent income hypothesis, and asset pricing. Perhaps surprisingly, there is a significant disagreement about the basic time series properties of consumption. First, while part of the profession holds a long-standing view that aggregate consumption follows a random walk, e.g., Hall (1978) and Campbell and Cochrane (1999), the recent literature on long-run risks (LRR), e.g., Bansal and Yaron (2004) and Hansen, Heaton, and Li (2008), emphasizes the presence of a small persistent component in consumption growth.[1] Second, while time-varying volatility was a feature that until recently was mainly associated with financial time series, there is now a rapidly growing literature stressing the importance of stochastic volatility in macroeconomic aggregates, e.g., Bansal and Yaron (2004), Bloom (2009), and Fernández-Villaverde and Rubio-Ramírez (2011), and the occurrence of rare disasters, e.g., Barro (2009) and Gourio (2012).

Studying consumption growth dynamics leads to the following challenge. On the one hand, it is difficult to identify time variation in volatility based on time-aggregated data, e.g., Drost and Nijman (1993), which favors the use of high-frequency monthly data. On the other hand, monthly consumption growth data are contaminated by measurement error, e.g., Slesnick (1998) and Wilcox (1992), which mask the dynamics of the underlying process. We address this challenge by developing and estimating a novel Bayesian state-space model with an elaborate measurement error component that is consistent with the view that annual and quarterly consumption data are more accurate than monthly data. The model is tailored toward monthly data, but a mixed-frequency approach enables us to accommodate the longest span of annual consumption growth data starting from the Great Depression era.

In the first part of our empirical analysis we provide strong evidence for a persistent component of consumption growth as well as its time-varying volatility, which contradicts the commonly held view that consumption follows a random walk. The combination of measurement errors and the local-level component in "true" consumption growth in our empirical model allows us to generate the strong second-order moving average (MA(2)) component in observed consumption growth. Our basic empirical finding is robust across a wide range of model specifications that include univariate models for consumption growth as well as bivariate models with either output or dividend growth as second observable. The bivariate models feature a common persistent factor in cash-flow growth rates. An important conclusion from our analysis is that plausible models of observed monthly con-

[1] The literature on robustness, e.g., Hansen and Sargent (2007), highlights that merely contemplating low-frequency shifts in consumption growth can be important for macroeconomic outcomes and asset prices.

sumption growth need to contain a MA(2) component, while macroeconomic models that confront monthly data should filter out the high-frequency movements attributable to measurement errors.

In the second part of our empirical analysis, we embed the cash-flow process into a representative agent endowment economy as in Bansal and Yaron (2004). This model is referred to as long run risks (LRR) model. When asset returns are added to the set of observables and the LRR model is jointly estimated with the dynamics of consumption and dividend growth, the credible intervals for a common persistent component in cash-flow growth rates are further sharpened and three separate volatility components are identified: one governing dynamics of the persistent cash-flow growth component, and the other two controlling temporally independent shocks to consumption and dividend growth. The stochastic volatility process for the persistent component is important for asset prices, while the other two volatility processes are important for tracking the data. We show that the estimated LRR model generates asset prices that are largely consistent with the data. Moreover, we demonstrate that if we replace the parameters of the cash-flow processes from the joint estimation by those obtained from the cash-flow-only estimation, the LRR model still has by-and-large realistic asset pricing implications.

In addition to the empirical results, our paper also contains an important technical innovation. To incorporate market returns and the risk-free rate into the state-space model that is used for the second part of the empirical analysis, we have to solve for the asset pricing implications of the LRR model to obtain measurement equations for these two series.[2] Unlike in the cash-flow-only specifications, the model with asset prices has the feature that the volatility processes also affect the conditional means of the asset prices. This considerably complicates the evaluation of the likelihood function with a nonlinear filter as well as the implementation of Bayesian inference. In fact, due to the high-dimensional state space that arises from our measurement error model and the mixed-frequency setting, this nonlinear filtering is a seemingly daunting task. We show how to exploit the partially linear structure of the state-space model to derive a very efficient sequential Monte Carlo (particle) filter.

Unlike the generalized method of moments (GMM) approach that is common in the LRR literature, our sophisticated state-space approach lets us track the predictable component x_t as well as the stochastic volatilities over time. In turn, this allows us to construct period-by-period decompositions of risk premia and asset price variances. Our Bayesian approach allows us to account for three types of statistical uncertainties in a unified manner: parameter uncertainty, uncertainty

[2]In order to solve the model, we approximate the exponential Gaussian volatility processes by linear Gaussian processes such that the standard analytical solution techniques that have been widely used in the LRR literature can be applied. The approximation of the exponential volatility process is used only to derive the coefficients in the law of motion of the asset prices.

about the hidden states, and uncertainty about future (or hypothetical shocks). These three types of uncertainty feature prominently in our empirical results. Depending on the question at hand, we present in some instances credible bands for our results reflecting multiple sources of uncertainty, e.g., when we provide bands for the predictable component of cash flows; and in other instances, to facilitate clear comparisons across parameterizations, we focus on the dominating source of uncertainty, e.g., shock uncertainty when we examine the model-implied sample moments of asset prices or the sampling distribution of R^2's from predictability regressions.

Our empirical analysis starts with the estimation of a state-space model according to which consumption growth is the sum of an *iid* and an AR(1) component, focusing on the persistence ρ of the AR(1) component. We show that once we include monthly measurement errors that average out at the annual frequency, the fit of the model improves significantly, and we obtain an estimate of ρ around 0.92.[3] Using a battery of model specifications, we show that our measurement error model in which measurement errors account for half of the variation in monthly consumption is the preferred one. We further show that the estimation of the monthly model with measurement errors leads to a more accurate estimate of ρ than the estimation with time-aggregated data. Importantly, adding stochastic volatility leads to a further improvement in model fit, a reduction in the posterior uncertainty about ρ, and an increase in the point estimate of ρ to 0.95. Because consumption is generally viewed as being influenced by output fluctuations, we use our framework to show that a similar persistent component is also important for characterizing quarterly GDP dynamics. When we estimate a common persistent component in consumption and output growth (imposing cointegration) inference regarding ρ is essentially the same as in the consumption-only specifications. When we augment the state-space model to include a measurement equation for dividend growth as a precursor to ultimately pricing equity, the joint estimation based on consumption and dividend growth based on post-1959 data leads the estimate of ρ to rise to 0.97.

The LRR model used in the second part of the empirical analysis distinguishes itself from the existing LRR literature in two important dimensions: our model for the cash flows includes measurement errors and three volatility processes to improve the fit. Moreover, we specify an additional process for variation in the time rate of preference as in Albuquerque, Eichenbaum, Luo, and Rebelo (2016), which generates risk-free rate variation that is independent of cash flows and leads to an improved fit for the risk-free rate. The estimation of the LRR model delivers several important empirical findings. First, the estimate of ρ, i.e., the autocorrelation of the persistent cash-flow component, is 0.987, which is higher than what we obtained based on the cash-flow-only estimation.[4]

[3] Without accounting for measurement errors, the estimate of ρ using monthly consumption growth data is insignificantly different from 0 which can partly account for some view that consumption growth is an *iid* process.

[4] The corresponding half-lives of the cash-flow-only (0.97) and asset pricing based (0.987) estimates for ρ are 1.9

Importantly, we show that the time path of the persistent component looks very similar with and without asset price data.

Second, the volatility processes partly capture heteroskedasticity of innovations, and in part they break some of the tight links that the model imposes on the conditional mean dynamics of asset prices and cash flows. This feature significantly improves the model implications for consumption and return predictability. An important feature of our estimation is that the likelihood focuses on conditional correlations between the risk-free rate and consumption — a dimension often not directly targeted in the literature. We show that because consumption growth and its volatility determine the risk-free rate dynamics, one requires another independent process to account for the weak correlation between consumption growth and the risk-free rate. The independent time rate of preference shocks mute the model-implied correlation further and improve the model fit in regard to the risk-free rate dynamics.

Third, it is worth noting that the median posterior estimate for risk aversion is around 9 while it is around 2 for the intertemporal elasticity of substitution (IES). These estimates are broadly consistent with the parameter values highlighted in the LRR literature (see Bansal, Kiku, and Yaron (2012), and Bansal, Kiku, and Yaron (2014)). Fourth, at the estimated preference parameters and those characterizing the consumption and dividend dynamics, the model is able to successfully generate many key asset-pricing moments, and improve model performance relative to previous LRR models along several dimensions. In particular, the posterior median of the equity premium is 8.2%, while the model's posterior predictive distribution is consistent with the observed large volatility of the price-dividend ratio at 0.45, and the R^2s from predicting returns and consumption growth by the price-dividend ratio.

Our paper is connected to several strands of the literature. In terms of the LRR literature, Bansal, Kiku, and Yaron (2014) utilize data that are time-aggregated to annual frequency to estimate the LRR model by GMM and Bansal, Gallant, and Tauchen (2007) pursue an approach based on the efficient method of moments (EMM). Both papers use cash flow *and* asset price data jointly for the estimation of the parameters of the cash flow process. Our likelihood-based approach provides evidence which is broadly consistent with the results highlighted in those paper and other calibrated LRR models, e.g., Bansal, Kiku, and Yaron (2012). Our likelihood function implicitly utilizes a broader set of moments than earlier GMM or EMM estimation approaches. These moments include the entire sequence of autocovariances as well as higher-order moments of the time series used in the estimation and let us measure the time path of the predictable component of cash flows as well as the time path of the innovation volatilities. Rather than asking the model to fit a few selected

and 4.4 years respectively.

moments, we are raising the bar and force the model to track cash flow and asset return time series. Finally, it is worth noting that our paper distinguishes itself from previous LRR literature in showing that even by just using monthly consumption growth data with an appropriate measurement error structure, we are able to estimate the highly persistent predictable component. In complementary research Nakamura, Sergeyev, and Steinsson (2015) show that an estimation based on a long cross-country panel of annual consumption data also yields large estimates of the autocorrelation of the persistent component.

To implement Bayesian inference, we embed a particle-filter-based likelihood approximation into a Metropolis-Hastings algorithm as in Fernández-Villaverde and Rubio-Ramírez (2007) and Andrieu, Doucet, and Holenstein (2010). This algorithm belongs to the class of particle Markov chain Monte Carlo (MCMC) algorithms. Because our state-space system is linear conditional on the volatility states, we can use Kalman-filter updating to integrate out a subset of the state variables. The genesis of this idea appears in the mixture Kalman filter of Chen and Liu (2000). Particle filter methods are also utilized in Johannes, Lochstoer, and Mou (2016), who estimate an asset pricing model in which agents have to learn about the parameters of the cash flow process from consumption growth data. While Johannes, Lochstoer, and Mou (2016) examine the role of parameter uncertainty for asset prices, which is ignored in our analysis, they use a more restrictive version of the cash flow process and do not utilize mixed-frequency observations.[5]

Our state-space setup makes it relatively straightforward to utilize data that are available at different frequencies. The use of state-space systems to account for missing monthly observations dates back to at least Harvey (1989) and has more recently been used in the context of dynamic factor models (see, e.g., Mariano and Murasawa (2003) and Aruoba, Diebold, and Scotti (2009)) and VARs (see, e.g., Schorfheide and Song (2015)). Finally, there is a growing and voluminous literature in macro and finance that highlights the importance of volatility for understanding the macroeconomy and financial markets (see, e.g., Bansal, Khatacharian, and Yaron (2005), Bloom (2009), Fernández-Villaverde and Rubio-Ramírez (2011), Bansal, Kiku, and Yaron (2012), and Bansal, Kiku, Shaliastovich, and Yaron (2014)). Our volatility specification that accommodates three processes further contributes to identifying the different uncertainty shocks in the economy.

The remainder of the paper is organized as follows. Section 2 introduces the state-space model for consumption growth and presents the empirical findings based on consumption growth data. In Section 3 we consider multivariate cash-flow models and examine the evidence for a predictable growth rate component in specifications that include GDP growth and dividend growth. Section 4

[5]Building on our approach, Creal and Wu (2015) use gamma processes to model time-varying volatilities and estimate a yield curve model using particle MCMC. Doh and Wu (2015) estimate a nonlinear asset pricing model in which all the asset prices and the consumption process are quadratic rather than linear function of the states.

introduces the LRR asset-pricing model, describes the model solution and the particle MCMC approach used to implement Bayesian inference. Section 5 discusses the empirical findings obtained from the estimation of the LRR model and Section 6 provides concluding remarks. A description of our data sources, analytical derivations, a detailed description of the state-space representations and posterior inference, and additional empirical results are relegated to an Online Appendix.

2 Modeling Consumption Growth

The first step of our analysis is to develop an empirical state-space model for consumption growth. We take the length of the period to be one month. The use of monthly data is important for identifying stochastic volatility processes. Unfortunately, consumption data are less accurate at monthly frequency than at the more widely-used quarterly or annual frequencies. In this regard, the main contribution in this section is a novel specification of a measurement error model for consumption growth, which has the feature that monthly measurement errors average out under temporal aggregation. Moreover, because monthly consumption data have only been published since 1959, we use annual consumption growth rates prior to 1959 and adapt the measurement equation to the data availability.[6] We develop our measurement error model in Section 2.1 and present the empirical results in Section 2.2.

2.1 A Measurement Equation for Consumption

We proceed in two steps. First, we derive a measurement equation for consumption growth at the annual frequency, which is used for pre-1959 data. Second, we specify a measurement equation for consumption growth at the monthly frequency, which is used for post-1959 data. We use C_t^o and C_t to denote the observed and the "true" level of consumption, respectively. Moreover, we represent the monthly time subscript t as $t = 12(j-1)+m$, where $m = 1, \ldots, 12$. Here j indexes the year and m the month within the year.

Measurement of Annual Consumption Growth. We define annual consumption as the sum of monthly consumption over the span of one year, i.e., $C^a_{(j)} = \sum_{m=1}^{12} C_{12(j-1)+m}$. Log-linearizing this relationship around a monthly value C_* and defining lowercase c as percentage deviations from the log-linearization point, i.e., $c = \log C/C_*$, we obtain $c^a_{(j)} = \frac{1}{12}\sum_{m=1}^{12} c_{12(j-1)+m}$. Defining monthly consumption growth as the log difference

$$g_{c,t} = c_t - c_{t-1},$$

[6] In principle we could utilize the quarterly consumption growth data from 1947 to 1959, but we do not in this version of the paper.

we can deduce that annual growth rates are given by

$$g^a_{c,(j)} = c^a_{(j)} - c^a_{(j-1)} = \sum_{\tau=1}^{23} \left(\frac{12 - |\tau - 12|}{12} \right) g_{c,12j-\tau+1}. \tag{1}$$

We assume a multiplicative *iid* measurement-error model for the level of annual consumption, which implies that, after taking log differences, observed annual consumption growth (o superscript)

$$g^{a,o}_{c,(j)} = g^a_{c,(j)} + \sigma^a_\epsilon \big(\epsilon^a_{(j)} - \epsilon^a_{(j-1)} \big). \tag{2}$$

Measurement of Monthly Consumption Growth. Consistent with the practice of the Bureau of Economic Analysis (BEA), we assume that the levels of monthly consumption are constructed by distributing annual consumption over the 12 months of a year. We approximate the BEA's data construction by assuming that this distribution is based on an observed monthly proxy series z_t, where z_t is a noisy measure of monthly consumption. The monthly levels of consumption are determined such that the growth rates of monthly consumption are proportional to the growth rates of the proxy series and monthly consumption adds up to annual consumption. A measurement-error model that is consistent with this assumption is the following (a detailed derivation is provided in the Online Appendix):

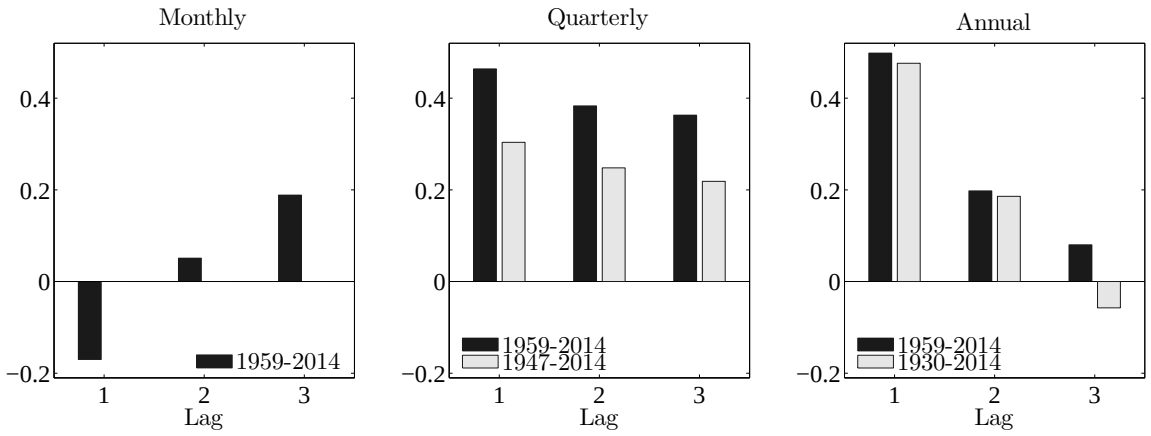
$$\begin{aligned} g^o_{c,12(j-1)+1} &= g_{c,12(j-1)+1} + \sigma_\epsilon \big(\epsilon_{12(j-1)+1} - \epsilon_{12(j-2)+12} \big) \\ &\quad - \frac{1}{12} \sum_{m=1}^{12} \sigma_\epsilon \big(\epsilon_{12(j-1)+m} - \epsilon_{12(j-2)+m} \big) + \sigma^a_\epsilon \big(\epsilon^a_{(j)} - \epsilon^a_{(j-1)} \big) \\ g^o_{c,12(j-1)+m} &= g_{c,12(j-1)+m} + \sigma_\epsilon \big(\epsilon_{12(j-1)+m} - \epsilon_{12(j-1)+m-1} \big), \quad m = 2, \ldots, 12. \end{aligned} \tag{3}$$

The term $\epsilon_{12(j-1)+m}$ can be interpreted as the error made by measuring the level of monthly consumption through the monthly proxy variable, that is, in log deviations $c_{12(j-1)+m} = z_{12(j-1)+m} + \epsilon_{12(j-1)+m}$. The summation of monthly measurement errors in the second line of (3) ensures that monthly consumption sums up to annual consumption. It can be verified that converting the monthly consumption growth rates into annual consumption growth rates according to (1) averages out the measurement errors and yields (2).

2.2 Empirical Analysis

We use the per capita series of real consumption expenditure on nondurables and services from the NIPA tables available from the Bureau of Economic Analysis. Annual observations are available from 1929 to 2014, quarterly from 1947:Q1 to 2014:Q4, and monthly from 1959:M1 to 2014:M12.

Figure 1: Sample Autocorrelation



Notes: Monthly data available from 1959:M2 to 2014:M12, quarterly from 1947:Q2 to 2014:Q4, annual from 1930 to 2014.

Growth rates of consumption are constructed by taking the first difference of the corresponding log series.

Autocorrelation of Consumption Growth. Figure 1 displays the sample autocorrelation of consumption growth for monthly, quarterly and annual data respectively. The figure clearly demonstrates that at the annual frequency consumption growth is strongly positively autocorrelated while at the monthly frequency consumption growth has a negative first autocorrelation. These autocorrelation plots provide prima facie evidence for a negative moving average component at the monthly frequency, which is consistent with the measurement error model described in Section 2.1. Our measurement error model can reconcile the monthly negative autocorrelation with a strongly positive autocorrelation for time aggregated annual consumption. The right panel in Figure 1 also shows that the strong positive autocorrelation in annual consumption growth is robust to using the long pre-war sample as well as the post war data. Given these features of the data, we focus our analysis of measurement errors in consumption using the post 1959 monthly series.

A State-Space Model for Consumption Growth. In our subsequent analysis we will consider several different laws of motion for "true" consumption growth. The benchmark specification takes the following form:

$$\begin{aligned} g_{c,t+1} &= \mu_c + x_t + \sigma_{c,t}\eta_{c,t+1}, \quad \eta_{c,t+1} \sim N(0,1) \\ x_{t+1} &= \rho x_t + \sqrt{1-\rho^2}\varphi_x \sigma_{c,t}\eta_{x,t+1}, \quad \eta_{x,t+1} \sim N(0,1) \\ \sigma_{c,t} &= \sigma \exp(h_{c,t}), \quad h_{c,t+1} = \rho_{h_c} h_{c,t} + \sigma_{h_c} w_{c,t+1}, \quad w_{c,t+1} \sim N(0,1). \end{aligned} \tag{4}$$

This specification is based on Bansal and Yaron (2004) and decomposes consumption growth $g_{c,t+1}$ into a persistent component, x_t, and a transitory component, $\sigma_{c,t}\eta_{c,t+1}$. The state-transition equation is augmented by the measurement equations (2) and (3) to form a state-space model.

The combination of measurement and state-transition equations leads to a high-dimensional state-space model; see the Online Appendix for details. The data that we are using for the estimation have the property that monthly consumption is consistent with annual consumption. While the statistical agency may have access to the monthly proxy series z_t in real time, it can only release the monthly consumption series that is consistent with the corresponding annual consumption observation at the end of each year. Thus, we specify the state-space model such that every 12 months the econometrician observes 12 consumption growth rates. This implies that in addition to tracking the monthly measurement errors $\epsilon_{12(j-1)+m}$ for $m = 1, \ldots, 12$, the state-space model also has to track 12 lags of x_t.

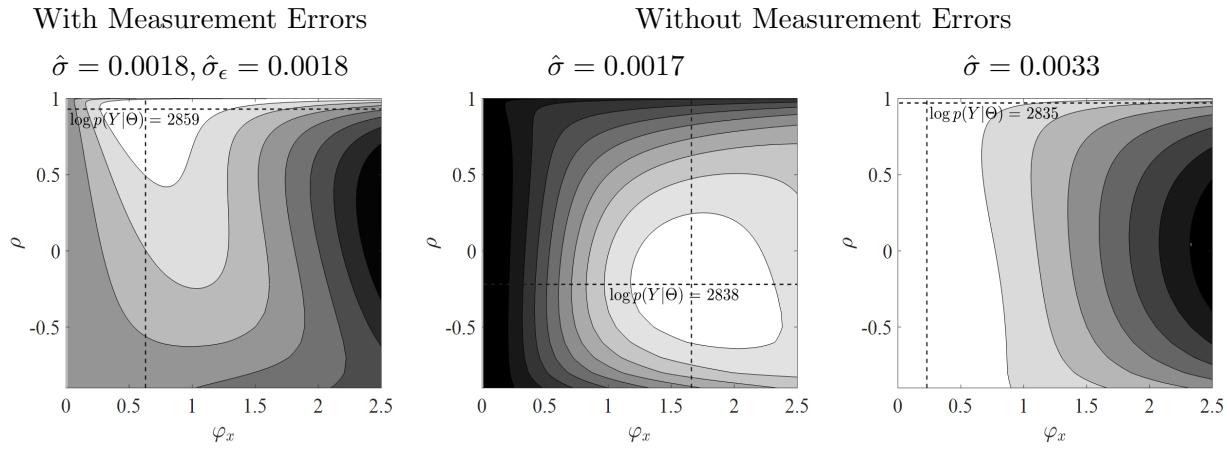
Throughout this paper we use Bayesian inference for the model parameters and the hidden states. In addition to the latent monthly consumption growth rates and measurement errors, the state space also comprises the latent volatility process $h_{c,t}$. Define the parameter vectors

$$\Theta_{cf} = [\mu_c, \rho, \varphi_x, \sigma]', \quad \Theta_h = [\rho_{h_c}, \sigma_{h_c}]$$

and the sequence of latent volatilities $H_{0:T-1}$. To sample from the posterior distribution of $(\Theta_{cf}, \Theta_h, H_{0:T-1})$ we use a Metropolis-within-Gibbs algorithm that iterates over three conditional distributions: First, a Metropolis-Hastings step is used to draw from the posterior of Θ_{cf} conditional on $\big(Y, \Theta_h^{(s-1)}, H_{0:T-1}^{(s-1)}\big)$. Here the likelihood $p(Y|\Theta_{cf}, H_{0:T-1}^{(s-1)})$ is evaluated with the Kalman filter. Second, we draw the sequence of stochastic volatilities $H_{0:T-1}$ conditional on $\big(Y, \Theta_{cf}^{(s)}, \Theta_h^{(s-1)}\big)$ using the algorithm developed by Kim, Shephard, and Chib (1998). This step involves the use of a simulation smoother (e.g., Carter and Kohn (1994)) for a linear state-space model to obtain draws from the conditional posterior distributions of the "residuals" $g_{c,t+1} - \mu_c - x_t$ and $x_{t+1} - \rho x_t$. Conditional on these residuals, it is possible to draw from the posterior of $H_{0:T-1}$. Finally, we draw from the posterior of the coefficients of the stochastic volatility processes, Θ_h, conditional on $\big(Y, H_{0:T-1}^{(s)}, \Theta_{cf}^{(s)}\big)$.

The Likelihood Function. We simplify the law of motion of consumption growth in (4) by assuming that the innovations are homoskedastic, i.e., $\sigma_{h_c} = 0$ and $h_{c,t} = 0$ for all t. In Figure 2 we plot likelihood function contours based on a sample of monthly consumption growth rates that ranges from 1959:M2 to 2014:M12. We consider two specifications: with and without measurement errors. To isolate the role of measurement errors for inference about ρ, we set μ_c to the sample mean and fix σ and σ_ϵ to their respective maximum likelihood estimates, while varying the two

Figure 2: Log-Likelihood Contour



Notes: We use maximum likelihood estimation to estimate the homoskedastic version ($\sigma_{h_c} = 0$, $h_{c,t} = 0$) of model (4) with and without allowing for measurement errors. We then fix $\sigma = \hat{\sigma}$ and $\sigma_\epsilon = \hat{\sigma}_\epsilon$ at their point estimates and vary ρ and φ_x to plot the log-likelihood function contour. Without measurement errors, we find that the log-likelihood function is bimodal at positive and negative values of ρ. Therefore, we obtain two sets of $\hat{\sigma}$.

parameters, ρ and φ_x, that govern the dynamics of x_t. In the absence of measurement errors the log-likelihood function is bimodal. The first mode is located at $\rho = -0.23$ which matches the negative monthly sample autocorrelation (see Figure 1). The location of the second mode is at $\rho = 0.96$, but the log-likelihood function is flat across a large set of values of ρ between -1 and 1. Importantly, when we allow for monthly measurement errors according to (3), setting $\sigma_\epsilon^a = 0$, the log likelihood function has a very sharp peak, displaying a very persistent expected consumption growth process with $\rho = 0.92$. Measurement errors at the monthly frequency help identify a large persistent component in consumption by allowing the model to simultaneously match the negative first-order autocorrelation observed at the monthly frequency and the large positive autocorrelation at the annual frequency.

Bayesian Estimation of Homoskedastic Models. We now proceed with the Bayesian estimation of variants of (4) using the monthly sample ranging from 1959:M2 to 2014:M12. Table 1 reports quantiles of the prior distribution[7] as well as posterior median estimates of the model parameters. Estimates for the benchmark specification with monthly and annual measurement errors are reported in column (1). We briefly comment on some important aspects of the prior distribution. The prior for ρ (persistence of x_t) is uniform over the interval $(-1, 1)$ and encompasses values that

[7]In general, our priors attempt to restrict parameter values to economically plausible magnitudes. The judgment of what is *economically plausible* is, of course, informed by some empirical observations, in the same way the choice of the model specification is informed by empirical observations.

Table 1: Posterior Median Estimates of Consumption Growth Processes

	Prior Distribution				Posterior Estimates					
					State-Space Model				IID	ARMA
					M&A	No ME AR(2)	M $\rho_\epsilon \neq 0$	M $\rho_\eta \neq 0$		(1,2)
	Distr.	5%	50%	95%	(1)	(2)	(3)	(4)	(5)	(6)
μ_c	N	-.007	.0016	.0100	.0016	.0016	.0016	.0016	.0016	.0016
ρ	U	-.90	0	.90	.918	-.684	.918	.919	-	.913
ρ_2	U	-.90	0	.90	-	-.353	-	-	-	-
φ_x	U	.05	0.5	.95	.681		.704	.644	-	-
	U	.1	1.0	1.9	-	.482	-	-	-	-
σ	IG	.0008	.0019	.0061	.0018	.0027	.0017	.0019	.0033	.0032
σ_ϵ	IG	.0008	.0019	.0061	.0018		.0019	.0018	-	-
σ_ϵ^a	IG	.0007	.0029	.0386	.0011	-	-	-	-	-
ρ_ϵ	U	-.90	0	.90	-	-	.060	-	-	-
ρ_η	U	-.90	0	.90	-	-	-	-.046	-	-
ζ_1	N	-8.2	0	8.2	-	-	-	-	-	-1.14
ζ_2	N	-8.2	0	8.2	-	-	-	-	-	.302
$\ln p(Y)$					2887.1	2870.3	2883.9	2885.8	2863.2	2884.0

Notes: The estimation sample is from 1959:M2 to 2014:M12. We denote the persistence of the growth component x_t by ρ (and ρ_2 if follows an AR(2) process), the persistence of the measurement errors by ρ_ϵ, and the persistence of $\eta_{c,t}$ by ρ_η. We report posterior median estimates for the following measurement error specifications of the state-space model: (1) monthly and annual measurement errors (M&A); (2) no measurement errors with AR(2) process for x_t (no ME AR(2)); (3) serially correlated monthly measurement errors (M, $\rho_\epsilon \neq 0$); (4) serially correlated consumption shocks $\eta_{c,t}$ (M, $\rho_\eta \neq 0$, $\rho > \rho_\eta$). In addition we report results for the following models: (5) consumption growth is *iid*; (6) consumption growth is ARMA(1,2).

imply near *iid* consumption growth as well as values for which x_t is almost a unit root process. The parameter φ_x can be interpreted as the square root of a "signal-to-noise ratio," meaning the ratio of the variance of x_t over the variance of the *iid* component $\sigma_c \eta_{c,t+1}$. We use a uniform prior for φ_x that allows for "signal-to-noise ratios" between 0 and 1. At an annualized rate, our *a priori* 90% credible interval for σ and σ_ϵ ranges from 0.3% to 2.1% and the prior for the σ_ϵ^a covers the interval 0.07% to 3.9%. For comparison, the sample standard deviations of annualized monthly consumption growth and annual consumption growth are approximately 1.1% and 2%.

The estimate of ρ obtained from our benchmark specification is approximately 0.92, pointing

toward a fairly persistent predictable component in consumption growth. The estimate of φ_x is 0.68 and implies that the variance of the variance of x_t is roughly 50% smaller than the variance of the *iid* component of consumption growth. At first glance, the large estimate of ρ in the benchmark model may appear inconsistent with the negative sample autocorrelation of monthly consumption growth reported in Figure 1. However, the sample moment confounds the persistence of the "true" consumption growth process and the dynamics of the measurement errors. Our state-space framework is able to disentangle the various components of the observed monthly consumption growth, thereby detecting a highly persistent predictable component x_t that is hidden under a layer of measurement errors.

To assess the robustness of this finding we now modify the benchmark specification in several dimensions. If we shut down the measurement errors and generalize x_t to an AR(2) process (see Column (2)), then the estimates of the autoregressive coefficients turn negative, thereby confirming the graphical pattern in Figure 2. Reverting back to an AR(1) process for x_t and allowing for serially correlated measurement errors (see Column (3)) does not change the estimates of the benchmark model. In fact the estimated autocorrelation for the measurement error is close to zero. Likewise, if we allow for some serial correlation in the transitory component of "true" consumption growth (see Column (4)), the estimate of ρ stays around 0.92. Finally, in the last two columns of Table 1 we report estimates for an *iid* model of consumption growth and an ARMA(1,2) model. We compute log marginal data densities for each specification. Differentials of $\ln p(Y)$ can be interpreted as log posterior odds (under the assumption that the prior odds are 1). The last row of Table 1 shows the benchmark specification dominates all of the alternatives. In particular, the *iid* specification and the state-space model without measurement errors are strongly dominated with log odds of 23.9 and 16.8 in favor of the benchmark.[8]

In order to examine the degree to which measurement errors contribute to the variation in the observed consumption growth, we conduct variance decomposition of monthly and annual consumption growth using measurement error specification of column (1) in Table 1. We find that more than half of the observed monthly consumption growth variation is due to measurement errors.[9] For annual consumption growth data, this fraction drops below 1%. On the other hand,

[8] In a preliminary data analysis we estimated a variety of ARMA(p,q) models using maximum likelihood estimation. We used the Schwarz Information Criterion (BIC) to estimate p and q which lead us to the ARMA(1,2) specification. In the Online Appendix we are reporting results for other variants of the benchmark state-space model. Among them, the specification in which the monthly measurement errors are not restricted to average out over the year is at par with the benchmark specification in terms of fit. For reasons explained above, we find the benchmark specification more appealing.

[9] Wilcox (1992) finds that more than a quarter of the variation in the retail sales series from Detroit and Philadelphia is due to measurement error. For New York this figure is around 50% and LA around 67%. Our estimates are of a

Table 2: Informational Gain Through High-Frequency Observations

Data Frequency	Posterior of ρ 5%	50%	95%	90% Intv. Width
Without Stochastic Volatility				
Monthly	.847	.918	.963	.116
Quarterly	.843	.917	.962	.119
Annual	.621	.893	.983	.362
With Stochastic Volatility				
Monthly	.904	.951	.980	.076
Quarterly	.872	.931	.971	.099

Notes: The estimation sample ranges from 1959:M2 to 2014:M12. The model frequency is monthly. For monthly data we use both monthly and annual measurement errors (specification (1) in Table 1). For quarterly (annual) data we use quarterly (annual) measurement errors only. The model specification is provided in (4).

the opposite pattern holds true for the persistent growth component. While the variation in the persistent growth component only accounts for 13% of the monthly consumption growth variation, this fraction increases to 87% for annual consumption growth data.

Informational Gain Through Temporal Disaggregation. The observation that monthly consumption growth data are strongly contaminated by measurement errors which to a large extent average out at quarterly or annual frequency, suggests that one might be able to estimate ρ equally well based on time-aggregated data. We examine this issue in Table 2. The first row reproduces the ρ estimate from Specification (1) of Table 1. However, we now also report the 5% and 95% quantile of the posterior distribution. Keeping the length of a period equal to a month in the state-space model, we change the measurement equation to link it with quarterly and annual consumption growth data. As the data frequency drops from monthly to annual, the posterior median estimate of ρ falls from 0.92 to 0.89. Moreover, the width of the equal-tail probability 90% credible interval increases from 0.12 to 0.36, highlighting that the use of high-frequency data sharpens inference about ρ.

Hansen, Heaton, and Li (2008) estimate a cointegration model for log consumption and log earnings to extract a persistent component in consumption. The length of a time period in their

similar order of magnitude.

reduced-rank vector autoregression (VAR) is a quarter and the model is estimated based on quarterly data. The authors find that the ratio of long-run to short-run response of log consumption to a persistent growth shock, $\eta_{x,t}$ in our notation, is about two, which would translate into an estimate of ρ of approximately 0.5 for a quarterly model. As a robustness check, we estimate two quarterly versions of the homoskedastic state-space model: without quarterly measurement errors and with quarterly measurement errors. The posterior median estimates of ρ are 0.649 and 0.676, respectively. These results are by and large consistent with the low value reported in Hansen, Heaton, and Li (2008) as well as the estimate in Hansen (2007) under the "loose" prior. Using a crude cube-root transformation, the quarterly ρ estimates translate into 0.866 and 0.878 at monthly frequency and thereby somewhat lower than the estimates obtained by estimating a monthly model on quarterly data.

Accounting for Stochastic Volatility. We now re-estimate the benchmark model (4) allowing for stochastic volatility. Our prior interval for the persistence of the volatility processes ranges from 0.27 to 0.999. The prior for the standard deviation of the consumption volatility process implies that the volatility may fluctuate either relatively little, over the range of 0.7 to 1.2 times the average volatility, or substantially, over the range of 0.4 to 2.4 times the average volatility.

According to Table 2 the width of the 90% credible interval for ρ shrinks from 0.116 to 0.076 for monthly data and from 0.119 to 0.099 for quarterly data.[10] At the same time the posterior median of ρ increases from 0.918 to 0.951 for monthly data and from 0.917 to 0.931 for quarterly data. Without stochastic volatility sharp movements in consumption growth must be accounted for by large temporary shocks reducing the estimate of ρ; however, the presence of stochastic volatility allows the model to account for these sharp movements by fluctuations in the conditional variance of the shocks enabling ρ to be large. We conclude that allowing for heteroskedasticity reduces the posterior uncertainty about ρ and raises the point estimate.

As a by-product, we also obtain an estimate for the persistence, ρ_{h_c}, of the stochastic volatility process in (4). The degree of serial correlation of the volatility also has important implications for asset pricing. Starting from a truncated normal distribution that implies a 90% prior credible set ranging from 0.27 to 0.99, based on monthly observations the posterior credible set ranges from 0.955 to 0.999, indicating that the data favor a highly persistent volatility process $h_{c,t}$. Once the observation frequency is reduced from monthly to quarterly the sample contains less information about the high frequency volatility process and there is less updating of the prior distribution. Now the 90% credible interval ranges from 0.73 to 0.97.[11]

[10] We found that the state-space model with stochastic volatility is poorly identified if the observation frequency is annual, which is why we do not report this case in Table 1.

[11] We conducted a small Monte Carlo experiment in which we repeatedly simulated data from a consumption

Table 3: Posterior Estimates: Consumption Only

		Prior			Posterior			Posterior		
					1959:M2-2014:M12			1930-1959 1960:M1-2014:M12		
	Distr.	5%	50%	95%	5%	50%	95%	5%	50%	95%
Consumption Growth Process										
μ_c	N	-.007	.0016	.0100	.0009	.0016	.0019	.0010	.0016	.0018
ρ	U	-.9	0	.9	.904	.951	.980	.891	.940	.971
φ_x	U	.05	.50	.95	.357	.509	.778	.369	.535	.759
σ	IG	.0008	.0019	.0061	.0017	.0021	.0025	.0017	.0022	.0028
ρ_{h_c}	N^T	.27	.80	.999	.955	.988	.999	.949	.984	.996
$\sigma^2_{h_c}$	IG	.0011	.0043	.0283	.0007	.0014	.0030	.0022	.0054	.0242
Consumption Measurement Error										
σ_ϵ	IG	.0008	.0019	.0061	.0010	.0013	.0016	.0010	.0013	.0016
σ^a_ϵ	IG	.0007	.0029	.0386	.0010	.0015	.0020	.0010	.0198	.0372

Notes: We report estimates of model (4). We adopt the measurement error model of Section 2.1. N, N^T, G, IG, and U denote normal, truncated (outside of the interval $(-1, 1)$) normal, gamma, inverse gamma, and uniform distributions, respectively. We allow for annual consumption measurement errors ϵ^a_t during the periods from 1930 to 1948. We impose monthly measurement errors ϵ_t when we switch from annual to monthly consumption data from 1960:M1 to 2014:M12.

Estimation Based on Mixed-Frequency Data. To measure the small persistent component in consumption growth, one would arguably want to use the longest span of data possibe. Adopting a mixed-frequency approach, we now add annual consumption growth data from 1930 to 1959 to our estimation sample. It is well known from Romer (1986) and Romer (1989) that prewar data on consumption are known to be measured with significantly greater error that exaggerates the size of cyclical fluctuations. To cope with the criticism, we allow for annual measurement errors during 1930-1948 but restrict them to be zero afterwards. This break in measurement errors is also motivated by Amir-Ahmadi, Matthes, and Wang (2016) who provide empirical evidence for larger measurement errors in the early sample before the end of World War II. Importantly, we always account for monthly measurement errors whenever we use monthly data.

growth model with stochastic volatility and then estimated models without and with stochastic volatility. For both specifications the estimate of ρ is downward biased and for the misspecified version without stochastic volatility, the downward bias is slightly larger.

Prior credible intervals and posterior estimates are presented in Table 3. Note that the ρ estimate reported under the 1959:M2-2014:M12 posterior is the same as the estimate reported in Table 2 based on monthly data and the model with stochastic volatility. Extending the sample period reduces the posterior median estimate of ρ slightly, from 0.95 to 0.94. We attribute this change to the large fluctuations around the time of the Great Depression. The width of the credible interval stays approximately the same. Note that at this stage we are adding 30 annual observations to a sample of 671 monthly observations (and we are losing 11 monthly observations from 1959). The standard deviation of the monthly measurement error σ_ϵ is estimated to be about half of σ and is robust to different estimation samples because it is solely identified from monthly consumption growth data. The standard deviation of the annual measurement error is larger than that of monthly measurement error by a factor of 4 (recall that to compare σ_ϵ and σ_ϵ^a one needs to scale the latter by $\sqrt{12}$). This finding is consistent with Amir-Ahmadi, Matthes, and Wang (2016) who find significant presence of measurement errors in output growth during 1930 and 1948.

3 Information From Other Cash-Flow Series

Because aggregate consumption is typically thought of as an endogenous variable that responds to fluctuations in aggregate income, we examine in Section 3.1 whether our evidence for a predictable component in consumption growth can be traced back to GDP and whether estimating a joint model for consumption and GDP has important effects on our inference. In Section 3.2 we include dividend growth data in the estimation of the cash-flow model to set the stage for the subsequent asset-pricing analysis. Finally, we provide a brief summary of the cash-flow estimation results in Section 3.3. Posterior inference for the models considered in this section is implemented with a Metropolis-within-Gibbs sampler that is similar to the one described in Section 2.2.

3.1 Real GDP

We begin the analysis with a monthly model for GDP growth $g_{y,t+1}$ that is identical to the benchmark consumption growth model in (4). Because GDP is only available at quarterly frequency, the measurement equation is

$$g^o_{y,t+1} = \sum_{j=1}^{5} \left(\frac{3 - |j-3|}{3} \right) g_{y,t+2-j}, \quad t = 1, 4, 7, \ldots \tag{5}$$

We estimate this model without measurement errors (this was the preferred specification based on marginal data density comparisons) using observations on per capita GDP growth from 1959:Q1 to

Table 4: Posterior Estimates: GDP Growth Only

		Prior			Posterior		
	Distr.	5%	50%	95%	5%	50%	95%
μ_y	N	-.007	.0016	.0100	.0011	.0017	.0022
ρ	U	.05	.50	.95	.698	.874	.966
φ_x	U	.05	.50	.95	.117	.259	.418
σ	IG	.0008	.0019	.0061	.0040	.0045	.0051
ρ_{h_y}	N^T	.27	.80	.999	.928	.970	.992
$\sigma^2_{h_y}$	IG	.0013	.0043	.0283	.0026	.0086	.0228

Notes: We report estimates of model (4) for GDP growth. N, N^T, IG, and U denote normal, truncated (outside of the interval $(-1,1)$) normal, inverse gamma, and uniform distributions, respectively. The estimation sample ranges from 1959:Q1 to 2014:Q4.

2014:Q4.[12] The estimation results are provided in Table 4. The posterior median is 0.874 and the equal-tail 90% credible interval ranges from 0.698 to 0.966. These estimates can be compared to those obtained from quarterly consumption growth reported in Table 2 where the posterior median estimate of ρ is 0.921 (with stochastic volatility) and the upper bound of the credible interval is 0.963. Thus, while the median of ρ for GDP is smaller than for consumption, the 95% quantiles are in fact very similar.

So far, we have considered univariate models of consumption and income growth. Next, we examine the joint dynamics of these two series. In most models, consumption and income are cointegrated. We impose this cointegration relationship in the empirical analysis below. Specifically, the consumption dynamics are given by (4), while the log income-consumption ratio $yc_t \equiv y_t - c_t$ takes the form:

$$yc_{t+1} = \mu_{yc} + \phi_{yc} x_{t+1} + s_{t+1}, \quad s_{t+1} = \rho_s s_t + \sqrt{1-\rho_s^2}\sigma_{s,t}\eta_{s,t+1}, \quad \eta_{s,t+1} \sim N(0,1). \tag{6}$$

We assume that the log of stochastic volatility $\sigma_{s,t}$ follows an AR(1) process and adopt the measurement error model of Section 2.1 for consumption growth. For GDP, the measurement equation time-aggregates monthly growth rates $g_{y,t} = g_{c,t} + \Delta yc_t$ to average quarterly growth rates as in (5).

The estimated parameters for the cointegration model based on monthly consumption growth and quarterly GDP growth data are reported in Table 5. The posterior median estimate of ρ is 0.948 and the equal-tail probability 90% credible interval ranges from 0.913 to 0.970. Here, the

[12] We take log differences of the real GDP per capita series provided by FRED.

Table 5: Posterior Estimates: Consumption and Output

		Prior			Posterior		
	Distr.	5%	50%	95%	5%	50%	95%
				Consumption			
μ	N	-.007	.0016	.0100	.0012	.0016	.0020
ρ	U	-.9	0	.9	.913	.948	.970
φ_x	U	.05	.50	.95	.419	.593	.796
σ	IG	.0008	.0019	.0061	.0015	.0018	.0022
σ_ϵ	IG	.0008	.0019	.0061	.0013	.0015	.0018
ρ_{h_c}	N^T	.27	.80	.999	.949	.979	.996
$\sigma^2_{h_c}$	IG	.0013	.0043	.0283	.0024	.0091	.0235
				Output			
ρ_s	U	-.9	0	.9	.943	.965	.983
φ_s	U	5	50	95	6.65	8.82	13.25
ϕ_{yc}	U	-90	0	90	-1.83	-1.57	-1.18
ρ_{h_s}	N^T	.27	.80	.999	.943	.982	.996
$\sigma^2_{h_s}$	IG	.0013	.0043	.0283	.0015	.0039	.0187

Notes: The estimation sample ranges from 1959 to 2014. We report estimates of model (6). N, N^T, IG, and U denote normal, truncated (outside of the interval $(-1, 1)$) normal, inverse gamma, and uniform distributions, respectively.

strong evidence in monthly consumption in favor of a predictable component x_t seems to dominate the estimation result. There is no information in GDP growth that contradicts this information. The log-GDP consumption ratio itself is fairly persistent with median estimate of ρ_s of 0.965. Thus, deviations from the steady-state ratio are relatively long-lived.

How does our evidence relate to common views on GDP dynamics? U.S. GDP growth is well described as an AR(1) model with an autocorrelation coefficient of about 0.3. In our cointegration model the implied posterior predictive quantiles (0.05%, 0.50%, and 0.95%) for the autocorrelation of output growth at the quarterly frequency are 0.156, 0.273, and 0.389, which is consistent with this conventional wisdom. Thus, on balance, we view the dynamics of output and consumption to be consistent with our LRR specification with both series containing a small persistent component, and with models that imply a transmission from income to consumption.[13]

[13]It is well known that in production models consumption is not a Martingale sequence and the predictable

3.2 Dividends

As our subsequent asset pricing analysis focuses on the U.S. aggregate equity market, we now include dividend growth data in the estimation of the cash-flow model. We use monthly observations of dividends of the CRSP value-weighted portfolio of all stocks traded on the NYSE, AMEX, and NASDAQ. Dividend series are constructed on the per share basis as in Campbell and Shiller (1988b) and Hodrick (1992). Following Robert Shiller, we smooth out dividend series by aggregating three months values of the raw nominal dividend series.[14] We then compute real dividend growth as log difference of the adjusted nominal dividend series and subtract CPI inflation. Further details are provided in the Online Appendix.

Measurement Equation for Dividend Growth. Dividend data are available at monthly frequency for the estimation period from 1930 to 2014. There is a consensus in the finance literature that aggregate dividend series for a broad cross section of stocks exhibit a strong seasonality. This seasonality is generated by payout patterns which are not uniform over the calendar year. Much of this seasonality, in particular its deterministic component, can be removed by averaging observed dividend growth over the span of a year. To do so, we utilize the same "tent" function as for consumption growth in (1) and define

$$g^{a,o}_{d,t+1} = \sum_{j=1}^{23} \left(\frac{12 - |j - 12|}{12} \right) g^{o}_{d,t-j+2}, \quad g^{a}_{d,t+1} = \sum_{j=1}^{23} \left(\frac{12 - |j - 12|}{12} \right) g_{d,t-j+2}. \tag{7}$$

Our measurement equation then takes the form

$$g^{a,o}_{d,t+1} = g^{a}_{d,t+1} + \sigma^{a}_{d,\epsilon} \epsilon^{a}_{d,t+1}, \quad \epsilon^{a}_{d,t+1} \sim N(0,1). \tag{8}$$

For computational reasons that arise in the estimation of the asset pricing model in Section 4, we allow for some additional measurement errors, which we assume to be *iid* across periods. We fix these measurement errors at 1% of the sample variance of dividend growth rates. Note that (8) does not imply $g^{o}_{d,t+1} = g_{d,t+1}$, even for $\sigma^{a}_{d,\epsilon} = 0$. For instance, there could be a deterministic seasonal pattern in the observed monthly dividend growth data $g^{o}_{d,t+1}$ that is not part of the model-implied process $g_{d,t+1}$. The tent-shaped transformation would remove the seasonal component from observed data such that we are effectively equating the non-seasonal component of the observed data to the model-implied data.

component in consumption growth can be generated by a predictable component in productivity growth; see Croce (2014). In a frictionless environment both labor and capital income are determined by their respective marginal products, which in turn depend on the exogenous productivity process.

[14] We follow Shiller's approach despite the use of the annualization in (8) because we found that the annualization did not remove all the anomalies in the data.

State-Transition Equation. We will model consumption and dividend growth as a joint process. The law of motion for consumption growth is identical to (4), except for the fact that now we will have separate volatility processes for the persistent and transitory components. Dividend streams have levered exposures to both x_t and $\eta_{x,t+1}$, which is captured by the parameters ϕ and π, respectively. We allow $\sigma_{d,t}\eta_{d,t+1}$ to capture idiosyncratic movements in dividend streams. Overall, the cash-flow dynamics follow:

$$\begin{aligned} g_{c,t+1} &= \mu_c + x_t + \sigma_{c,t}\eta_{c,t+1} \\ x_{t+1} &= \rho x_t + \sqrt{1-\rho^2}\sigma_{x,t}\eta_{x,t+1} \\ g_{d,t+1} &= \mu_d + \phi x_t + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1}, \\ \sigma_{i,t} &= \varphi_i \sigma \exp(h_{i,t}), \quad h_{i,t+1} = \rho_{h_i} h_{i,t} + \sigma_{h_i} w_{i,t+1}, \quad i = \{c, x, d\} \end{aligned} \tag{9}$$

where the shocks are assumed to be $\eta_{i,t+1}, w_{i,t+1} \sim N(0,1)$, $i = \{c, x, d\}$ and we impose the normalization $\varphi_c = 1$. For now, we will also restrict $h_{x,t} = h_{c,t}$ and only report estimates for $\rho_{h,c}$ and $\sigma^2_{h,c}$.

Estimation Results. Table 6 provides percentiles of the prior distribution and the posterior distribution for the post 1959 estimation sample and for the mixed frequency sample starting in 1930. The priors for ϕ and π, parameters that determine the comovement of dividend and consumption growth, are uniform distributions on the interval $[-10, 10]$. The parameter φ_d determines the standard deviation of the *iid* component of dividend growth relative to consumption growth. Here we use a prior that is uniform on the interval [0, 10], thereby allowing for dividends to be much more volatile than consumption. The prior for the standard deviation of the dividend volatility process implies that the volatility may fluctuate either relatively little, over the range of 0.5 to 2.1 times the average volatility, or substantially, over the range of 0.1 to 13 times the average volatility. Finally, we fix the measurement error variance $(\sigma^a_{d,\epsilon})^2$ at 1% of the sample variance of dividend growth.

The most important finding is that the posterior median ρ increases as we add dividend growth data in the estimation. In addition, we find significant reduction in our uncertainty about ρ captured by the distance between 95% and 5% posterior quantiles. The posterior median of ρ is around 0.97 for the post 1959 sample and is 0.95 for the longer sample, both of which are higher than those in Table 3. The 5-95% distance dropped from 0.076 to 0.054 as we include dividend growth in the estimation (compare with Table 3). The posterior median of the standard deviation of the unconditional volatility of the persistent component φ_x is around 0.43, slightly lower than before.

The dividend leverage ratio on expected consumption growth ϕ is estimated to be around 2.8-3.5 and the standard deviation of the idiosyncratic dividend shocks φ_d is around 4.8. The estimation results also provide strong evidence for stochastic volatility. According to the posteriors reported

Table 6: Posterior Estimates: Consumption and Dividend Growth

		Prior			Posterior			Posterior		
								1930-1959		
					1959:M2-2014:M12			1960:M1-2014:M12		
	Distr.	5%	50%	95%	5%	50%	95%	5%	50%	95%
Consumption Growth Process										
ρ	U	-.90	0	.90	.937	.967	.991	.923	.952	.978
φ_x	U	.05	.50	.95	.285	.430	.834	.291	.430	.684
σ	IG	.0008	.0019	.0061	.0019	.0022	.0025	.0021	.0029	.0036
ρ_{h_c}	N^T	.27	.80	.999	.952	.985	.997	.976	.992	.998
$\sigma^2_{h_c}$	IG	.0013	.0043	.0283	.0015	.0053	.0185	.0013	.0034	.0132
Dividend Growth Process										
ϕ	U	-9.0	0.0	9.0	2.13	2.85	3.55	3.31	3.52	3.64
π	U	-9.0	0.0	9.0	.136	.358	.751	.642	.819	.932
φ_d	U	.50	5.0	9.5	3.51	4.69	6.16	3.31	4.82	7.66
ρ_{h_d}	N^T	.27	.80	.999	.939	.977	.994	.951	.977	.992
$\sigma^2_{h_d}$	IG	.015	.0445	.208	.0166	.0418	.1076	.0146	.0357	.0835
Consumption Measurement Error										
σ_ϵ	IG	.0008	.0019	.0062	.0009	.0011	.0014	.0009	.0012	.0015
σ^a_ϵ	IG	.0007	.0029	.0389	-	-	-	.0006	.0067	.0134

Notes: We utilize the mixed-frequency approach in the estimation: For consumption we use annual data from 1930 to 1959 and monthly data from 1960:M1 to 2014:M12; we use monthly dividend annual growth data from 1930:M1 to 2014:M12. For consumption we adopt the measurement error model of Section 2.1. We allow for annual consumption measurement errors ϵ^a_t during the periods from 1930 to 1948. We impose monthly measurement errors ϵ_t when we switch from annual to monthly consumption data from 1960:M1 to 2014:M12. We fix $\mu_c = 0.0016$ and $\mu_d = 0.0010$ at their sample averages. Moreover, we fix the measurement error variance $(\sigma^a_{d,\epsilon})^2$ at 1% of the sample variance of dividend growth. N, N^T, G, IG, and U denote normal, truncated (outside of the interval $(-1, 1)$) normal, gamma, inverse gamma, and uniform distributions, respectively.

in Table 6, both $\sigma_{c,t}$ and $\sigma_{d,t}$ exhibit significant time variation. The posterior medians of ρ_{h_c} and ρ_{h_d} range from 0.98 to 0.99.

Cointegration of Dividends and Consumption. In our analysis aggregate consumption is measured per capita and dividends are computed per share. Thus, there is no theoretical reason for the two series to be cointegrated. Nonetheless, we examined the presence of a cointegration

relationship between the observed series c_t^o and d_t^o. First, we conducted two frequentist cointegration tests based on post-1959 monthly data. The first test is an augmented Dickey-Fuller test that imposes the cointegration vector $[1, -1]$ and the second test is an Engle-Granger test based on an estimated cointegration vector of $[1, -0.55]$. None of these tests can reject the null hypothesis of no cointegration. Second, we estimate a modified state-space model with a hardwired cointegration restriction. This model retains the consumption growth dynamics of (9), but the law of motion of dividends is modified as follows:

$$dc_{t+1} = \mu_{dc} + \phi_{dc} x_{t+1} + s_{t+1}, \quad s_{t+1} = \rho_s s_t + \sqrt{1-\rho_s^2}\sigma_{s,t}\eta_{s,t+1}, \quad \eta_{s,t+1} \sim N(0,1), \tag{10}$$

where $dc_{t+1} \equiv d_{t+1} - c_{t+1}$ is the error correction representation for dividends and consumption. dc_{t+1} loads on x_{t+1}, and a stationary AR(1) process s_{t+1}, which has its own stochastic volatility process $\sigma_{s,t}$. Under this structure, it can be easily verified that dividend growth can be written as $g_{d,t+1} = \Delta dc_{t+1} + g_{c,t+1}$. The measurement equation for dividends then follows equation (8). A full set of estimates of the cointegration specification is reported in the Online Appendix. The estimate of ρ, as well as the other estimates of the consumption parameters are essentially unaffected by the cointegration specification. The marginal likelihood for the cointegration specification is 6,041.1 whereas the marginal likelihood for the original specification is 6,101.2. Thus, there is no evidence in the data in favor of the cointegration restriction.

3.3 Summary of Cash-Flow-Only Analysis

Aggregate consumption is a key macroeconomic variable, and it is therefore important for macroeconomists to understand its dynamic properties. There are several important implications that are robust across our analyses of consumption, consumption and output, and consumption and dividends. At the monthly frequency, consumption growth has a very strong MA(2) component. Ignoring this MA(2) component distorts inference. There is clear evidence against the hypothesis that consumption is a random walk at monthly frequency. Our interpretation of the MA(2) component is that it is generated by MA(1) measurement errors and a highly persistent "local level" component.[15] Empirically, our measurement error specification is preferred to the ARMA(1,2) specification. Thus, if the goal is to create a reduced-form model of consumption, it is important to capture the MA component. If the goal is to confront a macro model with monthly consumption data, it is important to apply a "filter" that removes the high-frequency movements in consumption

[15]This interpretation is consistent with studies that examine the quality of consumption data, e.g., Wilcox (1992), but from a pure time series perspective, we cannot rule out that the MA(2) component is partly due to transient dynamics in "true" consumption growth.

growth that we attribute to measurement error, because a typical macro model is not equipped to capture these dynamics. Overall, the posterior interval for the parameter estimates essentially encompass those used in the LRR literature (e.g Bansal, Kiku, and Yaron (2012)). Importantly, the various estimation results (univariate consumption, consumption and GDP, and consumption and dividends) provide supportive evidence for a small persistent component in both consumption growth rate and its stochastic volatility. Consistent with previous LRR work, this evidence is distinctly different from a commonly held view in which consumption growth is an *iid* process.

4 The Long-Run Risks Model

We now embed the cash flow process (9) into an endowment economy, which allows us to price financial assets. The preferences of the representative household are described in Section 4.1. Section 4.2 describes the model solution. Section 4.3 presents the state-space representation of the asset-pricing model and its Bayesian estimation.

4.1 Representative Agent's Optimization

We consider an endowment economy with a representative agent that has Epstein and Zin (1989) recursive preferences and maximizes her lifetime utility,

$$V_t = \max_{C_t} \left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta\big(\mathbb{E}_t[V_{t+1}^{1-\gamma}]\big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}},$$

subject to budget constraint

$$W_{t+1} = (W_t - C_t)R_{c,t+1},$$

where W_t is the wealth of the agent, $R_{c,t+1}$ is the return on all invested wealth, γ is risk aversion, $\theta = \frac{1-\gamma}{1-1/\psi}$, and ψ is intertemporal elasticity of substitution. As highlighted in Albuquerque, Eichenbaum, Luo, and Rebelo (2016), we also allow for a preference shock, λ_t, to the time rate of preference. The endowment stream is given by the law of motion for consumption and dividend growth in (9), and the growth rate of the preference shock, denoted by $x_{\lambda,t}$, follows an AR(1) process with shocks that are independent of the shocks to cash flows:

$$x_{\lambda,t+1} = \rho_\lambda x_{\lambda,t} + \sigma_\lambda \eta_{\lambda,t+1}, \quad \eta_{\lambda,t+1} \sim N(0,1). \tag{11}$$

The Euler equation for any asset $r_{i,t+1}$ takes the form

$$\mathbb{E}_t\left[\exp\left(m_{t+1} + r_{i,t+1}\right)\right] = 1, \tag{12}$$

where $m_{t+1} = \theta \log \delta + \theta x_{\lambda,t+1} - \frac{\theta}{\psi} g_{c,t+1} + (\theta - 1) r_{c,t+1}$ is the log of the real stochastic discount factor (SDF), and $r_{c,t+1}$ is the log return on the consumption claim. We reserve $r_{m,t+1}$ for the log market return – the return on a claim to the market dividend cash flows.[16]

4.2 Solution

Our goal is to devise a solution method that strikes a balance between accuracy and computational time. The solution described subsequently meets this requirement: it can be computed quickly because it relies on analytical approximations; it leads to a conditionally linear state-space representation for which the likelihood function can be efficiently evaluated with a particle filter (see below); and it is accurate for the empirically relevant parameter values.

Conditional on the cash-flow dynamics in (9) and the Euler equation (12), we have to derive the asset prices for the model economy. In order to fit the cash-flow specification to consumption and dividend growth data, we assumed that the volatilities follow log Gaussian processes: $\sigma_{i,t} = \varphi_i \sigma \exp(h_{i,t})$, where $h_{i,t}$ is a linear autoregressive process with normally-distributed innovations. This specification has been empirically successful in capturing conditional heteroskedasticity in a broad set of financial and macroeconomic time series.

The advantage of the exponential transformation is that it ensures that volatilities are non-negative. The disadvantage is that under this specification the expected value of the level of consumption and dividends is infinite, which creates problems with the existence of continuation values in the endowment economy. This issue has been pointed out, for instance, in Chernov, Gallant, Ghysels, and Tauchen (2003) and Andreasen (2010), who proposed to splice the exponential transformation of $h_{i,t}$ together with a non-exponential function, e.g., a square-root function, for $h_{i,t}$ exceeding some large threshold $\bar{h}_i$. To obtain a solution for the asset prices we proceed slightly differently, by taking a linear approximation of the exponential transformation $\sigma_{i,t}^2 = (\varphi_i \sigma)^2 \exp(2h_{i,t})$ around the steady state $h_{i,*} = 0$ and replacing the innovation variances in (9) with a process that follows Gaussian dynamics:

$$\sigma_{i,t+1}^2 \quad = \quad (\varphi_i \sigma)^2 (1 - \rho_{h_i}) + \rho_{h_i} \sigma_{i,t}^2 + \sigma_{w_i} w_{i,t+1}, \quad i = \{c, x, d\}. \tag{13}$$

[16] Formally, markets are complete in the sense that all income and assets are tradable and add up to total wealth for which the return is $R_{c,t}$. In particular, let $R_{j,t+1} = (d_{j,t+1} + p_{j,t+1})/p_{j,t}$ be the return to a claim that pays the dividend stream $\{d_{j,\tau}\}_{\tau=t}^{\infty}$ and has the price $p_{j,t}$. Let $q_{j,t}$ be the number of shares. Then $W_t - C_t = \sum_j p_{j,t} q_{j,t}$. Wealth next period, W_{t+1}, equals $\sum_j p_{j,t} q_{j,t} R_{j,t+1}$, and it follows that $R_{c,t+1} = \frac{\sum_j p_{j,t} q_{j,t} R_{j,t+1}}{\sum_j p_{j,t} q_{j,t}}$. As in Lucas (1978), we normalize all shares $q_{j,t}$ to one and the risk free asset to be in zero net supply such that in equilibrium $C_t = D_m + D_o$, where D_m are the dividends to all tradable financial assets and D_o are dividends on all other assets (e.g., labor, housing etc.). R_m, the return we utilize in our empirical work, is the return on the claims that pay dividends D_m.

Just as the above-mentioned approaches of modifying the exponential transformation, the linear approximation effectively generates thinner tails for the variance processes and facilitates valuations to be finite.[17] After this linearization, the volatility process is identical to the one used in Bansal and Yaron (2004) and the subsequent work that builds on their paper.

Asset prices can now be derived by using the approximate analytical solution described in Bansal, Kiku, and Yaron (2012) which utilizes the Campbell and Shiller (1988a) log-linear approximation for returns. This solution serves our purpose well, because it can be computed very quickly, which facilitates the Bayesian estimation below. The log price-consumption ratio takes the form

$$pc_t = A_0 + A_1 x_t + A_{1,\lambda} x_{\lambda,t} + A_{2,c}\sigma^2_{c,t} + A_{2,x}\sigma^2_{x,t}. \tag{14}$$

As discussed in Bansal and Yaron (2004), $A_1 = \frac{1-\frac{1}{\psi}}{1-\kappa_1\rho}$, the elasticity of prices with respect to growth prospects, will be positive whenever the IES, ψ, is greater than 1. $A_{1,\lambda} = \frac{\rho_\lambda}{1-\kappa_1\rho_\lambda}$, the elasticity of prices with respect to the growth rate of the preference shock, is always positive. Further, the elasticity of pc_t with respect to the two volatility processes $\sigma^2_{c,t}$ and $\sigma^2_{x,t}$ is $\frac{\theta}{2}\frac{(1-\frac{1}{\psi})^2}{1-\kappa_1\nu_c}$ and $\frac{\theta}{2}\frac{(\kappa_1 A_1)^2}{1-\kappa_1\nu_x}$ respectively; both will be negative — namely, prices will decline with uncertainty — whenever θ is negative. A condition that guarantees a negative θ is that $\gamma > 1$, and $\psi > 1$ – a configuration relevant for our parameter estimates and one in which agents exhibit preference for early resolution of uncertainty. The innovation to the log stochastic discount factor (SDF) are linear in the shocks to consumption growth η_c, η_x, the preference shocks, η_λ, and the shocks to volatilities w_c, and w_x. Denoting λs as their respective market prices of risk, it is instructive to note that $\lambda_c = \gamma$, $\lambda_x = \frac{(\gamma-\frac{1}{\psi})\kappa_1}{1-\kappa_1\rho}$, $\lambda_\lambda = -\frac{\theta-\kappa_1\rho_\lambda}{1-\kappa_1\rho_\lambda}$ (and λ_{w_c} and λ_{w_x}) are positive (negative) whenever $\gamma > 1$ and $\psi > 1$. Furthermore, when preferences are time separable, namely, when $\theta = 1$, λ_x, λ_{w_x}, and λ_{w_c} are all zero.

Risk premia are determined by the negative covariation between the innovations to returns and the innovations to the SDF. It can be shown that the risk premium for the market return, $r_{m,t+1}$, is

$$\begin{aligned}\mathbb{E}_t(r_{m,t+1} - r_{f,t}) + \frac{1}{2}var_t(r_{m,t+1}) &= -cov_t(m_{t+1}, r_{m,t+1}) \\ = \underbrace{\beta_{m,c}\lambda_c\sigma^2_{c,t}}_{\text{short-run risk}} + \underbrace{\beta_{m,x}\lambda_x\sigma^2_{x,t}}_{\text{long-run growth risk}} &+ \underbrace{\beta_{m,\lambda}\lambda_\lambda\sigma^2_\lambda}_{\text{preference risk}} + \underbrace{\beta_{m,w_x}\lambda_{w_x}\sigma^2_{w_x} + \beta_{m,w_c}\lambda_{w_c}\sigma^2_{w_c}}_{\text{volatility risks}},\end{aligned} \tag{15}$$

where the βs reflect the exposures of the market return to the underlying consumption risks. Equation (15) highlights that the conditional equity premium can be attributed to (i) short-run consumption growth, (ii) long-run growth, (iii) preference shock, (iv) short-run and long-run volatility risks.

[17] A quantitative comparison among the various approaches of thinning the tails of the $\sigma^2_{i,t}$ processes is beyond the scope of this paper.

A key variable for identifying the model parameters is the risk-free rate. Under the assumed dynamics in (9), the risk-free rate is affine in the state variables and follows

$$r_{f,t} = B_0 + B_1 x_t + B_{1,\lambda} x_{\lambda,t} + B_{2,c}\sigma^2_{c,t} + B_{2,x}\sigma^2_{x,t}. \tag{16}$$

It can be shown that $B_1 = \frac{1}{\psi} > 0$ and the risk-free rate rises with good economic prospects, while $B_{1,\lambda} = -\rho_\lambda < 0$ and the risk-free rate falls with positive preference shock. Under $\psi > 1$ and $\gamma > 1$, $B_{2,c}$ and $B_{2,x}$ are negative so the risk-free rate declines with a rise in economic uncertainty. Further details of the solution are provided in the Online Appendix.

The accuracy of the log-linearization depends on the parameterization of the LRR model. Taking the linear volatility process in (13) as given, Pohl, Schmedders, and Wilms (2016) compared the quantitative implications of our model solution to that of a nonlinear solution obtained by a projection method. They find that discrepancies for key asset pricing moments between the solutions are small (less than 6%) conditional on a parameterization that is similar to the posterior medians reported in Table 7 below. However, the discrepancies become larger if the persistence parameters ρ and ρ_h are increased and pushed toward the upper bound of credible sets derived from marginal posterior distributions. Thus, strictly speaking, the parameter estimates that we are reporting below should be interpreted as parameter estimates for the approximation.[18]

4.3 State-Space Representation and Bayesian Inference

While the state-space models for the cash-flow dynamics analyzed in Sections 2 and 3 can be analyzed with a fairly straightforward Metropolis-within-Gibbs sampler, posterior computations for the model with asset returns are considerably more complicated because the stochastic volatility process $h_t = [h_{c,t}, h_{x,t}, h_{d,t}]'$ affects the conditional mean of the asset prices.

The measurement equation can be expressed as

$$y_{t+1} = A_{t+1}\big(D + Z s_{t+1} + Z^v s^v_{t+1}(h_{t+1}, h_t) + \Sigma^u u_{t+1}\big), \quad u_{t+1} \sim N(0, I). \tag{17}$$

The vector of observables y_t comprises consumption and dividend growth, the observed market return $r^o_{m,t}$ and the risk-free rate $r^o_{f,t}$. u_{t+1} is a vector of measurement errors and A_{t+1} is a

[18] A similar issue arises in the literature on dynamic stochastic general equilibrium (DSGE) models: the vast majority of DSGE models are estimated based on log-linear approximations, which facilitate a speedy evaluation of the likelihood function with the Kalman filter. The caveat that strictly speaking the resulting parameter estimates are estimates for the log-linear approximation has been widely accepted in this literature. An immediate consequence is that one should not plug parameter estimates obtained from a log-linear approximation into the nonlinear version of the model, becase it will lead to a mismatch between model-implied moments and the moments in the data; see An and Schorfheide (2007) for more details.

selection matrix that accounts for deterministic changes in the data availability. $s^v_{t+1}(\cdot)$ is a vector of conditional variances that depend on the log volatilities of the cash flows, h_{t+1} and h_t. The remaining "linear" state variables are collected in the vector s_{t+1}, which essentially consists of the persistent cash flow component x_t (see (9)) and the preference shock $x_{\lambda,t}$. However, in order to express the observables y_{t+1} as a linear function of s_{t+1}, to capture the elaborate measurement error model of consumption, and to account for potentially missing observations, it is necessary to augment s_{t+1} by lags of x_t and $x_{\lambda,t}$ as well as the innovations for the cash-flow process and measurement errors. This leads to a high dimensional state vector s_t (see Online Appendix for a precise definition).

The solution of the LRR model sketched in Section 4.2 provides the link between the state variables and the observables y_{t+1}. The state variables themselves follow vector autoregressive processes of the form

$$s_{t+1} = \Phi s_t + v_{t+1}(h_t), \quad h_{t+1} = \Psi h_t + \Sigma_h w_{t+1}, \quad w_{t+1} \sim N(0, I), \tag{18}$$

where $v_{t+1}(\cdot)$ is an innovation process with a variance that is a function of the log volatility process h_t and w_{t+1} is the innovation of the stochastic volatility process.

The key difference between the nonlinear state-space model given by (17) and (18) and the state-space models estimated in Sections 2 and 3 is that the volatility states (h_{t+1}, h_t) enter the conditional mean of y_{t+1} through the model-implied asset returns. This means that the Metropolis-within-Gibbs sampler that we used previously is not valid for the model with asset prices. Instead, we will use a particle filter to approximate the likelihood function of the state-space model and then embed the likelihood approximation into a Metropolis-Hastings algorithm.

Our particle filter exploits the particular structure of the state-space model. Conditional on the volatility states (h_{t+1}, h_t), the model is linear. Building on ideas in Chen and Liu (2000), we use Kalman filtering steps to track the Gaussian distribution of $s_t|(h^j_t, Y_{1:t})$, where $\{h^j_t, W^j_t\}_{j=1}^M$ is a set of particle values and weights for the volatility states. Because conditional on the 3-dimensional volatility vector h^j_t one can integrate over the high-dimensional vector s_t analytically (Rao-Blackwellization), the particle filter approximation $\hat{p}(Y|\Theta)$ of the likelihood function tends to be sufficiently accurate so that it can be embedded into a random walk Metropolis-Hastings algorithm. Here Θ comprises the parameters of the cash-flow process, the volatility parameters, and the preference parameters of the representative household. The resulting sampler belongs to the class of particle MCMC samplers. Andrieu, Doucet, and Holenstein (2010) have shown that the use of $\hat{p}(Y|\Theta)$ in MCMC algorithms can still deliver draws from the actual posterior $p(\Theta|Y)$ because these approximation errors essentially average out as the Markov chain progresses. Further details of the posterior sampler are provided in the Online Appendix.

5 Empirical Results Based on the Long-Run Risks Model

We now turn to the empirical analysis based on the LRR model. Section 5.1 describes the asset price data that are used in addition to the cash-flow data. We discuss the estimation results in Section 5.2 and present the asset pricing implications of the estimated model in Section 5.3.

5.1 Data

In addition to the consumption and dividend data used in Sections 2 and 3 we now also use financial market data from 1930:M1 to 2014:M12. This includes monthly observations of returns and prices of the CRSP value-weighted portfolio of all stocks traded on the NYSE, AMEX, and NASDAQ. Prices are also constructed on the per share basis as in Campbell and Shiller (1988b) and Hodrick (1992). The stock market data are converted to real using the consumer price index (CPI) from the Bureau of Labor Statistics. Finally, the ex-ante real risk-free rate is constructed as a fitted value from a projection of the ex-post real rate on the current nominal yield and inflation over the previous year. To run the predictive regression, we use monthly observations on the three-month nominal yield from the CRSP Fama Risk Free Rate tapes and CPI series. Data sources and summary statistics are available in the Online Appendix.

5.2 Model Estimation

Parameter Estimates. The prior distribution for the parameters associated with the exogenous cash flow process are the same as the ones used in Section 3.2. Thus, we focus on the preference parameters that affect the asset pricing implications of the model. Percentiles for the prior are reported in the left-side columns of Table 7. The prior for the discount rate δ reflects beliefs about the magnitude of the risk-free rate. For the asset pricing implications of our model, it is important whether the IES is below or above 1. Thus, we choose a prior that covers the range from 0.3 to 3.5. The 90% prior credible interval for the risk-aversion parameter γ ranges from 3 to 15, encompassing the values that are regarded reasonable in the asset pricing literature. The prior for the persistence and the innovation standard deviation of the preference shock is identical to the prior for the cash flow parameters ρ and σ. Finally, we fix the variance $\sigma^2_{f,\epsilon}$ of the measurement error of the risk free rate at 1% of the risk-free rate's sample variance.

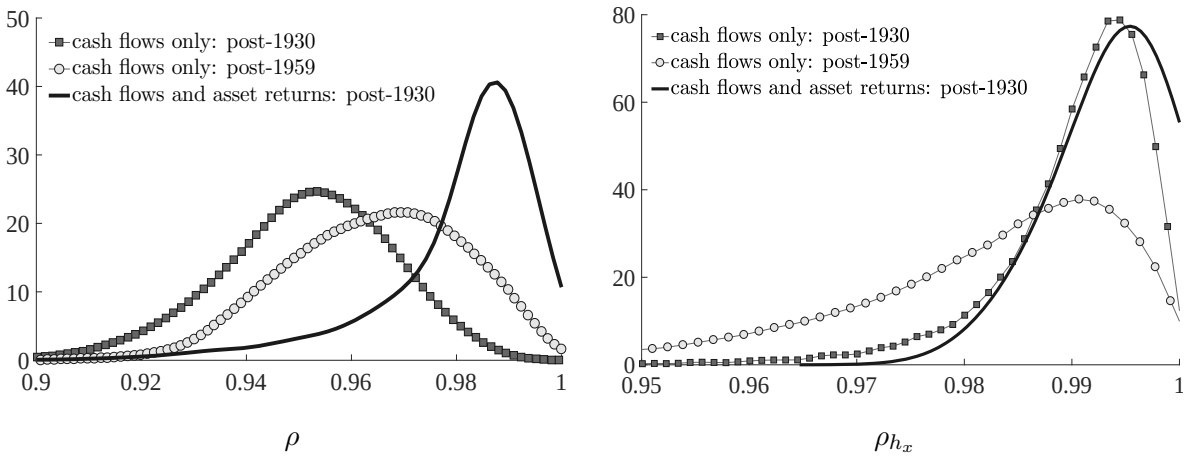
The remaining columns of Table 7 summarize the percentiles of the posterior distribution for the model parameters. While the estimated cash flow parameters are, by and large, similar to those reported in Table 6 when asset prices are not utilized, a few noteworthy differences emerge. First,

Table 7: Prior and Posterior Estimates

		Prior			Posterior		
	Distr.	5%	50%	95%	5%	50%	95%
Household Preferences							
ψ	G	.30	1.30	3.45	1.25	1.97	3.22
γ	G	2.75	7.34	15.46	5.44	8.89	14.44
Preference Risk							
ρ_λ	U	-.90	0	.90	.933	.959	.974
σ^2_λ	IG	.0003	.0005	.0015	.0003	.0004	.0005
Consumption Growth Process							
ρ	U	-.90	0	.90	.949	.987	.997
φ_x	U	.05	.50	.95	.120	.215	.382
σ	IG	.0008	.0019	.0061	.0027	.0035	.0042
ρ_{h_c}	N^T	.27	.80	.999	.977	.991	.998
$\sigma^2_{h_c}$	IG	.0011	.0043	.0283	.0075	.0096	.0109
ρ_{h_x}	N^T	.27	.80	.999	.982	.992	.998
$\sigma^2_{h_x}$	IG	.0011	.0043	.0283	.0022	.0039	.0044
Dividend Growth Process							
ϕ	N	-9.0	0.0	9.0	2.14	3.65	6.43
π	N	-9.0	0.0	9.0	.75	1.47	2.37
φ_d	U	.50	5.0	9.5	3.19	4.54	6.55
ρ_{h_d}	N^T	.28	.80	.999	.943	.969	.974
$\sigma^2_{h_d}$	IG	.015	.0445	.208	.0404	.0447	.0565
Consumption Measurement Error							
σ_ϵ	IG	.0008	.0019	.0062	.0009	.0014	.0020
σ^a_ϵ	IG	.0007	.0029	.0389	.0038	.0141	.0213

Notes: The estimation results are based on annual consumption growth data from 1930 to 1960 and monthly consumption growth data from 1960:M1 to 2014:M12. We allow for annual consumption measurement errors ϵ^a_t during the periods from 1930 to 1948. We impose monthly measurement errors ϵ_t when we switch from annual to monthly consumption data from 1960:M1 to 2014:M12. For the other three series we use monthly data from 1930:M1 to 2014:M12. We fix $\delta = 0.999$. We fix $\mu_c = 0.0016$ and $\mu_d = 0.0010$ at their sample averages. We also fix the measurement error variances $(\sigma^a_{d,\epsilon})^2$ and $(\sigma_{f,\epsilon})^2$ at 1% of the sample variance of dividend growth and the risk-free rate, respectively. B, N, N^T, G, and IG are beta, normal, truncated (outside of the interval $(-1, 1)$) normal, gamma, and inverse gamma distributions, respectively.

Figure 3: Posterior Distribution of ρ and ρ_{h_x}



Notes: We plot posterior densities of ρ from the estimation with cash flow data only from post-1930 (squared-line) and from post-1959 samples (circled-line), respectively, and from the estimation with cash flow and asset return data from post-1930 sample (solid-line).

the estimate of ρ, the persistence of the predictable cash flow component, increases from 0.952 to 0.987 to better capture the equity premium and persistence of the price-dividend ratio. The left panel of Figure 3 overlays the posterior densities of ρ obtained with (post-1930 sample) and without asset prices (post-1930 and post-1959 samples, respectively).[19] Interestingly, the figure shows that although the mode of the posterior increases and shifts to the right when asset prices are used in estimation, the 90% credible interval ranging from 0.949 to 0.997 contains the posterior medians of ρ from the cash-flow-only estimations.[20]

Second, the right panel of Figure 3 shows the posterior distribution of ρ_{h_x}, the persistence of the stochastic volatility process for x_t. The modes of the three posteriors are quite similar, with the cashflow-only posteriors having a longer left tail. Again, the posterior becomes more concentrated as asset returns are added to the estimation. Third, in the cash-flow-only estimation, we imposed a common stochastic log volatility process for the transitory and persistent component of consumption growth, i.e., $h_{x,t} = h_{c,t}$, which lead to an estimate $\hat{\sigma}^2_{h_c} = 0.0034$. Once we add the returns to the set

[19]Results from the post-1959 sample with asset prices are virtually identical to the results from the post-1930 sample. For this reason, they are not plotted separately in Figure 3.

[20]In the Online Appendix we present additional misspecification tests for the consumption dynamics. To assess the extent to which the increase in ρ leads to a decrease in fit of the consumption growth process, we re-estimate model (4) conditional on various choices of ρ between 0.90 and 0.99 and re-compute the marginal data density for consumption growth. The key finding is that the drop in the marginal data density by changing ρ from $\hat{\rho}$ to 0.99 is small, indicating that there essentially is no tension between the parameter estimates obtained with and without asset prices.

of observables and remove the restriction, we obtain $\hat{\sigma}^2_{h_x} = 0.0039$ and $\hat{\sigma}^2_{h_c} = 0.0096$, reflecting asset price information about the volatility of volatilities. Fourth, the estimate of φ_x drops from 0.430 to 0.215, which reduces the model-implied predictability of consumption growth by the price-dividend ratio and brings it more in line with the data. Finally, the estimate of σ increases somewhat from .0029 to .0035 to explain the highly volatile asset prices data.

Overall, the information from the market returns and risk-free rate reduces the posterior uncertainty about the cash flow parameters and strengthens the evidence in favor of a time-varying conditional mean of cash flow growth rates as well as time variation in the volatility components. Table 7 also provides the estimated preference parameters. Importantly, the IES is estimated above 1 with a credible interval ranging from 1.3 to 3.2, while the posterior median estimate of the risk aversion parameter γ is 8.9 and its interval estimate is 5.4 to 14.4.

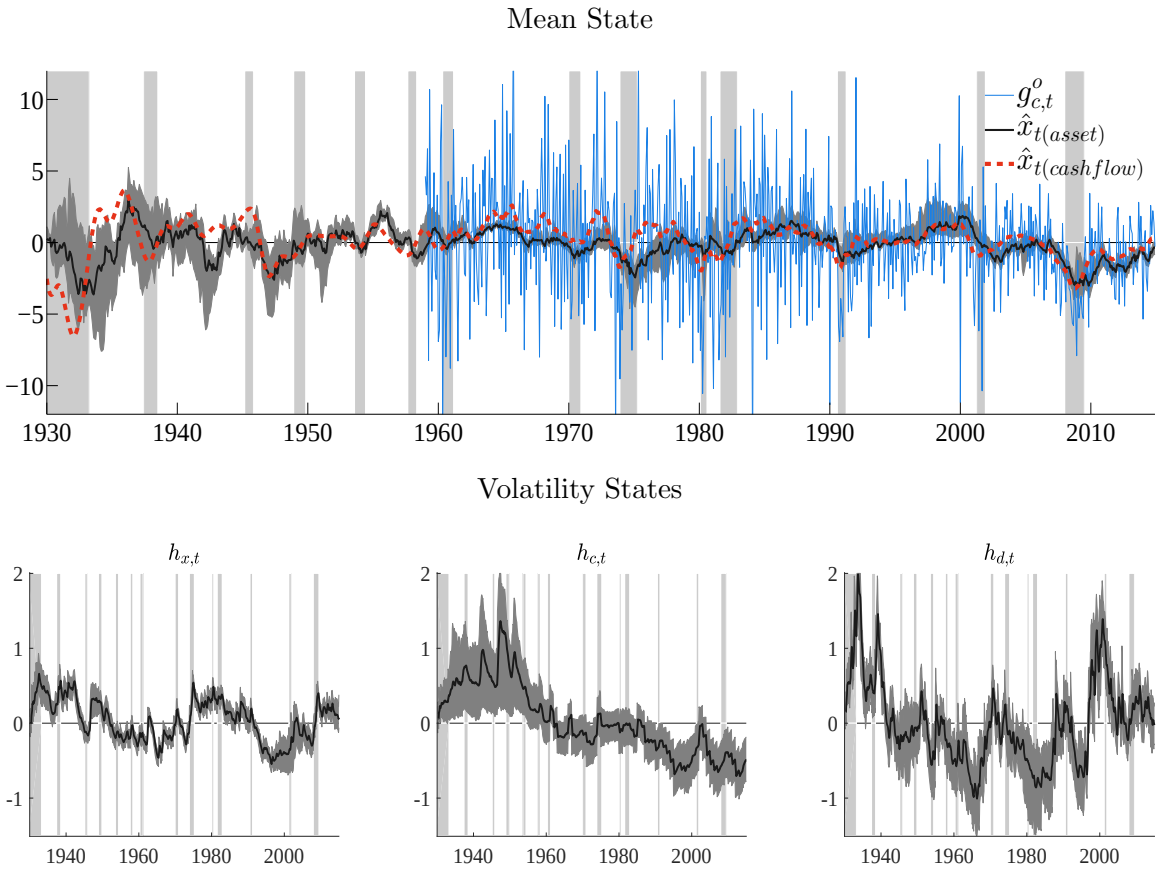
Smoothed Mean and Volatility States. Figure 4 depicts smoothed estimates of the predictable growth component x_t. Because the estimate of x_t is, to a large extent, determined by the time path of consumption, the 90% credible bands (reflecting uncertainty about parameters and the latent states) are much wider prior to 1960, when only annual consumption growth data were used in the estimation. Post 1959, x_t tends to fall in recessions (indicated by the shaded bars in Figure 4), but periods of falling x_t also occur during expansions. We overlay the smoothed estimate of x_t obtained from the estimation without asset price data. It is very important to note that the two estimates are similar, which highlights that x_t is, in fact, detectable based on cash flow data only. We also depict the monthly consumption growth data post 1959, which confirms that x_t indeed captures low-frequency movements in consumption growth.

The smoothed volatility processes are plotted in Figure 4. Recall that our model has three independent volatility processes, $h_{c,t}$, $h_{d,t}$, and $h_{x,t}$, associated with the innovations to consumption growth, dividend growth, and the predictable component, respectively. The most notable feature of $h_{c,t}$ is that it captures a drop in consumption growth volatility that occurred between 1940 and 1960. In magnitude, this drop in volatility is much larger than a subsequent decrease around 1984, the year typically associated with the Great Moderation. The stochastic volatility process for dividend growth $h_{d,t}$ seems to exhibit more medium- and high-frequency movements than $h_{c,t}$. Finally, the volatility of the persistent component, $h_{x,t}$, exhibits substantial fluctuations over our sample period, and it tends to peak during NBER recessions.

5.3 Asset Pricing Implications

Risk-Free Rate Estimate and Preference Shock. Figure 5 overlays the actual risk-free rate, which is assumed to be subject to measurement errors, and the smoothed "true" or model-implied

Figure 4: Smoothed States

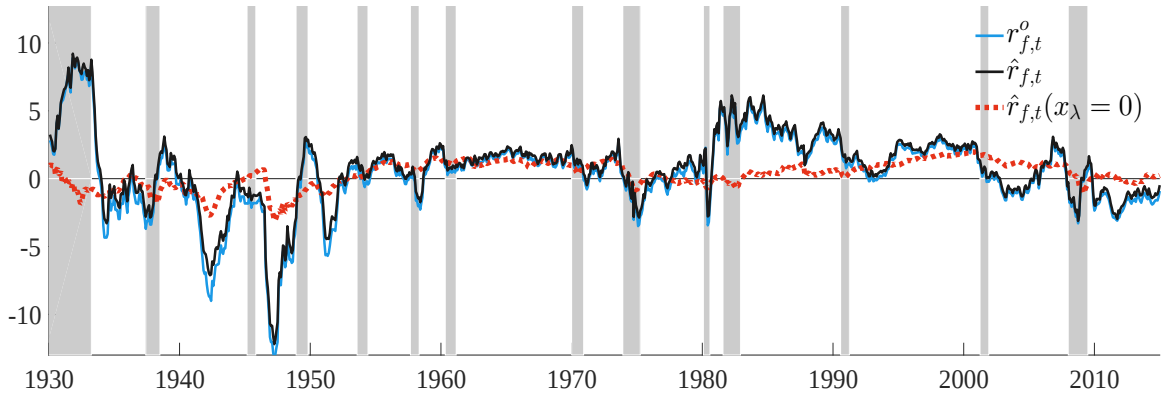


Notes: Black lines represent posterior medians of smoothed states and gray-shaded areas correspond to 90% credible intervals. Shaded bars indicate NBER recession dates. In the top panel, we overlay the smoothed state x_t obtained from the estimation without asset prices (red dashed line) and monthly consumption growth data (blue solid line).

risk-free rate. We find that the measurement errors are fairly small. To highlight the importance of the preference shock, we also plot a counterfactual risk-free rate that would prevail in the absence of $x_{\lambda,t}$. It turns out that *ex-post* much of the risk-free rate fluctuations are explained by the preference shock. In the absence of the preference shock, the process for the expected stochastic discount factor implied by the predictable component of cash flow growth and the stochastic volatilities is too smooth relative to the observed risk-free rate. The preference shock can generate additional fluctuations in the expected discount factor without having a significant impact on asset returns (as we will see below).

We assumed that the preference shock is independent of cash flows. In a production economy this assumption will typically not be satisfied. Stochastic fluctuations in the discount factor generate fluctuations in consumption and investment, which in turn affect cash flows. We assess the inde-

Figure 5: Model-Implied Risk-Free Rate



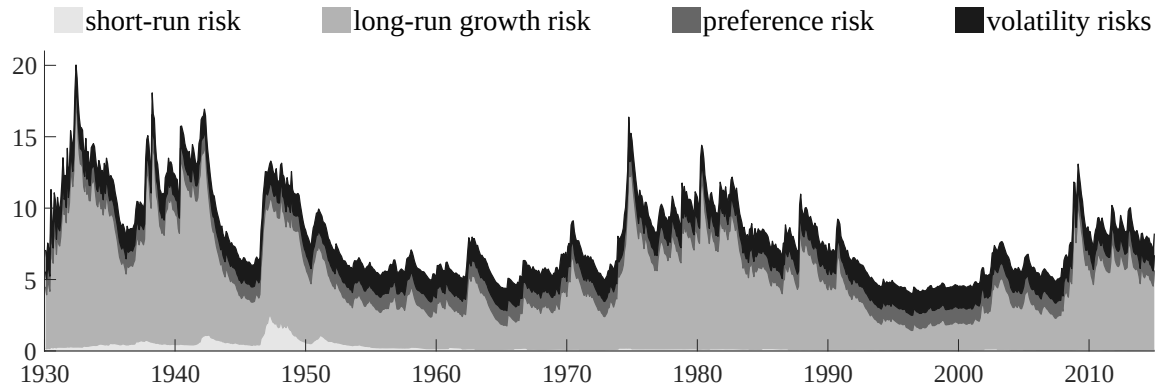
Notes: Blue line depicts the actual risk-free rate, and black line depicts the smoothed, model-implied risk-free rate without measurement errors. Red dashed line depicts the model-implied risk-free rate with $x_{\lambda,t} = 0$. The parameters are fixed at their posterior median estimates.

pendence as follows. First, we compute the *ex-post* correlation between the smoothed preference shock innovations $\eta_{\lambda,t}$ and the cash-flow innovations $\eta_{c,t}$ and $\eta_{x,t}$. We can do so for every parameter draw Θ^s from the posterior distribution. The 90% posterior predictive intervals range from -0.09 to 0.03 for the correlation between $\eta_{\lambda,t}$ and $\eta_{c,t}$ and from 0 to 0.2 for the correlation between $\eta_{\lambda,t}$ and $\eta_{x,t}$. Second, we re-estimate our model under the assumption that $\eta_{\lambda,t}$ and $\eta_{x,t}$ are negatively correlated. The resulting parameter estimates as well as the asset pricing moments are essentially unaltered. According to a marginal data density comparison the more parsimonious specification in which preference shocks and cash flows are independent is preferred. Based on these results we conclude that there is no evidence that contradicts the independence assumption.

Determinants of the Equity Risk Premium. Figure 6 depicts the contribution of short-run risk, $\sigma^2_{c,t}$, the long-run growth risk, $\sigma^2_{x,t}$, the preference risk, σ^2_λ, and the volatility risks, $\sigma^2_{w_c}$ and $\sigma^2_{w_x}$, to the risk premium at the posterior median parameter estimates; see (15). We compute βs and λs based on the median posterior parameter estimates and multiply them by the median volatility state estimates to construct the risk premium. The total (annualized) equity risk premium is around 8.2%.[21] The two major sources of the risk premium are the long-run growth risk and the volatility risks and when combined they account for 83% of the risk premium. More specifically, the 8.2% equity premium can be decomposed as follows. On average, the long-run growth risk

[21]The gross equity premium is $\mathbb{E}[r_{m,t+1} - r_{f,t}] + 1/2\sigma^2_{r_m} \approx 0.0615 + 0.5 * 0.226^2 - 0.0047 = 8.2\%$. If we were to attribute the moving-average fluctuations in observed monthly consumption growth to "true" consumption growth instead of measurement errors, the asset pricing implications of the model would essentially remain unchanged. The equity premimum would rise by approximately 0.06%.

Figure 6: Decomposition of the Equity Risk Premium



Notes: We provide the decomposition of the risk premium (15). We compute βs and λs based on the median posterior parameter estimates and multiply with the median volatility state estimates $\hat{\sigma}^2_{c,t}$ and $\hat{\sigma}^2_{x,t}$ to construct the model-implied risk premium. On average, the risk premium is accounted for by the short-run risk (0.3%), long-run growth risk (5.0%), preference risk (1.1%), and volatility risk (1.8%), respectively. The total in-sample market risk premium (annualized) is around 8.2%.

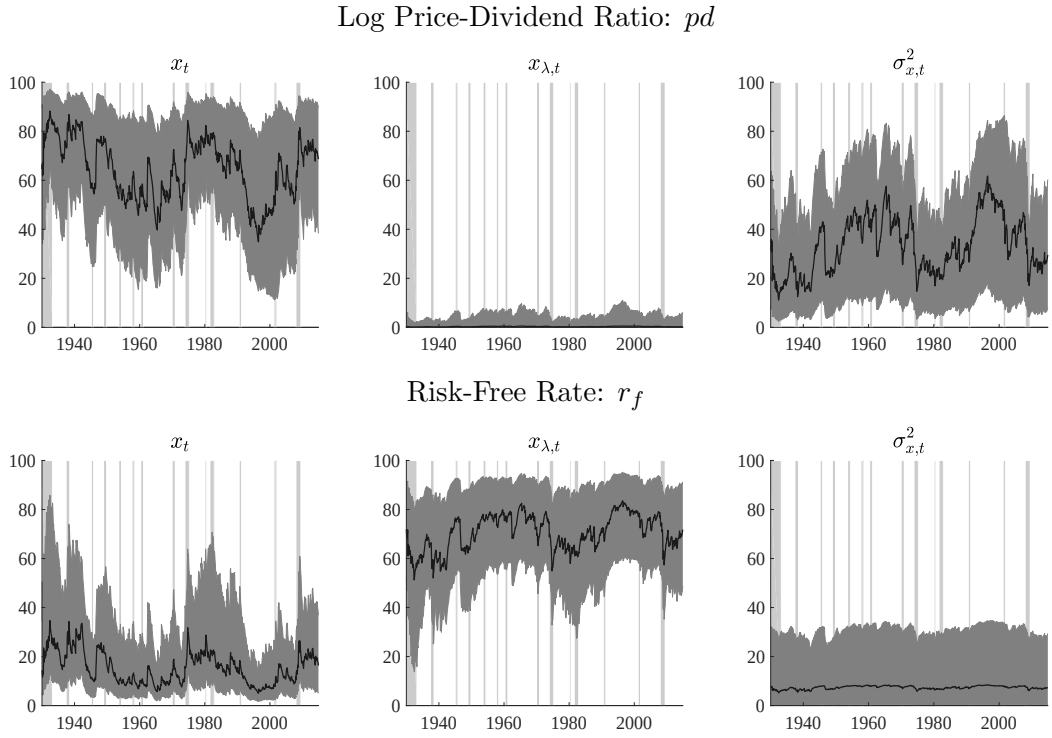
generates a premium of 5.0%, the volatility risks account for 1.8%, the preference shock generates 1.1%, and the short-run volatility risk contributes 0.3%.

Determinants of Asset Price Volatility. Figure 7 depicts the time-varying contribution of the fluctuations in growth prospects, x_t, the preference shock, $x_{\lambda,t}$, and the conditional variability of growth prospects, $\sigma_{x,t}$, to the volatility of the price-dividend ratio and the risk-free rate.[22] We generate counterfactual volatilities by shutting down the estimated x_t, $x_{\lambda,t}$, and $\sigma_{x,t}$ processes, respectively. The ratios of the counterfactual and the actual asset price volatilities measure the contribution of the non-omitted risk factors. We subtract this ratio from 1 to obtain the relative contribution of the omitted risk factor shown in Figure 7. The credible bands reflect parameter uncertainty and uncertainty about the latent states. While the preference shocks are important for the risk-free rate, they contribute very little to the variance of the price-dividend ratio. Most of the variability of the price-dividend ratio is, in equal parts, due to the variation in x_t and $\sigma_{x,t}$. The remaining risk factors $\sigma^2_{c,t}$ and $\sigma^2_{d,t}$ have negligible effects (less than 1% on average) on the asset price volatilities, but are important for tracking the consumption and dividend growth data.

Matching Asset Price Moments. While asset pricing moments implicitly enter the likelihood function of our state-space model, it is instructive to examine the extent to which sample moments implied by the estimated state-space model mimic the sample moments computed from our actual

[22]The decomposition of market return volatility (not shown in Figure 7) is qualitatively similar to that of the price-dividend ratio.

Figure 7: Variance Decomposition for Market Returns and Risk-Free Rate



Notes: Fraction of volatility fluctuations (in percent) in the price-dividend ratio and the risk-free rate that is due to x_t, $x_{\lambda,t}$, and $\sigma^2_{x,t}$, respectively. We do not present the graphs for $\sigma^2_{c,t}, \sigma^2_{d,t}$ since their time-varying shares are less than 1% on average. See the main text for computational details.

data set. To do so, we report percentiles of the posterior predictive distribution for various sample moments based on simulations from the posterior distribution of the same length as the data.[23] Typically, the posterior predictive distribution is computed to reflect both parameter and shock uncertainty. In our application the effect of the parameter uncertainty is an order of magnitude smaller than the effect of the shock (or sampling) uncertainty. Thus, we decided to fix the parameters at their posterior median values as this facilitates a clear comparison between the two types of model parameterization.

Results are summarized in Table 8. Means and standard deviations refer to annualized asset prices. We first focus on the results from estimating the full model based on cash-flow data and asset returns (full model estimation). All of the "actual" sample moments are within the 5th

[23]This is called a posterior predictive check; see Geweke (2005) for a textbook treatment. Specifically, the percentiles are obtained using the following simulation: draw parameters Θ^s from the posterior distribution; for each Θ^s simulate a trajectory Y^s (same number of observations as in the actual sample) and compute the sample statistics $\mathcal{S}(Y^s)$ of interest.

Table 8: Asset Return Moments

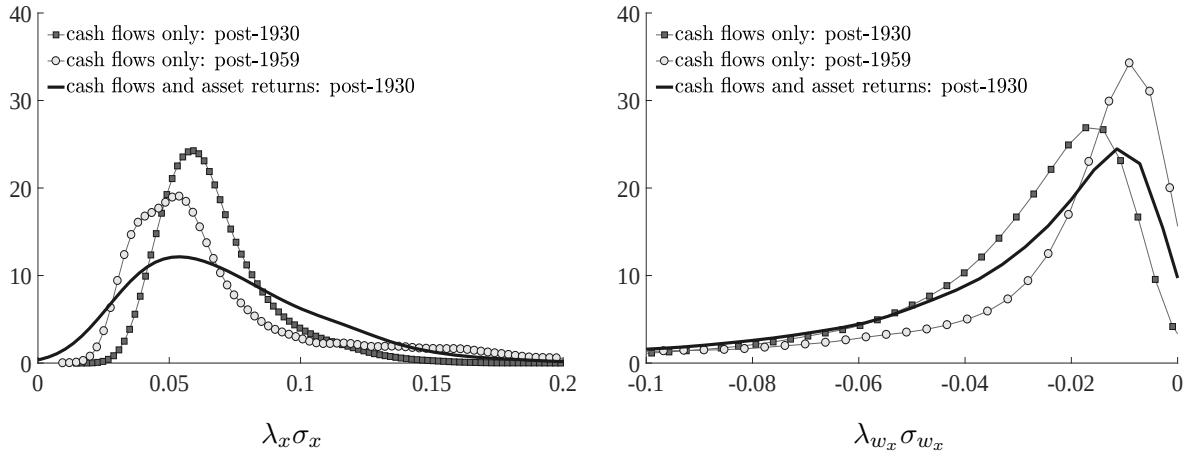
		Parameter Estimates are Based On					
	Data	Cash Flows & Asset Returns			Cash Flows Only		
		5%	50%	95%	5%	50%	95%
Mean (r_m)	6.06	2.56	6.15	10.29	1.99	4.61	7.55
StdDev (r_m)	19.8	14.9	22.6	46.1	11.0	16.6	28.3
AC1 (r_m)	-0.01	-0.30	-0.05	0.18	-0.29	-0.02	0.19
Corr ($\Delta c, r_m$)	0.11	-0.10	0.10	0.29	-0.10	0.12	0.32
Mean (pd)	3.40	2.63	3.14	3.41	3.26	3.42	3.51
StdDev (pd)	0.45	0.17	0.32	0.76	0.11	0.18	0.39
AC1 (pd)	0.87	0.49	0.75	0.89	0.32	0.62	0.82
Mean (r_f)	0.37	-0.56	0.47	1.32	0.22	0.97	1.64
StdDev (r_f)	2.85	1.54	2.09	2.90	1.61	1.93	2.41
AC1 (r_f)	0.64	0.38	0.57	0.73	0.35	0.52	0.67

Notes: We present descriptive statistics for log returns of the aggregate stock market (r_m), its correlation with consumption growth (Δc), the log risk-free rate (r_f), and the log price-dividend ratio (pd). We report means (Mean), standard deviations (StdDev), first-order sample autocorrelations (AC1), and correlations (Corr). Market returns, the risk-free rate, and the price-dividend ratio refer to 12-month averages (in percent). Computing asset pricing implications for the cash-flow-only estimates requires calibration of the preference parameters and the preference shock $x_{\lambda,t}$. We set $\delta, \psi, \gamma, \rho_\lambda, \sigma_\lambda^2$ to the median posterior estimates from Table 7.

and the 95th percentile of the corresponding posterior predictive distribution.[24] In particular, the model generates a sizable mean log market return with median value of 6.2%, and a sizeable equity risk premium with a median value of about 8.2%. Consistent with the data, the model's return variability is about 22%. The high volatility of the market returns translates into a large variability of the sample moments. As in the data, the model generates both a highly variable and persistent price-dividend ratio. The median and 95th percentile of the price-dividend volatility distribution are significantly larger than in other LRR calibrated models with Gaussian shocks. This feature owes in part to the fact the model contains an independent dividend volatility process. Finally, partly due to the preference shocks, the model is able to reproduce the observed sample moments of the risk-free rate.

[24] Although not reported in the table this is also the case for the mean, standard deviation and first autocorrelation moments of consumption and dividend growth.

Figure 8: Posterior Distribution of Market Prices of Risks



Notes: We plot posterior densities of $\lambda_x\sigma_x$ and $\lambda_{w_x}\sigma_{w_x}$ from the estimation with cash flow data only from post-1930 (squared-line) and from post-1959 samples (circled-line), respectively, and from the estimation with cash flow and asset return data from post-1930 sample (solid-line).

In Section 5.2 we noted that the parameter estimates for the cash flow processes change a bit once asset pricing data are included. To assess the economic implications of the parameter differentials, we combine the cash flow process parameter estimates reported in Table 6 (1930-2014 sample) with the posterior median estimates of the preference parameters and the preference shock $x_{\lambda,t}$ from the full estimation. Because the cash-flow-only model was estimated without the third volatility process $\sigma^2_{x,t}$, we set $h_{c,t} = h_{x,t}$ when re-computing the asset pricing implications of the LRR model. The last three columns of Table 8 show that, due to a lower persistence ρ, the cash-flow-only estimates generate a slightly lower mean and variance for the market return, and a slightly higher and less volatile price-dividend ratio and risk-free rate. The standard deviation and autocorrelation of the price-dividend ratio and the standard deviation of the risk-free rate lie just outside the posterior predictive bands, whereas all other sample moments continue to fall within the bands. This confirms that even the cash-flow-only estimates of the endowment process parameters can generate realistic asset price fluctuations.[25]

Figure 8 compares the posterior distributions for the appropriately-scaled market prices of risk, $\lambda_x\sigma_x$ and $\lambda_{w_x}\sigma_{w_x}$, based on the estimation with and without asset prices. The posterior densities are remarkably similar. In fact, the modes of the distributions are almost identical; the main difference lies in the dispersion of the densities. This indicates that the lower risk premium (see

[25]Chen, Dou, and Kogan (2015) formalize this comparison by developing a measure of model fragility, roughly speaking based on the discrepancy between the posterior medians obtained under the cash-flow-only estimation and the estimation with asset returns.

Table 8) obtained under the cash-flow-only estimate is due to smaller return exposures to the shocks (βs in (15)). [26]

Consumption Growth and Excess Return Predictability. Asset pricing models are often evaluated based on their implications for the predictability of future cash flows and returns. In the model the price-dividend ratio is determined by multiple state variables. Consequently, a VAR-based predictive regression is a natural starting point. As in Bansal, Kiku, and Yaron (2012) we estimate a first-order VAR that includes consumption growth, the price-dividend ratio, the real risk-free rate, and the market return. Based on the estimated VAR coefficients we compute R^2's for cumulative H-step-ahead consumption growth and excess returns:

$$\sum_{h=1}^{H} \Delta c_{t+h} \quad \text{and} \quad \sum_{h=1}^{H} (r_{m,t+h} - r_{f,t+h-1}).$$

While the VAR-based predictive checks are appealing from a theoretical perspective, much of the empirical literature focuses on R^2's from univariate predictive regressions using the price-dividend ratio as the only regressor. We subsequently consider both multivariate and univariate regressions.

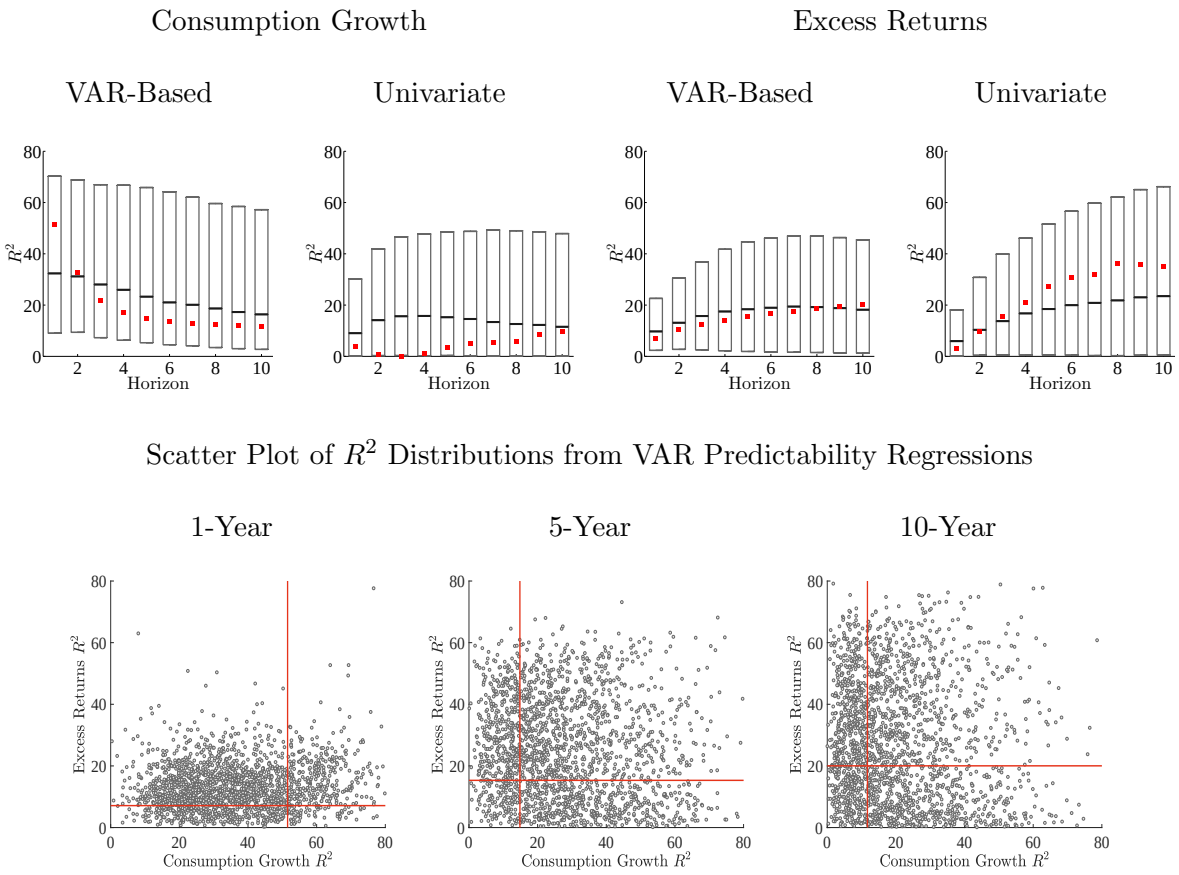
Predictive checks are graphically summarized in Figure 9. We begin with a discussion of the results depicted in the four panels of the top row of the figure. The sample statistics considered are the R^2 values obtained from predictability regressions. The top and bottom ends of the boxes correspond to the 5th and 95th percentiles, respectively, of the posterior predictive distribution, and the horizontal bars signify the medians. The predictive intervals reflect the fact that we are repeatedly generating data from the model and computing a sample statistic for each of these simulated trajectories.[27] The small squares correspond to R^2 statistics computed from "actual" U.S. data.

The top left panel of Figure 9 depicts results for the VAR-based predictability regressions for consumption growth. Based on multiple variables, consumption growth is *highly* predictable in the data. At the one-year horizon the R^2 is about 52% (see also Bansal, Kiku, Shaliastovich, and Yaron (2014)). While the predictability diminishes over time, it is still nontrivial with an R^2 of 12% at

[26] For completeness, using the SDF decomposition in Equation (A.18) in the Online Appendix, we also report λ_i and $\lambda_i \sigma_i$ for $i = \{c, x, \lambda, w_w, w_c\}$ at our median parameter estimates. The resulting values are $\{8.9, 695.2, 406.3, -26572824, -3391.0\}$ and $\{0.03, 0.07, 0.16, -0.03, -0.01\}$ respectively. These figures are consistent with the variance decomposition of the risk-free rate and equity return presented earlier, whereby the former has a relatively large exposure (β) to the preference shock while the latter has a large exposure to the growth and volatility shock.

[27] To ease subsequent comparisons, we condition on the the posterior median estimates of the LRR model. This is innocuous because the contribution of parameter uncertainty to the variability of the posterior predictive distributions is small.

Figure 9: Predictability Checks



Notes: We fix the parameters at their posterior median estimates and simulate data sets. The red squares represent R^2 values obtained from the actual data. The boxes represent 90% posterior predictive intervals and the horizontal lines represent medians. The "Benchmark" case is based on simulations with all five state variables x_t, $x_{\lambda,t}$, $\sigma^2_{x,t}$, $\sigma^2_{c,t}$, and $\sigma^2_{d,t}$; The horizon is measured in years. The VAR-Based R^2s are constructed as in Hodrick (1992). In the bottom panel, the intersection of the solid lines indicates the R^2 values obtained from the actual data.

the 10-year horizon. The key finding is that the data R^2s lie within the 90% credible intervals constructed from the model-implied predictive distribution. At the one-year horizon the median of the model-implied R^2 is somewhat lower than its data estimate, whereas over horizons of three years or more, the medians are slightly larger than the data estimates.

Panel 2 in the top row of Figure 9, labeled "Univariate", provides results for univariate consumption growth predictability regressions. As for the VAR-based predictability checks, we simulate the LRR model with all of its five state variables: x_t, $x_{\lambda,t}$, $\sigma^2_{x,t}$, $\sigma^2_{c,t}$, and $\sigma^2_{d,t}$. However, we only use the price-dividend ratio to predict future consumption growth. As is well known, when the price-dividend ratio is used as a single regressor, it produces low R^2s. They are less than 5% for

horizons from one to eight years and reach almost 10% at the ten-year horizon.[28] The median R^2 values obtained from regressions on model-generated data are between 10% to 15%, slightly higher than in the actual data. However, the posterior predictive intervals range from 0 to 30% for the one-year horizon and from 0 to about 50% for horizons longer than three years, which means that there is no evidence in the data that contradicts our estimated asset pricing model.

Panels 3 and 4 in the top row of Figure 9 show the respective VAR and univariate predictive R^2s for future excess returns. It is noteworthy that the VAR median R^2s of the model-based estimates are almost perfectly aligned with the data-based estimates. The model also performs quite well in terms of the univariate excess return predictability regressions. Specifically, for all horizons the median of the model-implied distribution of R^2s are quite close to actual data R^2s and the model-based credible intervals contain the R^2 obtained from the actual data. The good performance is obtained because, according to the model, the price-dividend ratio is the most important predictor of long-horizon excess returns among the observables.[29]

While our model passes the predictive checks, the credible intervals depicted in Figure 9 are wide. The high variability of the sampling distribution of the R^2 measures under the LRR model implies that despite their popularity the predictability regressions have little power to detect model misspecifications. The diffuse and skewed sampling distributions of the R^2 statistics are caused by various non-standard features of predictive regressions. Due to overlapping time periods, residuals are typically serially correlated and lagged residuals may be correlated with the predictor. Moreover, the persistent component of the dependent variable (consumption growth or excess returns) is dominated by *iid* shocks and the right-hand-side regressor (price-dividend ratio) is highly persistent – a feature that can render the predictive regressions spurious (see Hodrick (1992) and Stambaugh (1999)).[30]

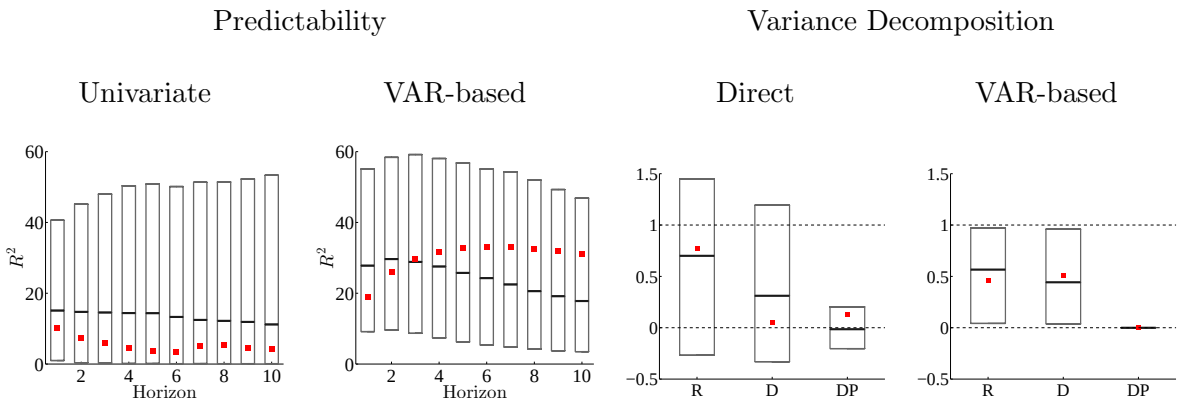
As a final check, the bottom row of Figure 9 illustrates the model-implied *joint* distribution of R^2s for predicting consumption growth and excess returns. Each dot in these scatter plots is obtained

[28]The univariate-based low R^2s for the first several years are consistent with the findings in Beeler and Campbell (2012) Table 4—the slight differences attributed to the longer sample available here.

[29]In the Online Appendix we explore the relative importance of "growth" and "volatility" risks by simulating model specifications that are only driven by (i) x_t and $\sigma^2_{x,t}$ or (ii) x_t. In Case (i) the posterior predictive distributions are quite close to the ones in the two "univariate" subplots of Figure 9 because x_t and $\sigma^2_{x,t}$ represent the key pricing state variables. In Case (ii) the credible intervals are often too small and do not encompass the data estimates. Drive by x_t only, the model generates too much consumption predictability, thereby highlighting that volatility shocks play an important role in lowering the model-implied predictability to a more realistic level.

[30]Valkanov (2003) derived an asymptotic distribution of the R^2 under the assumption that the regressor follows a local-to-unity process. He shows that the goodness-of-fit measure converges to a random limit as the sample size increases. More recently, Bauer and Hamilton (2015) studied the sampling distribution of R^2 measures in predictive regressions for bond returns, which exhibit similar distortions.

Figure 10: Dividend Growth Predictability and Dividend Yield Variance Decomposition



Notes: (Predictability) We fix the parameters at their posterior median estimates and simulate data sets. The horizon is measured in years. We run a univariate regression with the price-dividend ratio as predictor of future dividend growth. For the multivariate regression, we consider a first-order VAR that includes consumption growth, dividend growth, the price-dividend ratio, and the real risk-free rate. Based on the estimated coefficients we compute R^2's for cumulative H-step-ahead dividend growth. The red squares represent R^2 values obtained from the actual data. The boxes represent 90% posterior predictive intervals and the horizontal lines represent medians. The VAR-based R^2s are constructed as in Hodrick (1992). (Variance Decomposition, Direct) We regress 15-year ex post returns, dividend growth, and dividend yield, respectively, on a constant term and the dividend yield. (Variance Decomposition, VAR-based) We infer long-run coefficients ($k \to \infty$) from 1-year coefficients of the same VAR used for the predictability analysis. Using the Campbell-Shiller approximation, the fractions of dividend yield variation attributed to each source are provided as $1 \approx \frac{Cov(dp_t, \sum_{j=1}^{k} \varrho^{j-1} r_{t+j})}{Var(dp_t)} - \frac{Cov(dp_t, \sum_{j=1}^{k} \varrho^{j-1} \Delta d_{t+j})}{Var(dp_t)} + \frac{\varrho^k Cov(dp_t, dp_{t+k})}{Var(dp_t)}$. These components are marked as R, D, and DP respectively.

by computing the two R^2s based on a single model simulation. The intersection of the solid lines indicates the R^2 values computed from the actual data. The figure shows that the R^2 values at the 1-year and 5-year horizon are almost uniformly distributed over a rectangle. For every horizon the observed R^2s do not lie in the far tails of the posterior predictive distribution, which means that the model is also able to jointly generate the observed consumption growth and excess return predictability.

Dividend-Growth Predictability. Cochrane (2011) argues that there is very little dividend-growth predictability at all horizons. This view is based on a univariate regression with the price-dividend ratio as a predictor of future dividend growth. The data feature modest predictability, with an R^2 in the range of 4% to 9%, depicted by the red squares in the left panel of Figure 10. However, dividend growth is found to be highly predictable both at short and long horizons, once additional predictors are included in a VAR based predictive regression, with adjusted R^2s as large as 35% at the 10-year horizon (see Column 2 of Figure 10).[31] Importantly, in both the univariate

[31]This evidence is consistent with Lettau and Ludvigson (2005), Koijen and van Binsbergen (2010), and Jagannathan and Liu (2016) who report R^2 values from a VAR-based regression that range from 10% to 40%.

and VAR-based predictive regressions, the model implications for dividend growth predictability line up with the data and cover the data R^2s.

The strong evidence for dividend growth predictability has important implications for the variability of the log dividend yield dp_t. Based on the Campbell and Shiller (1988a) approximate present value identity it follows that

$$dp_t \approx \sum_{j=1}^{k} \varrho^{j-1} r_{t+j} - \sum_{j=1}^{k} \varrho^{j-1} \Delta d_{t+j} + \varrho^k dp_{t+k}, \tag{19}$$

where ϱ is an approximation constant based on the average dividend yield. Multiplying both sides of (19) by the log dividend yield and taking expectations implies that the variance of the current dividend yield can be attributed to its covariance with expected future returns, dividend growth rates, and the expected future dividend yield, respectively, marked as "R", "D", and "DP" in Figure 10 (see figure notes for details). As k approaches infinity, the dividend yield variability is explained completely by covariation with expected returns and cash flow growth. We compute the fraction of variability explained by the three covariances via "Direct" regression (setting k equal to 15 years and separately regressing the "R", "D", and "DP" components on the dividend yield) and "VAR-based" regression (inferring the $k = \infty$ decomposition from the coefficients of a VAR estimated based on annual data). The estimates based on the direct regressions attribute much of the variation in dividend yield to variation in discount rates (although not entirely), whereas the point estimates of the VAR attribute about half of the variation to discount rates and the other to dividend growth. Again, it is important to note that in both cases the model credible intervals contain the data point estimates. Moreover, in both cases the credible intervals around the point estimates are consistent with a view in which a large portion (about half) of the dividend yield variability is driven by cash flows.[32]

6 Conclusion

We developed a non-linear Bayesian state-space model that utilizes mixed frequency data to study the time series dynamics of consumption and its implications for asset pricing. We show that after accounting for monthly measurement errors there is strong evidence for both a small persistent predictable component as well as a stochastic volatility component in consumption growth. Importantly, this evidence emerges when the estimation uses information just from cash flows, namely,

[32] Albuquerque, Eichenbaum, Luo, and Rebelo (2016) also examine the 7- and 10-year correlations between cumulative return and cumulative consumption and dividend growth. For brevity, we defer this analysis to the Online Appendix where we show the model's credible confidence bands contain the data estimates.

consumption, consumption and output, and consumption and dividends. It is further reinforced and sharpened when the estimation uses consumption, dividends, and asset return data jointly. The estimation identifies three volatility processes: one governing dynamics of the persistent cash-flow growth component, and the other two controlling temporally independent shocks to consumption and dividend growth. The model is able to successfully capture many asset pricing moments and improve upon key predictability moments of previous LRR models.

Our findings raise the broader question of whether DSGE models, more generally, should have a predictable component built into one or more of the exogenous processes that drive macroeconomic fluctuations. If the goal of the modeling endeavor is to capture business cycle fluctuations at the quarterly level, then the answer is no, because the signal in the data is not strong enough to render macroeconomic predictions from a model without this predictable component to be inaccurate. But if the goal is to rely on long-horizon implications of the model, for instance, with respect to asset prices, then the answer is affirmative.

References

Albuquerque, R., M. Eichenbaum, V. Luo, and S. Rebelo (2016): "Valuation Risk and Asset Pricing," *Journal of Finance, forthcoming*.

Amir-Ahmadi, P., C. Matthes, and M.-C. Wang (2016): "Drifts and Volatilities under Measurement Error: Assessing Monetary Policy Shocks over the Last Century," *Quantitative Economics, forthcoming*.

An, S., and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4), 113–172.

Andreasen, M. (2010): "Stochastic Volatility and DSGE Models," *Economics Letters*, 108, 7–9.

Andrieu, C., A. Doucet, and R. Holenstein (2010): "Particle Markov Chain Monte Carlo Methods (with Discussion)," *Journal of the Royal Statistical Society, Series B*, 72, 1–33.

Aruoba, S., F. Diebold, and C. Scotti (2009): "Real-Time Measurement of Business Conditions," *Journal of Business and Economic Statistics*, 27(4), 417–427.

Bansal, R., A. Gallant, and G. Tauchen (2007): "Rational Pessimism, Rational Exuberance, and Asset Pricing Models," *Review of Economic Studies*, 74, 1005–1033.

Bansal, R., V. Khatacharian, and A. Yaron (2005): "Interpretable Asset Markets?," *European Economic Review*, 49, 531–560.

Bansal, R., D. Kiku, I. Shaliastovich, and A. Yaron (2014): "Volatility, the Macroeconomy and Asset Prices," *Journal of Finance*, 69, 2471–2511.

Bansal, R., D. Kiku, and A. Yaron (2012): "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices," *Critical Finance Review*, 1, 183–221.

——— (2014): "Risks for the Long Run: Estimation with Time Aggregation," Manuscript, University of Pennsylvania and Duke University.

Bansal, R., and A. Yaron (2004): "Risks For the Long Run: A Potential Resolution of Asset Pricing Puzzles," *Journal of Finance*, 59, 1481–1509.

Barro, R. (2009): "Rare Disasters, Asset Prices, and Welfare Costs," *American Economic Review*, 99, 243–264.

Bauer, M., and J. Hamilton (2015): "Robust Bond Risk Premia," Federal Reserve Bank of San Francisco Working Paper 2015-15.

Beeler, J., and J. Campbell (2012): "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment," *Critical Finance Review*, 1, 141–182.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77, 623–685.

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

Campbell, J., and R. Shiller (1988a): "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–227.

——— (1988b): "Stock Prices, Earnings, and Expected Dividends," *Journal of Finance*, 43, 661–676.

Carter, C. K., and R. Kohn (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81(3), 541–553.

Chen, H., W. D. Dou, and L. Kogan (2015): "Measuring the "Dark Matter" in Asset Pricing Models," *Manuscript, MIT*.

Chen, R., and J. Liu (2000): "Mixture Kalman Filters," *Journal of the Royal Statistical Society Series B*, 62, 493–508.

Chernov, M., R. Gallant, E. Ghysels, and G. Tauchen (2003): "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225–257.

Cochrane, J. (2011): "Presidential Address: Discount Rates," *Journal of Finance*, 66(4), 1047–1108.

Creal, D. D., and J. C. Wu (2015): "Bond Risk Premia in Consumption-based Models," *Manuscript, Chicago Booth*.

Croce, M. (2014): "Long-run Productivity Risk: A New Hope for Production-based Asset Pricing?," *Journal of Monetary Economics*, 66, 13–31.

Doh, T., and S. Wu (2015): "Cash Flow and Risk Premium Dynamics in an Equilibrium Asset Pricing Model with Recursive Preferences," *FRB Kansas City Research Working Paper*, 15-12.

Drost, F., and T. Nijman (1993): "Temporal Aggregation of Garch Processes," *Econometrica*, 61, 909–927.

Epstein, L., and S. Zin (1989): "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937–969.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74(4), 1059–1087.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez (2011): "Macroeconomics and Volatility: Data, Models, and Estimation," in *Advances in Economics and Econometrics: Theory and Applications, Tenth World Congress of the Econometric Society*, ed. by D. Acemoglu, M. Arellano, and E. Deckel. Cambridge University Press.

Geweke, J. (2005): *Contemporary Bayesian Econometrics and Statistics*. New Jersey: John Wiley & Sons.

Gourio, F. (2012): "Disaster Risk and Business Cycles," *American Economic Review*, 102(6), 2734?2766.

Hall, R. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86(6), 971?986.

Hansen, L. (2007): "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk," *American Economic Review*, 97(2), 1–30.

Hansen, L., and T. Sargent (2007): *Robustness*. Princeton University Press.

Hansen, L. P., J. C. Heaton, and N. Li (2008): "Consumption Strikes Back? Measuring Long-Run Risk," *Journal of Political Economy*, 116(2), 260–302.

Harvey, A. (1989): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press.

Herbst, E., and F. Schorfheide (2015): *Bayesian Estimation of DSGE Models*. Princeton University Press.

Hodrick, R. (1992): "Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement," *Review of Financial Studies*, 5, 357–386.

Jagannathan, R., and B. Liu (2016): "Dividend Dynamics, Learning, and Expected Stock Index Returns," Manuscript, Northwestern University.

Johannes, M., L. Lochstoer, and Y. Mou (2016): "Learning about Consumption Dynamics," *Journal of Finance, forthcoming*.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361–393.

Koijen, R., and J. van Binsbergen (2010): "Predictive Regressions: A Present-Value Approach," *Journal of Finance*, 65(4), 1439–1471.

Lettau, M., and S. Ludvigson (2005): "Expected Returns and Expected Dividend Growth," *Journal of Financial Economics*, 76(3), 583–626.

Lucas, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46(6), 1429–1445.

Mariano, R., and Y. Murasawa (2003): "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series," *Journal of Applied Econometrics*, 18, 427–443.

Nakamura, E., D. Sergeyev, and J. Steinsson (2015): "Growth-Rate and Uncertainty Shocks in Consumption: Cross-Country Evidence," *Manuscript, Columbia University*.

Pohl, W., K. Schmedders, and O. Wilms (2016): "Higher-Order Effects in Asset-Pricing Models with Long-Run Risks," *Manuscript, University of Zurich*.

Romer, C. (1986): "New Estimates of Prewar Gross National Product and Unemployment," *The Journal of Economic History*, 46(2), 341–352.

Romer, C. (1989): "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1869-1908," *Journal of Political Economy*, 97(1), 1–37.

Schorfheide, F., and D. Song (2015): "Real-Time Forecasting with a Mixed-Frequency VAR," *Journal of Business and Economic Statistics*, 33(3), 366–380.

Slesnick, D. (1998): "Are Our Data Relevant to the Theory? The Case of Aggregate Consumption," *Journal of Business and Economic Statistics*, 16(1), 52–61.

Stambaugh, R. F. (1999): "Predictive Regressions," *Journal of Financial Economics*, 54, 375–421.

Valkanov, R. (2003): "Long-Horizon Regressions: Theoretical Results and Applications," *Journal of Financial Economics*, 68, 201–232.

Wilcox, D. (1992): "The Construction of the U.S. Consumption Data: Some Facts and Their Implications for Empirical Work," *American Economic Review*, 82, 922–941.

OnlineAppendix

Identifying Long-Run Risks: A Bayesian Mixed-Frequency Approach

Frank Schorfheide, Dongho Song, and Amir Yaron

The Online Appendix contains supplementary material and consists of the following sections:

A. Data Sources

B. The Measurement Error Model for Consumption

C. Solving the Long-Run Risks Model

D. State-Space Representations of the Empirical Models

E. Posterior Inference

F. Supplementary Figures and Tables

A Data Source

A.1 Nominal PCE

We download seasonally adjusted data for nominal PCE from NIPA Tables 2.3.5 and 2.8.5. We then compute within-quarter averages of monthly observations and within-year averages of quarterly observations.

A.2 Real PCE

We use Table 2.3.3., Real Personal Consumption Expenditures by Major Type of Product, Quantity Indexes (A:1929-2014)(Q:1947:Q1-2014:Q4) to extend Table 2.3.6., Real Personal Consumption Expenditures by Major Type of Product, Chained Dollars (A:1995-2014) (Q:1995:Q1-2014:Q4). Monthly data are constructed analogously using Table 2.8.3. and Table 2.8.6.

A.3 Real Per Capita PCE: ND+S

The LRR model defines consumption as per capita consumer expenditures on nondurables and services. We download mid-month population data from NIPA Table 7.1.(A:1929-2014)(Q:1947:Q1-2014:Q4) and from Federal Reserve Bank of St. Louis' FRED database (M:1959:M1-2014:M12). We convert consumption to per capita terms.

A.4 Dividend and Market Returns Data

Data are from the Center for Research in Security Prices (CRSP). The three monthly series from CRSP are the value-weighted with-, RN_t, and without-dividend nominal returns, RX_t, of CRSP stock market indexes (NYSE/AMEX/NASDAQ/ARCA), and the CPI inflation rates, π_t. The sample period is from 1929:M1 to 2014:M12. The monthly real dividend series are constructed as in Hodrick (1992):

1. A normalized nominal value-weighted price series is produced by initializing $P_0 = 1$ and recursively setting $P_t = (1 + RX_t)P_{t-1}$.

2. A normalized nominal divided series, D_t^{Raw}, is obtained by recognizing that $D_t^{Raw} = (RN_t - RX_t)P_{t-1}$.

3. Following Robert Shiller we smooth out dividend series by aggregating 3 months values of the raw nominal dividend series $D_t = \sum_{i=0}^{2} D_{t-i}^{Raw}$ and apply the following quarterly interpolation. Here, D_t, D_{t-3}, ... is the last month of the quarter.

$$D_{t-m} = D_t - \frac{m}{3}(D_t - D_{t-3}), \quad m \in \{0, 1, 2\}. \tag{A.1}$$

4. We then compute the real dividend growth $g_{d,t}$ by subtracting the actual inflation from the interpolated nominal dividend growth

$$g_{d,t} = \log(D_t) - \log(D_{t-1}) - \pi_t. \tag{A.2}$$

   Here inflation rates are computed using the log differences of the consumer price index (CPI) from the Bureau of Labor Statistics.

Market returns, RN_{t+1}, are also converted from nominal to real terms using the CPI inflation rates and denoted by $r_{m,t+1}$.

A.5 Ex Ante Risk-Free Rate

The ex ante risk-free rate is constructed as in the online appendix of Beeler and Campbell (2012). Nominal yields to calculate risk-free rates are the CRSP Fama Risk Free Rates. Even though our model runs in monthly frequencies, we use the three-month yield because of the larger volume and higher reliability. We subtract annualized three-month inflation, $\pi_{t,t+3}$, from the nominal yield, $i_{f,t}$, to form a measure of the ex post (annualized) real three-month interest rate. The ex ante real

risk-free rate, $r_{f,t}$, is constructed as a fitted value from a projection of the ex post real rate on the current nominal yield, $i_{f,t}$, and inflation over the previous year, $\pi_{t-12,t}$:

$$\begin{aligned} i_{f,t} - \pi_{t,t+3} &= \beta_0 + \beta_1 i_{f,t} + \beta_2 \pi_{t-12,t} + \varepsilon_{t+3} \\ r_{f,t} &= \hat{\beta}_0 + \hat{\beta}_1 i_{f,t} + \hat{\beta}_2 \pi_{t-12,t}. \end{aligned}$$

The ex ante real risk-free rates are available from 1929:M1 to 2014:M12.

B The Measurement-Error Model for Consumption

For expositional purposes, we assume that the accurately measured low-frequency observations are available at quarterly frequency (instead of annual frequency as in the main text). Correspondingly, we define the time subscript $t = 3(j-1)+m$, where month $m = 1, 2, 3$ and quarter $j = 1, \ldots$. We use uppercase C to denote the level of consumption and lowercase c to denote percentage deviations from some log-linearization point. Growth rates are approximated as log differences and we use a superscript o to distinguish observed from "true" values.

The measurement-error model presented in the main text can be justified by assuming that the statistical agency uses a high-frequency proxy series to determine monthly consumption growth rates. We use $Z_{3(j-1)+m}$ to denote the monthly value of the proxy series and $Z^q_{(j)}$ the quarterly aggregate. Suppose the proxy variable provides a noisy measure of monthly consumption. More specifically, we consider a multiplicative error model of the form

$$Z_{3(j-1)+m} = C_{3(j-1)+m} \exp(\epsilon_{3(j-1)+m}). \tag{A.3}$$

The interpolation is executed in two steps. In the first step we construct a series $\tilde{C}^o_{3(j-1)+m}$, and in the second step we rescale the series to ensure that the reported monthly consumption data add up to the reported quarterly consumption data within the period. In Step 1, we start from the level of consumption in quarter $j-1$, $C^q_{(j-1)}$, and define

$$\begin{aligned}
\tilde{C}^o_{3(j-1)+1} &= C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+1}}{Z^q_{(j-1)}} \right) \\
\tilde{C}^o_{3(j-1)+2} &= C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+1}}{Z^q_{(j-1)}} \right) \left(\frac{Z_{3(j-1)+2}}{Z_{3(j-1)+1}} \right) = C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+2}}{Z^q_{(j-1)}} \right) \\
\tilde{C}^o_{3(j-1)+3} &= C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+1}}{Z^q_{(j-1)}} \right) \left(\frac{Z_{3(j-1)+2}}{Z_{3(j-1)+1}} \right) \left(\frac{Z_{3(j-1)+3}}{Z_{3(j-1)+2}} \right) = C^{q,o}_{(j-1)} \left(\frac{Z_{3(j-1)+3}}{Z^q_{(j-1)}} \right).
\end{aligned} \tag{A.4}$$

Thus, the growth rates of the proxy series are used to generate monthly consumption data for quarter q. Summing over the quarter yields

$$\tilde{C}^{q,o}_{(j)} = \sum_{m=1}^{3} \tilde{C}^o_{3(j-1)+m} = C^{q,o}_{(j-1)} \left[\frac{Z_{3(j-1)+1}}{Z^q_{(j-1)}} + \frac{Z_{3(j-1)+2}}{Z^q_{(j-1)}} + \frac{Z_{3(j-1)+3}}{Z^q_{(j-1)}} \right] = C^{q,o}_{(j-1)} \frac{Z^q_{(j)}}{Z^q_{(j-1)}}. \tag{A.5}$$

In Step 2, we adjust the monthly estimates $\tilde{C}^o_{3(j-1)+m}$ by the factor $C^{q,o}_{(j)}/\tilde{C}^{q,o}_{(j)}$, which leads to

$$\begin{aligned} C^o_{3(j-1)+1} &= \tilde{C}^o_{3(j-1)+1}\left(\frac{C^{q,o}_{(j)}}{\tilde{C}^{q,o}_{(j)}}\right) = C^{q,o}_{(j)}\frac{Z_{3(j-1)+1}}{Z^q_{(j)}} \\ C^o_{3(j-1)+2} &= \tilde{C}^o_{3(j-1)+2}\left(\frac{C^{q,o}_{(j)}}{\tilde{C}^{q,o}_{(j)}}\right) = C^{q,o}_{(j)}\frac{Z_{3(j-1)+2}}{Z^q_{(j)}} \\ C^o_{3(j-1)+3} &= \tilde{C}^o_{3(j-1)+3}\left(\frac{C^{q,o}_{(j)}}{\tilde{C}^{q,o}_{(j)}}\right) = C^{q,o}_{(j)}\frac{Z_{3(j-1)+3}}{Z^q_{(j)}} \end{aligned} \tag{A.6}$$

and guarantees that

$$C^{q,o}_{(j)} = \sum_{m=1}^{3} C^o_{3(j-1)+m}.$$

We now define the growth rates $g^o_{c,t} = \log C^o_t - \log C^o_{t-1}$ and $g_{c,t} = \log C_t - \log C_{t-1}$. By taking logarithmic transformation of (A.3) and (A.6) and combining the resulting equations, we can deduce that the growth rates for the second and third month of quarter q are given by

$$\begin{aligned} g^o_{c,3(j-1)+2} &= g_{c,3(j-1)+2} + \epsilon_{3(j-1)+2} - \epsilon_{3(j-1)+1} \\ g^o_{c,3(j-1)+3} &= g_{c,3(j-1)+3} + \epsilon_{3(j-1)+3} - \epsilon_{3(j-1)+2}. \end{aligned} \tag{A.7}$$

The derivation of the growth rate between the third month of quarter $j-1$ and the first month of quarter j is a bit more cumbersome. Using (A.6), we can write the growth rate as

$$\begin{aligned} g^o_{c,3(j-1)+1} &= \log C^{q,o}_{(j)} + \log Z_{3(j-1)+1} - \log Z^q_{(j)} \\ &\quad - \log C^{q,o}_{(j-1)} - \log Z_{3(j-2)+3} + \log Z^q_{(j-1)}. \end{aligned} \tag{A.8}$$

To simplify (A.8) further, we are using a log-linear approximation. Suppose we log-linearize an equation of the form

$$X^q_{(j)} = X_{3(j-1)+1} + X_{3(j-1)+2} + X_{3(j-1)+3}$$

around X^q_* and $X_* = X^q_*/3$, using lowercase variables to denote percentage deviations from the log-linearization point. Then,

$$x^q_{(j)} \approx \frac{1}{3}(x_{3(j-1)+1} + x_{3(j-1)+2} + x_{3(j-1)+3}).$$

Using (A.3) and the definition of quarterly variables as sums of monthly variables, we can apply the log-linearization as follows:

$$\log C^{q,o}_{(j)} - \log Z^q_{(j)} = \log(C^q_*/Z^q_*) + \epsilon^q_{(j)} - \frac{1}{3}\big(\epsilon_{3(j-1)+1} + \epsilon_{3(j-1)+2} + \epsilon_{3(j-1)+3}\big). \tag{A.9}$$

Substituting (A.9) into (A.8) yields

$$\begin{aligned} g^o_{c,3(j-1)+1} &= g_{c,3(j-1)+1} + \epsilon_{3(j-1)+1} - \epsilon_{3(j-2)+3} + \epsilon^q_{(j)} - \epsilon^q_{(j-1)} \\ &\quad -\frac{1}{3}\big(\epsilon_{3(j-1)+1} + \epsilon_{3(j-1)+2} + \epsilon_{3(j-1)+3}\big) + \frac{1}{3}\big(\epsilon_{3(j-2)+1} + \epsilon_{3(j-2)+2} + \epsilon_{3(j-2)+3}\big). \end{aligned} \tag{A.10}$$

An "annual" version of this equation appears in the main text.

C Solving the Long-Run Risks Model

This section provides solutions for the consumption and dividend claims for the endowment process:

$$\begin{aligned}
g_{c,t+1} &= \mu_c + x_t + \sigma_{c,t}\eta_{c,t+1} \\
g_{d,t+1} &= \mu_d + \phi x_t + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1} \\
x_{t+1} &= \rho x_t + \sigma_{x,t}\eta_{x,t+1} \\
x_{\lambda,t+1} &= \rho_\lambda x_{\lambda,t} + \sigma_\lambda \eta_{\lambda,t+1} \\
\sigma^2_{c,t+1} &= (1-\nu_c)(\varphi_c\bar{\sigma})^2 + \nu_c\sigma^2_{c,t} + \sigma_{w_c} w_{c,t+1} \\
\sigma^2_{x,t+1} &= (1-\nu_x)(\varphi_x\bar{\sigma})^2 + \nu_x\sigma^2_{x,t} + \sigma_{w_x} w_{x,t+1} \\
\sigma^2_{d,t+1} &= (1-\nu_d)(\varphi_d\bar{\sigma})^2 + \nu_d\sigma^2_{d,t} + \sigma_{w_d} w_{d,t+1} \\
&\quad \eta_{i,t+1}, \eta_{\lambda,t+1}, w_{i,t+1} \sim N(0,1), \quad i \in \{c,x,d\}.
\end{aligned} \tag{A.11}$$

The Euler equation for the economy is

$$\mathbb{E}_t\left[\exp\left(m_{t+1} + r_{i,t+1}\right)\right] = 1, \quad i \in \{c,m\}, \tag{A.12}$$

where

$$m_{t+1} = \theta \log\delta + \theta x_{\lambda,t+1} - \frac{\theta}{\psi} g_{c,t+1} + (\theta-1) r_{c,t+1} \tag{A.13}$$

is the log of the real stochastic discount factor (SDF), $r_{c,t+1}$ is the log return on the consumption claim, and $r_{m,t+1}$ is the log market return. (A.13) is derived in Section C.5 below. Returns are given by the approximation of Campbell and Shiller (1988a):

$$\begin{aligned}
r_{c,t+1} &= \kappa_0 + \kappa_1 pc_{t+1} - pc_t + g_{c,t+1} \\
r_{m,t+1} &= \kappa_{0,m} + \kappa_{1,m} pd_{t+1} - pd_t + g_{d,t+1}.
\end{aligned} \tag{A.14}$$

The risk premium on any asset is

$$\mathbb{E}_t(r_{i,t+1} - r_{f,t}) + \frac{1}{2} Var_t(r_{i,t+1}) = -Cov_t(m_{t+1}, r_{i,t+1}). \tag{A.15}$$

In Section C.1 we solve for the law of motion for the return on the consumption claim, $r_{c,t+1}$. In Section C.2 we solve for the law of motion for the market return, $r_{m,t+1}$. The risk-free rate is derived in Section C.3. All three solutions depend on linearization parameters that are derived in Section C.4. Finally, as mentioned above, the SDF is derived in Section C.5.

C.1 Consumption Claim

In order to derive the dynamics of asset prices, we rely on approximate analytical solutions. Specifically, we conjecture that the price-consumption ratio follows

$$pc_t = A_0 + A_1 x_t + A_{1,\lambda} x_{\lambda,t} + A_{2,c}\sigma_{c,t}^2 + A_{2,x}\sigma_{x,t}^2 \tag{A.16}$$

and solve for A's using (A.11), (A.12), (A.14), and (A.16).

From (A.11), (A.14), and (A.16)

$$\begin{aligned} r_{c,t+1} &= \left\{\kappa_0 + A_0(\kappa_1 - 1) + \mu_c + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2\right\} \\ &+ \frac{1}{\psi}x_t + A_{1,\lambda}(\kappa_1\rho_\lambda - 1)x_{\lambda,t} + A_{2,x}(\kappa_1\nu_x - 1)\sigma_{x,t}^2 + A_{2,c}(\kappa_1\nu_c - 1)\sigma_{c,t}^2 \\ &+ \sigma_{c,t}\eta_{c,t+1} + \kappa_1 A_1\sigma_{x,t}\eta_{x,t+1} + \kappa_1 A_{1,\lambda}\sigma_\lambda\eta_{\lambda,t+1} + \kappa_1 A_{2,x}\sigma_{w_x}w_{x,t+1} + \kappa_1 A_{2,c}\sigma_{w_c}w_{c,t+1} \end{aligned} \tag{A.17}$$

and from (A.11), (A.12), (A.14), and (A.16)

$$\begin{aligned} m_{t+1} &= (\theta-1)\left\{\kappa_0 + A_0(\kappa_1 - 1) + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2\right\} \\ &- \gamma\mu + \theta\log\delta - \frac{1}{\psi}x_t + \rho_\lambda x_{\lambda,t} + (\theta-1)A_{2,x}(\kappa_1\nu_x - 1)\sigma_{x,t}^2 + (\theta-1)A_{2,c}(\kappa_1\nu_c - 1)\sigma_{c,t}^2 \\ &- \gamma\sigma_{c,t}\eta_{c,t+1} + (\theta-1)\kappa_1 A_1\sigma_{x,t}\eta_{x,t+1} + \left\{(\theta-1)\kappa_1 A_{1,\lambda} + \theta\right\}\sigma_\lambda\eta_{\lambda,t+1} \\ &+ (\theta-1)\kappa_1 A_{2,x}\sigma_{w_x}w_{x,t+1} + (\theta-1)\kappa_1 A_{2,c}\sigma_{w_c}w_{c,t+1}. \end{aligned} \tag{A.18}$$

The solutions for A's that describe the dynamics of the price-consumption ratio are determined from

$$\mathbb{E}_t\left[m_{t+1} + r_{c,t+1}\right] + \frac{1}{2}Var_t\left[m_{t+1} + r_{c,t+1}\right] = 0$$

and they are

$$A_1 = \frac{1 - \frac{1}{\psi}}{1 - \kappa_1\rho}, \quad A_{1,\lambda} = \frac{\rho_\lambda}{1 - \kappa_1\rho_\lambda}, \quad A_{2,x} = \frac{\frac{\theta}{2}(\kappa_1 A_1)^2}{1 - \kappa_1\nu_x}, \quad A_{2,c} = \frac{\frac{\theta}{2}(1 - \frac{1}{\psi})^2}{1 - \kappa_1\nu_c} \tag{A.19}$$

and $A_0 = \frac{A_0^1 + A_0^2}{1-\kappa_1}$, where

$$\begin{aligned} A_0^1 &= \log\delta + \kappa_0 + \mu(1 - \frac{1}{\psi}) + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2 \\ A_0^2 &= \frac{\theta}{2}\left\{(\kappa_1 A_{1,\lambda} + 1)^2\sigma_\lambda^2 + (\kappa_1 A_{2,x}\sigma_{w_x})^2 + (\kappa_1 A_{2,c}\sigma_{w_c})^2\right\}. \end{aligned}$$

For convenience, (A.18) can be rewritten as

$$m_{t+1} - \mathbb{E}_t[m_{t+1}] = \lambda_c\sigma_{c,t}\eta_{c,t+1} + \lambda_x\sigma_{x,t}\eta_{x,t+1} + \lambda_\lambda\sigma_\lambda\eta_{\lambda,t+1} + \lambda_{w_x}\sigma_{w_x}w_{x,t+1} + \lambda_{w_c}\sigma_{w_c}w_{c,t+1}.$$

Note that λs represent the market price of risk for each source of risk. To be specific,

$$\begin{aligned}\lambda_c &= \gamma, \quad \lambda_x = (\gamma - \frac{1}{\psi})\frac{\kappa_1}{1-\kappa_1\rho}, \quad \lambda_\lambda = -\frac{\theta - \kappa_1\rho_\lambda}{1-\kappa_1\rho_\lambda}, \\ \lambda_{w_x} &= \frac{\theta(\gamma-\frac{1}{\psi})(1-\frac{1}{\psi})\kappa_1}{2(1-\kappa_1\nu_x)}(\frac{\kappa_1}{1-\kappa_1\rho})^2, \quad \lambda_{w_c} = \frac{\theta(\gamma-\frac{1}{\psi})(1-\frac{1}{\psi})\kappa_1}{2(1-\kappa_1\nu_c)}.\end{aligned} \tag{A.20}$$

Similarly, rewrite (A.17) as

$$r_{c,t+1} - \mathbb{E}_t[r_{c,t+1}] = \beta_{c,c}\sigma_{c,t}\eta_{c,t+1} + \beta_{c,x}\sigma_{x,t}\eta_{x,t+1} + \beta_{c,\lambda}\sigma_\lambda\eta_{\lambda,t+1} + \beta_{c,w_x}\sigma_{w_x}w_{x,t+1} + \beta_{c,w_c}\sigma_{w_c}w_{c,t+1}$$

where

$$\beta_{c,c} = 1, \quad \beta_{c,x} = \kappa_1 A_1, \quad \beta_{c,\lambda} = \kappa_1 A_{1,\lambda}, \quad \beta_{c,w_x} = \kappa_1 A_{2,x}, \quad \beta_{c,w_c} = \kappa_1 A_{2,c}. \tag{A.21}$$

The risk premium for the consumption claim is

$$\begin{aligned}\mathbb{E}_t(r_{c,t+1} - r_{f,t}) + \frac{1}{2}Var_t(r_{c,t+1}) &= -Cov_t(m_{t+1}, r_{c,t+1}) \\ &= \beta_{c,x}\lambda_x\sigma^2_{x,t} + \beta_{c,c}\lambda_c\sigma^2_{c,t} + \beta_{c,\lambda}\lambda_\lambda\sigma^2_\lambda + \beta_{c,w_x}\lambda_{w_x}\sigma^2_{w_x} + \beta_{c,w_c}\lambda_{w_c}\sigma^2_{w_c}.\end{aligned} \tag{A.22}$$

C.2 Market Return

Similarly, using the conjectured solution to the price-dividend ratio

$$pd_t = A_{0,m} + A_{1,m}x_t + A_{1,\lambda,m}x_{\lambda,t} + A_{2,x,m}\sigma^2_{x,t} + A_{2,c,m}\sigma^2_{c,t} + A_{2,d,m}\sigma^2_{d,t} \tag{A.23}$$

the market return can be expressed as

$$\begin{aligned}r_{m,t+1} &= \kappa_{0,m} + A_{0,m}(\kappa_{1,m}-1) + \mu_d + \kappa_{1,m}A_{2,x,m}(1-\nu_x)(\varphi_x\bar{\sigma})^2 \\ &+ \kappa_{1,m}A_{2,c,m}(1-\nu_c)(\varphi_c\bar{\sigma})^2 + \kappa_{1,m}A_{2,d,m}(1-\nu_d)(\varphi_d\bar{\sigma})^2 + \{\phi + A_{1,m}(\kappa_{1,m}\rho - 1)\}x_t \\ &+ (\kappa_{1,m}\rho_\lambda - 1)A_{1,\lambda,m}x_{\lambda,t} + A_{2,x,m}(\kappa_{1,m}\nu_x - 1)\sigma^2_{x,t} + A_{2,c,m}(\kappa_{1,m}\nu_c - 1)\sigma^2_{c,t} \\ &+ A_{2,d,m}(\kappa_{1,m}\nu_d - 1)\sigma^2_{d,t} + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1} + \kappa_{1,m}A_{1,m}\sigma_{x,t}\eta_{x,t+1} + \kappa_{1,m}A_{1,\lambda,m}\sigma_\lambda\eta_{\lambda,t+1} \\ &+ \kappa_{1,m}A_{2,x,m}\sigma_{w_x}w_{x,t+1} + \kappa_{1,m}A_{2,c,m}\sigma_{w_c}w_{c,t+1} + \kappa_{1,m}A_{2,d,m}\sigma_{w_d}w_{d,t+1}.\end{aligned} \tag{A.24}$$

Given the solution for A's, A_m's can be derived as follows:

$$\begin{aligned}
A_{0,m} &= \frac{A_{0,m}^{1st} + A_{0,m}^{2nd}}{1-\kappa_{1,m}} \\
A_{1,m} &= \frac{\phi - \frac{1}{\psi}}{1-\kappa_{1,m}\rho} \\
A_{1,\lambda,m} &= \frac{\rho_\lambda}{1-\kappa_{1,m}\rho_\lambda} \\
A_{2,x,m} &= \frac{\frac{1}{2}\left\{(\theta-1)\kappa_1 A_1 + \kappa_{1,m}A_{1,m}\right\}^2 + (\theta-1)(\kappa_1\nu_x - 1)A_{2,x}}{1-\kappa_{1,m}\nu_x} \\
A_{2,c,m} &= \frac{\frac{1}{2}(\pi-\gamma)^2 + (\theta-1)(\kappa_1\nu_c - 1)A_{2,c}}{1-\kappa_{1,m}\nu_c} \\
A_{2,d,m} &= \frac{\frac{1}{2}}{1-\kappa_{1,m}\nu_d},
\end{aligned} \tag{A.25}$$

where

$$\begin{aligned}
A_{0,m}^{1st} &= \theta\log\delta + (\theta-1)\left\{\kappa_0 + A_0(\kappa_1-1) + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2\right\} \\
&- \gamma\mu + \kappa_{0,m} + \mu_d + \kappa_{1,m}A_{2,x,m}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_{1,m}A_{2,c,m}(1-\nu_c)(\varphi_c\bar{\sigma})^2 \\
&+ \kappa_{1,m}A_{2,d,m}(1-\nu_d)(\varphi_d\bar{\sigma})^2 \\
A_{0,m}^{2nd} &= \frac{1}{2}\left(\kappa_{1,m}A_{2,x,m}\sigma_{w_x} + (\theta-1)\kappa_1 A_{2,x}\sigma_{w_x}\right)^2 + \frac{1}{2}\left(\kappa_{1,m}A_{2,c,m}\sigma_{w_c} + (\theta-1)\kappa_1 A_{2,c}\sigma_{w_c}\right)^2 \\
&+ \frac{1}{2}\left(\kappa_{1,m}A_{2,d,m}\sigma_{w_d}\right)^2 + \frac{1}{2}\left(\kappa_{1,m}A_{1,\lambda,m}\sigma_\lambda + (\theta-1)\kappa_1 A_{1,\lambda}\sigma_\lambda + \theta\sigma_\lambda\right)^2.
\end{aligned}$$

Rewrite the market-return equation (A.24) as

$$\begin{aligned}
&r_{m,t+1} - \mathbb{E}_t[r_{m,t+1}] \\
&= \beta_{m,c}\sigma_{c,t}\eta_{c,t+1} + \beta_{m,x}\sigma_{x,t}\eta_{x,t+1} + \beta_{m,d}\sigma_{d,t}\eta_{d,t+1} + \beta_{m,\lambda}\sigma_\lambda\eta_{\lambda,t+1} \\
&\quad + \beta_{m,w_x}\sigma_{w_x}w_{x,t+1} + \beta_{m,w_c}\sigma_{w_c}w_{c,t+1} + \beta_{m,w_d}\sigma_{w_d}w_{d,t+1},
\end{aligned}$$

where

$$\begin{aligned}
\beta_{m,c} &= \pi, \quad \beta_{m,x} = \kappa_{1,m}A_{1,m}, \quad \beta_{m,d} = 1, \quad \beta_{m,\lambda} = \kappa_{1,m}A_{1,\lambda,m}, \\
\beta_{m,w_x} &= \kappa_{1,m}A_{2,x,m}, \quad \beta_{m,w_c} = \kappa_{1,m}A_{2,c,m}, \quad \beta_{m,w_d} = \kappa_{1,m}A_{2,d,m}.
\end{aligned} \tag{A.26}$$

The risk premium for the dividend claim is

$$\begin{aligned}
&E_t(r_{m,t+1} - r_{f,t}) + \frac{1}{2}Var_t(r_{m,t+1}) \\
&= -Cov_t(m_{t+1}, r_{m,t+1}) \\
&= \beta_{m,x}\lambda_x\sigma_{x,t}^2 + \beta_{m,c}\lambda_c\sigma_{c,t}^2 + \beta_{m,\lambda}\lambda_\lambda\sigma_\lambda^2 + \beta_{m,w_x}\lambda_{w_x}\sigma_{w_x}^2 + \beta_{m,w_c}\lambda_{w_c}\sigma_{w_c}^2.
\end{aligned} \tag{A.27}$$

C.3 Risk-Free Rate

The model-driven equation for the risk-free rate is

$$\begin{aligned} r_{f,t} &= -\mathbb{E}_t\left[m_{t+1}\right] - \frac{1}{2}Var_t\left[m_{t+1}\right] \\ &= -\theta\log\delta - \mathbb{E}_t\left[x_{\lambda,t+1}\right] + \frac{\theta}{\psi}\mathbb{E}_t\left[g_{c,t+1}\right] + (1-\theta)\mathbb{E}_t\left[r_{c,t+1}\right] - \frac{1}{2}Var_t\left[m_{t+1}\right]. \end{aligned} \tag{A.28}$$

Subtract $(1-\theta)r_{f,t}$ from both sides and divide by θ,

$$r_{f,t} = -\log\delta - \frac{1}{\theta}\mathbb{E}_t\left[x_{\lambda,t+1}\right] + \frac{1}{\psi}\mathbb{E}_t\left[g_{c,t+1}\right] + \frac{(1-\theta)}{\theta}\mathbb{E}_t\left[r_{c,t+1} - r_{f,t}\right] - \frac{1}{2\theta}Var_t\left[m_{t+1}\right] \tag{A.29}$$

From (A.11) and (A.18)

$$r_{f,t} = B_0 + B_1 x_t + B_{1,\lambda}x_{\lambda,t} + B_{2,x}\sigma^2_{x,t} + B_{2,c}\sigma^2_{c,t},$$

where

$$B_1 = \frac{1}{\psi}, \quad B_{1,\lambda} = -\rho_\lambda, \quad B_{2,x} = -\frac{(1-\frac{1}{\psi})(\gamma-\frac{1}{\psi})\kappa_1^2}{2(1-\kappa_1\rho)^2}, \quad B_{2,c} = -\frac{1}{2}(\frac{\gamma-1}{\psi}+\gamma) \tag{A.30}$$

and

$$\begin{aligned} B_0 &= -\theta\log\delta - (\theta-1)\left\{\kappa_0 + (\kappa_1-1)A_0 + \kappa_1 A_{2,x}(1-\nu_x)(\varphi_x\bar{\sigma})^2 + \kappa_1 A_{2,c}(1-\nu_c)(\varphi_c\bar{\sigma})^2\right\} \\ &+ \gamma\mu - \frac{1}{2}\left\{(\theta-1)\kappa_1 A_{2,x}\sigma_{w_x}\right\}^2 - \frac{1}{2}\left\{(\theta-1)\kappa_1 A_{2,c}\sigma_{w_c}\right\}^2 - \frac{1}{2}\left\{((\theta-1)\kappa_1 A_{1,\lambda}+\theta)^2\sigma^2_\lambda\right\}. \end{aligned}$$

C.4 Linearization Parameters

For any asset, the linearization parameters are determined endogenously by the following system of equations:

$$\begin{aligned} \bar{pd}_i &= A_{0,i}(\bar{pd}_i) + \sum_{j\in\{c,x,d\}} A_{2,i,j}(\bar{pd}_i)\times(\varphi_j\bar{\sigma})^2 \\ \kappa_{1,i} &= \frac{\exp(\bar{pd}_i)}{1+\exp(\bar{pd}_i)} \\ \kappa_{0,i} &= \log(1+\exp(\bar{pd}_i)) - \kappa_{1,i}\bar{pd}_i. \end{aligned}$$

The solution is determined numerically by iteration until reaching a fixed point of $\bar{pd}_i$.

C.5 Deriving the Intertemporal Marginal Rate of Substitution (MRS)

We consider a representative-agent endowment economy modified to allow for time-preference shocks. The representative agent has Epstein and Zin (1989) recursive preferences and maximizes her lifetime utility

$$V_t = \max_{C_t} \left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta \big(\mathbb{E}_t [V_{t+1}^{1-\gamma}] \big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}$$

subject to budget constraint

$$W_{t+1} = (W_t - C_t) R_{c,t+1},$$

where W_t is the wealth of the agent, $R_{c,t+1}$ is the return on all invested wealth, γ is risk aversion, $\theta = \frac{1-\gamma}{1-1/\psi}$, and ψ is intertemporal elasticity of substitution. The ratio $\frac{\lambda_{t+1}}{\lambda_t}$ determines how agents trade off current versus future utility and is referred to as the time-preference shock (see Albuquerque, Eichenbaum, Luo, and Rebelo (2016)).

First conjecture a solution for $V_t = \phi_t W_t$. The value function is homogenous of degree 1 in wealth; it can now be written as

$$\phi_t W_t = \max_{C_t} \left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta \big(\mathbb{E}_t [(\phi_{t+1} W_{t+1})^{1-\gamma}] \big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}} \tag{A.31}$$

subject to

$$W_{t+1} = (W_t - C_t) R_{c,t+1}.$$

Epstein and Zin (1989) show that the above dynamic program has a maximum.

Using the dynamics of the wealth equation, we substitute W_{t+1} into (A.31) to derive

$$\phi_t W_t = \max_{C_t} \left[(1-\delta)\lambda_t C_t^{\frac{1-\gamma}{\theta}} + \delta (W_t - C_t)^{\frac{1-\gamma}{\theta}} \big(\mathbb{E}_t [(\phi_{t+1} R_{c,t+1})^{1-\gamma}] \big)^{\frac{1}{\theta}} \right]^{\frac{\theta}{1-\gamma}}. \tag{A.32}$$

At the optimum, $C_t = b_t W_t$, where b_t is the consumption-wealth ratio. Using (A.32) and shifting the exponent on the braces to the left-hand side, and dividing by W_t, yields

$$\phi_t^{\frac{1-\gamma}{\theta}} = (1-\delta)\lambda_t \left(\frac{C_t}{W_t} \right)^{\frac{1-\gamma}{\theta}} + \delta \left(1 - \frac{C_t}{W_t} \right)^{\frac{1-\gamma}{\theta}} \big(\mathbb{E}_t [(\phi_{t+1} R_{c,t+1})^{1-\gamma}] \big)^{\frac{1}{\theta}} \tag{A.33}$$

or simply

$$\phi_t^{\frac{1-\gamma}{\theta}} = (1-\delta)\lambda_t b_t^{\frac{1-\gamma}{\theta}} + \delta (1-b_t)^{\frac{1-\gamma}{\theta}} \big(\mathbb{E}_t [(\phi_{t+1} R_{c,t+1})^{1-\gamma}] \big)^{\frac{1}{\theta}}. \tag{A.34}$$

The first-order condition with respect to the consumption choice yields

$$(1-\delta)\lambda_t b_t^{\frac{1-\gamma}{\theta}-1} = \delta (1-b_t)^{\frac{1-\gamma}{\theta}-1} \big(\mathbb{E}_t [(\phi_{t+1} R_{c,t+1})^{1-\gamma}] \big)^{\frac{1}{\theta}}. \tag{A.35}$$

Plugging (A.35) into (A.34) yields

$$\phi_t = (1-\delta)^{\frac{\theta}{1-\gamma}} \lambda_t^{\frac{\theta}{1-\gamma}} \left(\frac{C_t}{W_t}\right)^{\frac{1-\gamma-\theta}{1-\gamma}} = (1-\delta)^{\frac{\psi}{\psi-1}} \lambda_t^{\frac{\psi}{\psi-1}} \left(\frac{C_t}{W_t}\right)^{\frac{1}{1-\psi}}. \tag{A.36}$$

The lifetime value function is $\phi_t W_t$, with the solution to ϕ_t stated above. This expression for ϕ_t is important: It states that the maximized lifetime utility is determined by the consumption-wealth ratio.

(A.35) can be rewritten as

$$(1-\delta)^{\theta} \lambda_t^{\theta} \left(\frac{b_t}{1-b_t}\right)^{-\frac{\theta}{\psi}} = \delta^{\theta} \mathbb{E}_t[(\phi_{t+1} R_{c,t+1})^{1-\gamma}]. \tag{A.37}$$

Consider the term $\phi_{t+1} R_{c,t+1}$:

$$\phi_{t+1} R_{c,t+1} = (1-\delta)^{\frac{\psi}{\psi-1}} \lambda_{t+1}^{\frac{\psi}{\psi-1}} \left(\frac{C_{t+1}}{W_{t+1}}\right)^{\frac{1}{1-\psi}} R_{c,t+1}. \tag{A.38}$$

After substituting the wealth constraint, $\frac{C_{t+1}}{W_{t+1}} = \frac{C_{t+1}/C_t}{W_t/C_t - 1} \cdot \frac{1}{R_{c,t+1}} = \frac{G_{t+1}}{R_{c,t+1}} \cdot \frac{b_t}{1-b_t}$, into the above expression, it follows that

$$\phi_{t+1} R_{c,t+1} = (1-\delta)^{\frac{\psi}{\psi-1}} \lambda_{t+1}^{\frac{\psi}{\psi-1}} \left(\frac{b_t}{1-b_t}\right)^{\frac{1}{1-\psi}} \left(\frac{G_{t+1}}{R_{c,t+1}}\right)^{\frac{1}{1-\psi}} R_{c,t+1}. \tag{A.39}$$

After some intermediate tedious manipulations,

$$\delta^{\theta} (\phi_{t+1} R_{c,t+1})^{1-\gamma} = \delta^{\theta} (1-\delta)^{\theta} \lambda_{t+1}^{\theta} \left(\frac{b_t}{1-b_t}\right)^{-\frac{\theta}{\psi}} G_{t+1}^{-\frac{\theta}{\psi}} R_{c,t+1}^{\theta}. \tag{A.40}$$

Taking expectations and substituting the last expression into (A.37) yields

$$\delta^{\theta} \mathbb{E}_t[\left(\frac{\lambda_{t+1}}{\lambda_t}\right)^{\theta} G_{t+1}^{-\frac{\theta}{\psi}} R_{c,t+1}^{\theta-1} R_{c,t+1}] = 1. \tag{A.41}$$

From here we see that the MRS in terms of observables is

$$M_{t+1} = \delta^{\theta} \left(\frac{\lambda_{t+1}}{\lambda_t}\right)^{\theta} G_{t+1}^{-\frac{\theta}{\psi}} R_{c,t+1}^{\theta-1}. \tag{A.42}$$

The log of MRS is

$$m_{t+1} = \theta \log \delta + \theta x_{\lambda,t+1} - \frac{\theta}{\psi} g_{t+1} + (\theta - 1) r_{c,t+1}, \tag{A.43}$$

where $x_{\lambda,t+1} = \log(\frac{\lambda_{t+1}}{\lambda_t})$.

D State-Space Representations of the Empirical Models

Below we describe the the state-space representation for the LRR model. The state-space representation for the cash-flow-only specifications can be obtained by eliminating the asset returns ($r_{m,t+1}$ and $r_{f,t}$) from the set of measurement equations.

D.1 Measurement Equations

In order to capture the correlation structure between the measurement errors at monthly frequency, we assumed in the main text that 12 months of consumption growth data are released at the end of each year. We will now present the resulting measurement equation. To simplify the exposition, we assume that the monthly consumption data are released at the end of the quarter (rather than at the end of the year). In the main text, the measurement equation is written as

$$y_{t+1} = A_{t+1}\Big(D + Zs_{t+1} + Z^v s^v_{t+1}(h_{t+1}, h_t) + \Sigma^u u_{t+1}\Big), \quad u_{t+1} \sim N(0, I). \tag{A.44}$$

The selection matrix A_{t+1} accounts for the deterministic changes in the vector of observables, y_{t+1}. Recall that monthly observations are available only starting in 1959:M1. For the sake of exposition, suppose prior to 1959:M1 consumption growth was available at quarterly frequency. We further assume that dividend growth data are always available in the form of time-aggregated quarterly data. Then (we are omitting some of the o superscripts for observed series that we used in the main text):

1. Prior to 1959:M1:

   (a) If $t+1$ is the last month of the quarter:

$$y_{t+1} = \begin{bmatrix} g^q_{c,t+1} \\ g^q_{d,t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix}, \quad A_{t+1} = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

   (b) If $t+1$ is not the last month of the quarter:

$$y_{t+1} = \begin{bmatrix} g^q_{d,t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix}, \quad A_{t+1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

2. From 1959:M1 to present:

(a) If $t+1$ is the last month of the quarter:

$$y_{t+1} = \begin{bmatrix} g_{c,t+1} \\ g_{c,t} \\ g_{c,t-1} \\ g^q_{d,t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix}, \quad A_{t+1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(b) If $t+1$ is not the last month of the quarter:

$$y_{t+1} = \begin{bmatrix} g^q_{d,t+1} \\ r_{m,t+1} \\ r_{f,t} \end{bmatrix}, \quad A_{t+1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

The relationship between observations and states (ignoring the measurement errors) is given by the approximate analytical solution of the LRR model described in Section C:

$$\begin{aligned} g_{c,t+1} &= \mu_c + x_t + \sigma_{c,t}\eta_{c,t+1} \\ g_{d,t+1} &= \mu_d + \phi x_t + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1} \\ r_{m,t+1} &= \{\kappa_{0,m} + (\kappa_{1,m}-1)A_{0,m} + \mu_d\} \\ &+ (\kappa_{1,m}A_{1,m})x_{t+1} + (\phi - A_{1,m})x_t + (\kappa_{1,m}A_{1,\lambda,m})x_{\lambda,t+1} - A_{1,\lambda,m}x_{\lambda,t} + \pi\sigma_{c,t}\eta_{c,t+1} + \sigma_{d,t}\eta_{d,t+1} \\ &+ (\kappa_{1,m}A_{2,x,m})\sigma^2_{x,t+1} - A_{2,x,m}\sigma^2_{x,t} + (\kappa_{1,m}A_{2,c,m})\sigma^2_{c,t+1} - A_{2,c,m}\sigma^2_{c,t} + (\kappa_{1,m}A_{2,d,m})\sigma^2_{d,t+1} - A_{2,d,m}\sigma^2_{d,t} \\ r_{f,t} &= B_0 + B_1 x_t + B_{1,\lambda}x_{\lambda,t} + B_{2,x}\sigma^2_{x,t} + B_{2,c}\sigma^2_{c,t} \\ & \eta_{i,t+1}, \eta_{\lambda,t+1}, w_{i,t+1} \sim N(0,1), \quad i \in \{c,x,d\}. \end{aligned} \tag{A.45}$$

In order to reproduce (A.45) and the measurement-error structure described in Sections 2.1 and 3.2, we define the vectors of states s_{t+1} and s^v_{t+1} as

$$s_{t+1} = \begin{bmatrix} x_{t+1} \\ x_t \\ x_{t-1} \\ x_{t-2} \\ x_{t-3} \\ x_{t-4} \\ \sigma_{c,t}\eta_{c,t+1} \\ \sigma_{c,t-1}\eta_{c,t} \\ \sigma_{c,t-2}\eta_{c,t-1} \\ \sigma_{c,t-3}\eta_{c,t-2} \\ \sigma_{c,t-4}\eta_{c,t-3} \\ \sigma_{\epsilon}\epsilon_{t+1} \\ \sigma_{\epsilon}\epsilon_{t} \\ \sigma_{\epsilon}\epsilon_{t-1} \\ \sigma_{\epsilon}\epsilon_{t-2} \\ \sigma_{\epsilon}\epsilon_{t-3} \\ \sigma_{\epsilon}\epsilon_{t-4} \\ \sigma^q_{\epsilon}\epsilon^q_{t+1} \\ \sigma^q_{\epsilon}\epsilon^q_{t} \\ \sigma^q_{\epsilon}\epsilon^q_{t-1} \\ \sigma^q_{\epsilon}\epsilon^q_{t-2} \\ \sigma_{d,t}\eta_{d,t+1} \\ \sigma_{d,t-1}\eta_{d,t} \\ \sigma_{d,t-2}\eta_{d,t-1} \\ \sigma_{d,t-3}\eta_{d,t-2} \\ \sigma_{d,t-4}\eta_{d,t-3} \\ x_{\lambda,t+1} \\ x_{\lambda,t} \end{bmatrix}, \quad s^v_{t+1} = \begin{bmatrix} \sigma^2_{x,t+1} \\ \sigma^2_{x,t} \\ \sigma^2_{c,t+1} \\ \sigma^2_{c,t} \\ \sigma^2_{d,t+1} \\ \sigma^2_{d,t} \end{bmatrix}. \tag{A.46}$$

It can be verified that the coefficient matrices D, Z, Z^v, and Σ^e are given by

$$
Z = \begin{bmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & $-\frac{1}{3}$ & $-\frac{1}{3}$ & $\frac{2}{3}$ & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & $\frac{\phi}{3}$ & $\frac{2\phi}{3}$ & ϕ & $\frac{2\phi}{3}$ & $\frac{\phi}{3}$ & $\frac{\pi}{3}$ & $\frac{2\pi}{3}$ & π & $\frac{2\pi}{3}$ & $\frac{\pi}{3}$ & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & $\frac{1}{3}$ & $\frac{2}{3}$ & 1 & $\frac{2}{3}$ & $\frac{1}{3}$ & 0 & 0 & 0 \\
$\mu_{r,1}$ & $\mu_{r,2}$ & 0 & 0 & 0 & 0 & 0 & $\mu_{r,3}$ & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & $\mu_{r,4}$ & 0 & 0 & 0 & 0 & $\mu_{r,5}$ & $\mu_{r,6}$ \\
0 & B_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & $B_{1,\lambda}$ & 0
\end{bmatrix}
$$

$$
Z^v = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
$\mu_{r,7}$ & $\mu_{r,8}$ & $\mu_{r,9}$ & $\mu_{r,10}$ & $\mu_{r,11}$ & $\mu_{r,12}$ & \\
0 & $B_{2,x}$ & 0 & $B_{2,c}$ & 0 & 0 &
\end{bmatrix},
\quad
D = \begin{bmatrix}
μ \\ μ \\ μ \\ μ \\ μ \\ $3\mu_d$ \\ $\mu_{r,0}$ \\ B_0
\end{bmatrix},
\quad
\Sigma^u = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & $\sigma_{d,\epsilon}^a$ & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & $\sigma_{f,\epsilon}$
\end{bmatrix}.
$$

The coefficients $\mu_{r,0}$ to $\mu_{r,12}$ are obtained from the solution of the LRR model:

$$\begin{bmatrix} \mu_{r,0} \\ \mu_{r,1} \\ \mu_{r,2} \\ \mu_{r,3} \\ \mu_{r,4} \\ \mu_{r,5} \\ \mu_{r,6} \end{bmatrix} = \begin{bmatrix} \kappa_{0,m} + A_{0,m}(\kappa_{1,m} - 1) + \mu_d \\ \kappa_{1,m} A_{1,m} \\ \phi - A_{1,m} \\ \pi \\ 1 \\ \kappa_{1,m} A_{1,\lambda,m} \\ -A_{1,\lambda,m} \end{bmatrix}, \quad \begin{bmatrix} \mu_{r,7} \\ \mu_{r,8} \\ \mu_{r,9} \\ \mu_{r,10} \\ \mu_{r,11} \\ \mu_{r,12} \end{bmatrix} = \begin{bmatrix} \kappa_{1,m} A_{2,x,m} \\ -A_{2,x,m} \\ \kappa_{1,m} A_{2,c,m} \\ -A_{2,c,m} \\ \kappa_{1,m} A_{2,d,m} \\ -A_{2,d,m} \end{bmatrix}.$$

D.2 State Transition Equations

Using the definition of s_{t+1} in (A.46), we write the state-transition equation as

$$s_{t+1} = \Phi s_t + v_{t+1}(h_t). \tag{A.47}$$

Conditional on the volatilities h_t, this equation reproduces the law of motion of the two persistent conditional mean processes

$$\begin{aligned} x_{t+1} &= \rho x_t + \sigma_{x,t} \eta_{x,t+1} \\ x_{\lambda,t+1} &= \rho_\lambda x_{\lambda,t} + \sigma_\lambda \eta_{\lambda,t+1} \end{aligned} \tag{A.48}$$

and it contains some trivial relationships among the measurement-error states. The matrices Φ and $v_{t+1}(h_t)$ are defined as

$$\Phi = \begin{bmatrix}
\rho & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho_\lambda & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0
\end{bmatrix}$$

and

$$v_{t+1}(h_t) = \begin{bmatrix} \sigma_{x,t}\eta_{x,t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma_{c,t}\eta_{c,t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma_\epsilon \epsilon_{t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma_\epsilon^q \epsilon_{t+1}^q \\ 0 \\ 0 \\ 0 \\ \sigma_{d,t}\eta_{d,t+1} \\ 0 \\ 0 \\ 0 \\ 0 \\ \sigma_\lambda \eta_{\lambda,t+1} \\ 0 \end{bmatrix}.$$

The law of motion of the three persistent conditional log volatility processes is given by

$$h_{t+1} = \Psi h_t + \Sigma_h w_{t+1}, \tag{A.49}$$

where

$$h_{t+1} = \begin{bmatrix} h_{x,t+1} \\ h_{c,t+1} \\ h_{d,t+1} \end{bmatrix}, \quad \Psi = \begin{bmatrix} \rho_{h_x} & 0 & 0 \\ 0 & \rho_{h_c} & 0 \\ 0 & 0 & \rho_{h_d} \end{bmatrix}$$

$$\Sigma_h = \begin{bmatrix} \sigma_{h_x}\sqrt{1-\rho_{h_x}^2} & 0 & 0 \\ 0 & \sigma_{h_c}\sqrt{1-\rho_{h_c}^2} & 0 \\ 0 & 0 & \sigma_{h_d}\sqrt{1-\rho_{h_d}^2} \end{bmatrix}, \quad w_{t+1} = \begin{bmatrix} w_{x,t+1} \\ w_{c,t+1} \\ w_{d,t+1} \end{bmatrix}.$$

We express

$$\sigma_{x,t} = \varphi_x \sigma \exp(h_{x,t}), \quad \sigma_{c,t} = \varphi_c \sigma \exp(h_{c,t}), \quad \sigma_{d,t} = \varphi_d \sigma \exp(h_{d,t}),$$

which delivers the dependence on h_t in the above definition of $v_{t+1}(\cdot)$. $\varphi_c = 1$ is normalized.

E Posterior Inference

E.1 Model With Asset Prices

To construct a posterior sampler for the LRR model (see Section 5 for estimation results), we use a particle-filter approximation of the likelihood function, constructed as follows. Our state-space representation, given the measurement equation (A.44) and the state transition equations (A.47) and (A.49), is linear conditional on the volatility states (h_{t+1}, h_t). The particle filter uses a swarm of particles $\{z_t^j, W_t^j\}_{j=1}^M$ to approximate

$$\mathbb{E}[h(z_t)|Y_{1:t}] \approx \frac{1}{M}\sum_{j=1}^{M} W_t^j h(z_t^j). \tag{A.50}$$

Throughout this section we omit the parameter vector Θ from the conditioning set. Here $h(\cdot)$ is an integrable function of z_t and the approximation $\approx$, under suitable regularity conditions, can be stated formally in terms of a strong law of large numbers and a central limit theorem. In general, z_t^j would be composed of h_t^j, h_{t-1}^j, and s_t^j. However, given that the state-space model is linear conditional on (h_t, h_{t-1}), we can replace s_t^j by

$$\left[vec\big(\mathbb{E}[s_t|h_t^j, h_{t-1}^j, Y_{1:t}]\big),\ vech\big(Var[s_t|h_t^j, h_{t-1}^j, Y_{1:t}]\big)\right]',$$

where $vech(\cdot)$ stacks the non-redundant elements of a symmetric matrix. The use of the vector of conditional means and covariance terms for s_t in the definition of the particle z_t^j leads to a variance reduction in the particle filter approximation of the likelihood function. The implementation of the particle filter is based on Algorithm 13 in Herbst and Schorfheide (2015). The particle-filter approximation of the likelihood function is embedded into a fairly standard random walk Metropolis-Hastings algorithm (see Chapter 9 of Herbst and Schorfheide (2015)).

E.2 Models Without Asset Prices

The estimation of the cash-flow only models in Sections 2 and 3 is considerably easier because the volatility states do not affect the conditional means of the observables. As before in the model with asset prices, the state variables are the model-implied monthly cash flows and the the latent volatility processes $h_{i,t}$. Let Θ_{cf} denote the parameters that denote that cash-flow processes, Θ_h the parameters that control the evolution of the volatility processes, and H^T the sequence of latent volatilities.

The MCMC algorithm iterates over three conditional distributions: First, a Metropolis-Hastings step is used to draw from the posterior of Θ_{cf} conditional on $\big(Y, (H^T)^{(s)}, \Theta_h^{(s-1)}\big)$. Second, we

draw the sequence of stochastic volatilities H^T conditional on $\left(Y, \Theta_{cf}^{(s)}, \Theta_h^{(s-1)}\right)$ using the algorithm developed by Kim, Shephard, and Chib (1998). It consists of transforming a nonlinear and non-Gaussian state space form into a linear and approximately Gaussian one, which allows the use of simulation smoothers such as those of Carter and Kohn (1994) to recover estimates of the residuals $\eta_{i,t}$. Finally, we draw from the posterior of the coefficients of the stochastic volatility processes, Θ_h, conditional on $\left(Y, H^{T(s)}, \Theta_{cf}^{(s)}\right)$.

F Supplementary Figures and Tables

This section provides supplementary empirical results that are referenced in the main paper.

- Table A-1: provides estimates of alternative specifications of the consumption growth model considered in Section 2.2 and supplements Table 1.
- Table A-2: provides estimates of a bivariate cash-flow model in which consumption and dividends are cointegrated. These estimates are referenced in the part *Cointegration of Dividends and Consumption.* of Section 3.2.
- Table A-3: This table appeared in the main text of an earlier version of the paper. Comparing the estimates of ρ from Table 6 based on cash-flow data only to the estimate obtained in Table 7 by estimating the LRR model based on cash flow and asset return data, we observed that the posterior mean increases from 0.94 and 0.95, respectively, to 0.99 once asset returns are included. To assess the extent to which the increase in ρ leads to a decrease in fit of the consumption growth process, we re-estimate model (4) conditional on various choices of ρ between 0.90 and 0.99 and re-compute the marginal data density for consumption growth. The results are summarized in the table. The key finding is that the drop in the marginal data density by changing ρ from $\hat{\rho}$ to 0.99 is small, indicating that there essentially is no tension between the parameter estimates obtained with and without asset prices.
- Figure A-1: contains further posterior predictive checks for the R^2 values associated with consumption and return predictability regressions. It supplements Figure 9 in Section 5.3 and shows model how the model-implied predictive distribution of the R^2's changes as different sources of risk are switched off. These results are mentioned in the main paper in Footnote 25.
- Figure A-2: This figure appeared in the main text of an earlier version of the paper. It examines the model's implication with respect to the long horizon correlation between consumption growth (dividend growth) and returns – that is the H-th horizon correlation

$$corr(\sum_{h=1}^{H} r_{m,t+h}, \sum_{h=1}^{H} \Delta c_{t+h}).$$

  Our model performs well along this dimension. Under the "Benchmark" specifications (all shocks are active), the 10-year consumption growth and 10-year return have a correlation of 0.3, but with a very wide credible interval that encompasses -0.2 to 0.7, which importantly contains the data estimate. The analogous correlation credible interval for dividend growth ranges from 0 to 0.8, with the data at 0.4 and again very close to the model median estimate.

It is noteworthy that these correlation features are primarily driven by "Growth and Volatility Risks." Albuquerque, Eichenbaum, Luo, and Rebelo (2016) highlight that preference shocks improve the LRR model-performance for these long horizon correlations. The "Preference Risk" subplots provide the correlations when all shocks except $x_{\lambda,t}$ are shutdown. These plots show that the preference shocks improve fit by generating lower credible intervals for consumption, yet deteriorate fit by generating way too large long horizon correlations for dividends.

Table A-1: Posterior Median Estimates of Consumption Growth Processes

	Prior Distribution				Posterior Estimates								
					State-Space Model / Measurement Error Specification							IID	ARMA
					M&A	No ME	No ME AR(2)	M	M $\rho_\epsilon \neq 0$	M $\rho_\eta \neq 0$	M NoAveOut		(1,2)
	Distr.	5%	50%	95%	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
μ_c	N	-.007	.0016	.0100	.0016	.0016	.0016	.0016	.0016	.0016	.0016	.0016	.0016
ρ	U	-.90	0	.90	.918	-.287	-.684	.918	.918	.919	.919	-	.913
ρ_2	U	-.90	0	.90	-	-	-.353	-	-	-	-	-	-
φ_x	U	.05	0.5	.95	.681	-		.669	.704	.644	.681	-	-
	U	.1	1.0	1.9	-	1.12	.482	-	-	-	-	-	-
σ	IG	.0008	.0019	.0061	.0018	.0022	.0027	.0018	.0017	.0019	.0018	.0033	.0032
σ_ϵ	IG	.0008	.0019	.0061	.0018	-		.0018	.0019	.0018	.0018	-	-
σ_ϵ^a	IG	.0007	.0029	.0386	.0011	-	-	-	-	-	-	-	-
ρ_ϵ	U	-.90	0	.90	-	-	-	-	.060	-	-	-	-
ρ_η	U	-.90	0	.90	-	-	-	-	-	-.046	-	-	-
ζ_1	N	-8.2	0	8.2	-	-	-	-	-	-	-	-	-1.14
ζ_2	N	-8.2	0	8.2	-	-	-	-	-	-	-	-	.302
$\ln p(Y)$					2887.1	2870.8	2870.3	2886.2	2883.9	2885.8	2886.5	2863.2	2884.0

Notes: The estimation sample is from 1959:M2 to 2014:M12. We denote the persistence of the growth component x_t by ρ (and ρ_2 if follows an AR(2) process), the persistence of the measurement errors by ρ_ϵ, and the persistence of $\eta_{c,t}$ by ρ_η. We report posterior median estimates for the following measurement error specifications of the state-space model: (1) monthly and annual measurement errors (M&A); (2) no measurement errors (no ME); (3) no measurement errors with AR(2) process for x_t (no ME AR(2)); (4) monthly measurement errors (M); (5) serially correlated monthly measurement errors (M, $\rho_\epsilon \neq 0$); (6) serially correlated consumption shocks $\eta_{c,t}$ (M, $\rho_\eta \neq 0$, $\rho > \rho_\eta$); (7) monthly measurement errors that do not average out at annual frequency (M, NoAveOut). In addition we report results for the following models: (8) consumption growth is *iid*; (9) consumption growth is ARMA(1,2).

Table A-2: Posterior Estimates: Cointegration of Consumption and Dividends

		Prior			Posterior		
	Distr.	5%	50%	95%	5%	50%	95%
				Consumption			
ρ	U	-.9	0	.9	.907	.951	.984
φ_x	U	.05	.50	.95	.314	.515	.946
σ	IG	.0008	.0019	.0061	.0022	.0028	.0034
ρ_{h_c}	N^T	.27	.80	.999	.976	.992	.999
$\sigma^2_{h_c}$	IG	.0013	.0043	.0283	.0012	.0037	.0117
				Dividends			
ϕ_{dc}	U	-9.0	0	9.0	-7.10	-5.66	-4.64
ρ_s	U	-.9	0	.9	.997	.998	.999
φ_s	U	15	150	285	86.5	148.0	241.2
ρ_{h_s}	N^T	.27	.80	.999	.995	.998	.999
$\sigma^2_{h_s}$	IG	.0007	.0029	.0392	.0008	.0014	.0028
				Measurement Errors			
σ_ϵ	IG	.0008	.0019	.0061	.0010	.0012	.0015
σ^a_ϵ	IG	.0008	.0029	.0387	.0005	.0044	.0109
$\sigma^a_{d,\epsilon}$	IG	.0008	.0029	.0387	-	.10	-

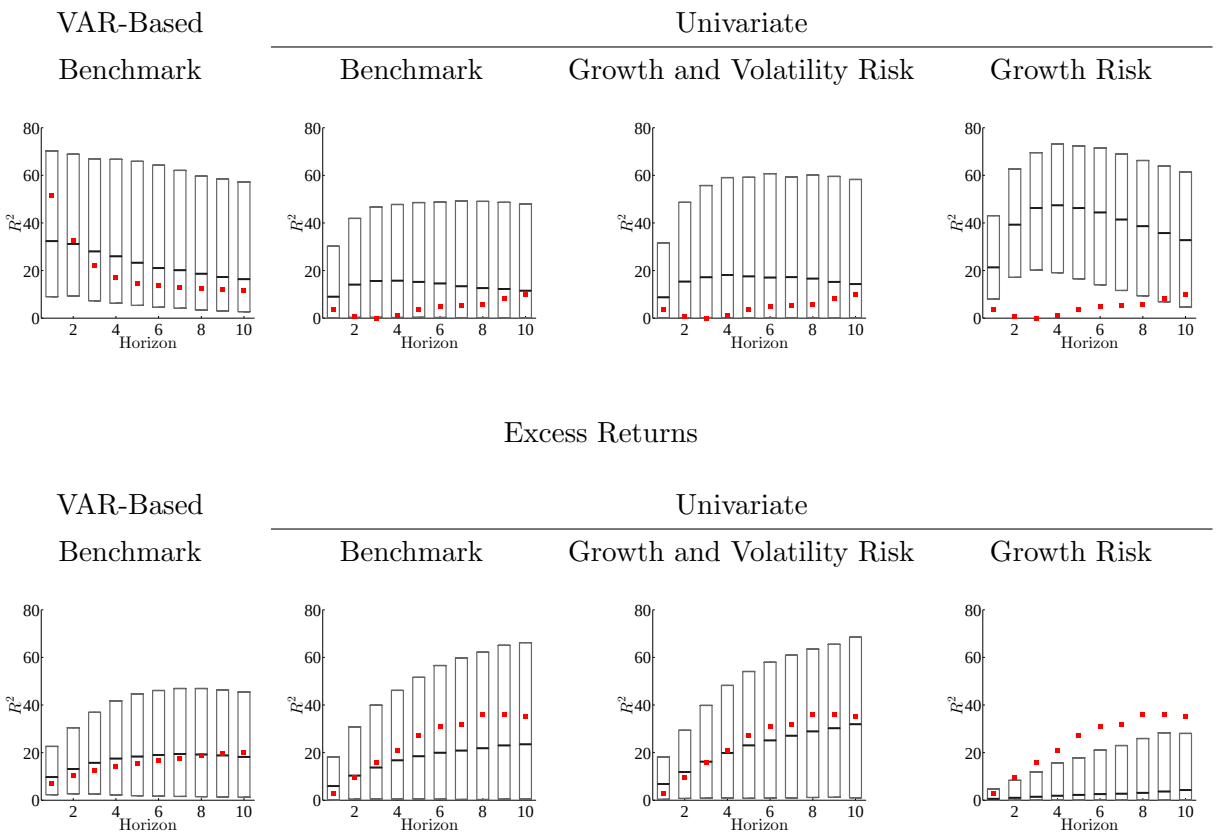
Notes: We utilize the mixed-frequency approach in the estimation: For consumption we use annual data from 1930 to 1959 and monthly data from 1960:M1 to 2014:M12; we use monthly dividend annual growth data from 1930:M1 to 2014:M12. For consumption we adopt the measurement error model of Section 2.1. We allow for annual consumption measurement errors ϵ^a_t during the periods from 1930 to 1948. We impose monthly measurement errors ϵ_t when we switch from annual to monthly consumption data from 1960:M1 to 2014:M12. We fix $\mu_c = 0.0016$ and $\mu_d = 0.0010$ at their sample averages. Moreover, we also fix the measurement error variances $(\sigma^a_{d,\epsilon})^2$ and $(\sigma_{f,\epsilon})^2$ at 1% of the sample variance of dividend growth and the risk-free rate, respectively. N, N^T, G, IG, and U denote normal, truncated (outside of the interval $(-1, 1)$) normal, gamma, inverse gamma, and uniform distributions, respectively.

Table A-3: Marginal Data Densities for Consumption Growth Model

Estimation	Fixed ρ					Estimated ρ
Sample	0.90	0.94	0.95	0.97	0.99	
1959-2014	2925.9	2935.9	2935.5	2934.8	2927.5	2930.1 ($\hat{\rho} = 0.95$)
1930-2014	2912.7	2914.2	2913.3	2912.1	2909.3	2909.9 ($\hat{\rho} = 0.94$)

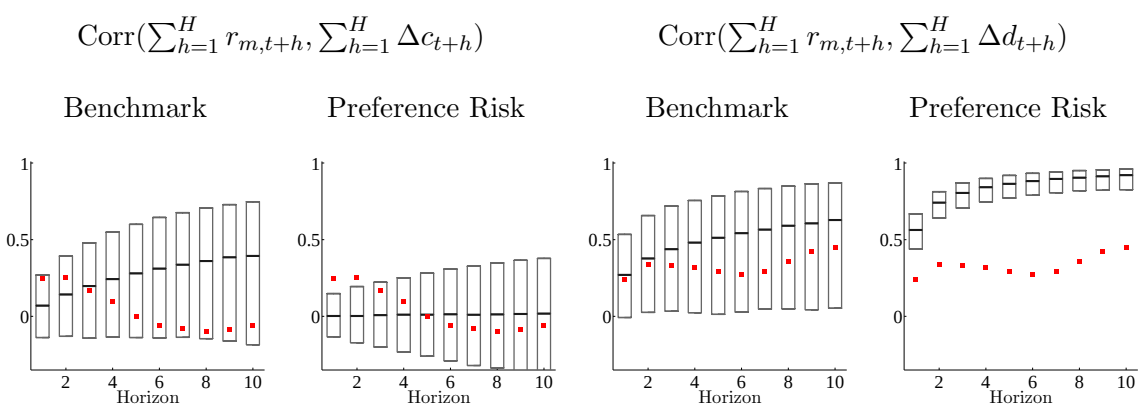
Notes: We estimate the consumption-only model (4) conditional on various choices of ρ ("Fixed ρ") and compute marginal data densities. We also report the marginal data densities for the estimated values of ρ ("Estimated ρ") based on the posterior mean estimates (in parentheses) from Table 3.

Figure A-1: Predictability Checks



Notes: We fix the parameters at their posterior median estimates. The red squares represent R^2 values obtained from the actual data. The boxes represent 90% posterior predictive intervals and the horizontal lines represent medians. The "Benchmark" case is based on simulations with all five state variables x_t, $x_{\lambda,t}$, $\sigma^2_{x,t}$, $\sigma^2_{c,t}$, and $\sigma^2_{d,t}$; "Growth and Volatility Risk" is based on x_t and $\sigma^2_{x,t}$ only; "Growth Risk" is based on x_t only. The horizon is measured in years. The VAR-Based R^2s are constructed as in Hodrick (1992).

Figure A-2: Correlation between Market Return and Cash-Flow Growth Rates



Notes: We fix the parameters at their posterior median estimates. The "Benchmark" case is based on simulations with all five state variables x_t, $x_{\lambda,t}$, $\sigma^2_{x,t}$, $\sigma^2_{c,t}$, and $\sigma^2_{d,t}$; "Preference Risk" is based on $x_{\lambda,t}$ only.