From Which Consumption-Based Asset Pricing Models Can Investors Profit? Evidence from Model-Based Priors*

Mathias S. Kruttli†

First draft: February 11, 2014

This draft: July 7, 2016

Abstract

This paper compares consumption-based asset pricing models based on the forecasting performance of investors who use the respective asset pricing theories to predict the equity premium with valuation ratios. Model-based priors are derived from three prominent consumption-based asset pricing models: Habit Formation, Long Run Risk, and Prospect Theory. A simple Bayesian framework is proposed through which the investors impose these model-based priors on the parameters of their predictive models. An investor whose prior beliefs are rooted in the Long Run Risk model achieves more accurate forecasts overall. The greatest difference in performance occurs during the bull market of the late 1990s. During this period, the weak predictability of the equity premium implied by the Long Run Risk model helps the investor to not prematurely anticipate falling stock prices.

JEL classification: G11, G12, G17

Keywords: Bayesian econometrics, consumption-based asset pricing, return predictability

*I thank Nicholas Barberis, John Campbell, Lubos Pastor, Andrew Patton, Tarun Ramadorai, Kevin Sheppard, Dimitri Vayanos, Jessica Wachter, Missaka Warusawitharana, Sumudu Watugala, Ivo Welch, Mungo Wilson, and seminar participants at the Oxford-Man Institute of Quantitative Finance, Saïd Business School, the 10th Swiss Economists Abroad End-of-Year Conference 2015, and the Federal Reserve Board of Governors for useful comments. The analysis and conclusions set forth are those of the author and do not indicate concurrence by other members of the Board of Governors of the Federal Reserve System. Part of this paper was completed when Mathias S. Kruttli was at the University of Oxford Economics Department. An earlier version of this paper was named "Model-Based Priors for Predicting the Equity Premium" (2014).

†Mathias S. Kruttli is at the Board of Governors of the Federal Reserve System and the Oxford-Man Institute of Quantitative Finance. Email: mathias.s.kruttli@frb.gov.

1 Introduction

Predicting aggregate stock returns has been of great interest to finance practitioners and academic finance economists alike. For an investor, knowing whether the equity premium is predictable is crucial for portfolio allocation decisions. An extensive literature uses a variety of variables to explain the time-variation of returns (see, for example, Campbell and Shiller (1988); Campbell (1987); Fama and French (1988 and 1989); Baker and Wurgler (2000); Lettau and Ludvigson (2001); Polk, Thompson, and Vuolteenaho (2006); Welch and Goyal (2008); Li, Ng, and Swaminathan (2013); Kruttli, Patton, and Ramadorai (2015)). Valuation ratios were initially found to have predictive power when forecasting the equity premium, but the set of forecasting variables has since been extended with variables such as, corporate payout, implied cost of capital, and yields on bonds and Treasury securities.

Welch and Goyal (2008) provide a comprehensive analysis of the in-sample and out-of-sample (OOS) predictive power of the major variables and question whether the equity premium is predictable OOS. Campbell and Thompson (2008) further investigate these findings by imposing economic restrictions when estimating the predictive model. They apply sign restrictions on the parameter estimates of the predictive model and a non-negativity restriction on the forecast of the equity premium. Campbell and Thompson (2008) find that through these restrictions, a real-time investor could profitably forecast the equity premium.

This paper imposes novel economic constraints derived from consumption-based asset pricing models. I propose a simple Bayesian econometric framework to implement these economic constraints as prior distributions on the parameters of single-variable predictive regressions. These prior distributions are named model-based priors. My approach relates to the macroeconometric literature, in which prior distributions from dynamic stochastic equilibrium models are imposed on vector autoregressions to predict macroeconomic variables (see, for example, Del Negro and Schorfheide (2011)). The three consumption-based asset pricing models that act as sources for the model-based priors are the Habit Formation (HF) model (see Campbell and Cochrane (1999)), the Prospect Theory (PT) model (see Barberis, Huang, and Santos (2001)), and the Long Run Risk (LRR) model (see Bansal and Yaron (2004)). All three models propose different theories that can explain the equity premium puzzle (Mehra and Prescott (1985)). The model-based priors allow me to assess whether an investor could have profited from knowing the respective theories and their implications for the predictability of the equity premium inherent in these consumption-based asset pricing models. I assume that an investor who forecasts the equity premium with valuation ratios has a prior belief about the parameter estimates of the predictive model that stems from one of the asset pricing models. The investor then updates her beliefs with empirical data and predicts the equity premium OOS based on the posterior parameter estimates. To my knowledge, prior distributions derived from asset pricing models have not been previously explored for the purpose of forecasting returns OOS. Unlike other papers in the equity premium prediction literature, the focus of this paper is to compare the performances of the model-based priors from the three asset pricing models with each other. Comparing the accuracy of the forecasts provides an assessment of how useful the theories developed by the asset pricing models are for a finance practitioner who attempts to time her investments in the aggregate stock market. This novel way of comparing consumption-based asset pricing models leads to insights that are not obtained when matching empirical data moments with model-based moments from Monte Carlo simulations, as is generally done.

Several other papers in this growing literature also make use of economically motivated parameter constraints for predicting the equity premium and implement them through a form of Bayesian framework. Pastor and Stambaugh (2009) employ a prior that implies a negative correlation between expected and unexpected return shocks. Shanken and Tamayo (2012) consider prior beliefs on mispricing as a driver of predictability and on the risk-return tradeoff. Pettenuzzo, Timmermann, and Valkanov (2014) propose a Bayesian methodology that imposes a non-negative equity premium and bounds on the conditional Sharpe ratio. Their constraints lead to forecasts of the equity premium that are substantially more accurate. Wachter and Warusawitharana (2009) model skepticism of an investor over the predictability of the equity premium as an informative prior over the R^2 and show that a skeptical investor achieves better forecasts. Wachter and Warusawitharana (2015) analyze whether an investor who is skeptical about the existence of equity premium predictability would update her prior and conclude that the equity premium is predictable when being confronted with historical data. Other Bayesian studies consider uncertainty about the predictive model parameters through uninformative priors (see, for example, Stambaugh (1999), Barberis (2000), and Brandt, Goyal, Santa-Clara, and Stroud (2005)) or investigate how parameter uncertainty affects the long run predictive variance (see, for example, Pastor and Stambaugh (2012) and Avramov, Cederburg, and Lucivjanska (2016)).

My sample comprises data from 1926 to 2014. I compare the predictive performance of hypothetical investors who had access to the three asset pricing models from 1926 onward and who try to time the market by forecasting the equity premium with the dividend-price ratio and the dividend yield. For my baseline results, the calibration of the asset pricing models is the same as proposed by the respective authors.[1] However, my results are robust to recalibrating the asset pricing models over a time period that has no overlap with the OOS period. I find a sharp distinction between the performance of the LRR model-based priors and the model-based priors derived from the HF and PT models. The LRR model-based priors perform particularly well from 1980 onward. The HF and PT model-based priors result in more accurate forecasts up to the 1980s. Over the whole data sample, an investor armed with the knowledge of the LRR model would have generally outperformed

[1]Because the authors of the asset pricing models use almost identical data sets for the calibration of their respective models, this assumption should not lead to distorted results when comparing the performance of the respective priors.

investors whose prior beliefs about the predictability of the equity premium were rooted in the HF or PT model. The differences in performance hold when comparing both the accuracy of the forecasts and the utility gains achieved by the investors. The key to the strong performance of the LRR prior over the total sample period is the bull market of the late 1990s, when low valuation ratios predicted negative stock returns that did not materialize for several years. The LRR model implies a lower predictive power of valuation ratios than the other two asset pricing models. Hence, an investor who uses the LRR model as guidance for her investment choices is reluctant to conclude that low valuation ratios imply an immediate decline in stock prices. This reluctance improves her forecast performance during the late 1990s, and this effect dominates less accurate forecasts of the LRR priors during episodes when the predictive power of valuation ratios was stronger, for example, in the 1970s.

The differences in forecast accuracy between the three asset pricing models are economically significant. I find that an investor with mean-variance preferences who allocates her portfolio based on equity premium forecasts would on average be willing to pay 26 basis points per year to have access to the LRR model-based priors instead of the HF model-based priors. For the PT priors, 67 basis points per year are on average necessary to achieve the same utility level as with the LRR priors. The limited equity premium predictability that the LRR model implies is often considered a shortcoming (see, for example, Beeler and Campbell (2012)). This paper shows that from the viewpoint of an investor, the weak predictability of the LRR model can be an advantage. Thus, the findings in this paper contribute to the current debates in consumption-based asset pricing and equity premium prediction.

The structure of this paper is as follows. Section 2 explains the Bayesian methodology used to impose the model-based priors. Section 3 reports the data used and the results. Section 4 discusses the utility gains that an investor with mean-variance preferences achieves when implementing the model-based priors. Section 5 analyzes the robustness of the results. Section 6 concludes the paper.

2 Methodology

This section describes how I impose economic constraints on the single-variable predictive regressions through priors derived from consumption-based asset pricing models and how these models are simulated to obtain the priors.

2.1 Equity premium prediction model

The equity premium at time $t+1$ is denoted by r_{t+1} and is defined as the rate of return on the stock market in excess of the prevailing short-term interest rate. As is common in the equity premium prediction literature, r_{t+1} is regressed on a constant and a predictor, x_t, which is lagged by one period:

$$r_{t+1} = \beta_0 + \beta_1 x_t + \epsilon_{t+1}, \text{ where } \epsilon_{t+1} \sim N(0, \sigma_\epsilon^2). \tag{1}$$

The OOS predictions of the equity premium are generated through recursive forecasts (see, for example, Campbell and Thompson (2008), Welch and Goyal (2008), and Pettenuzzo et al. (2014)). Hence, all available observations up to period t are used to estimate the model in equation (1). Based on the resulting estimates of the parameters $\beta = [\beta_0,\ \beta_1]'$ and σ_ϵ^2, and by observing x_t, one can forecast the equity premium in $t+1$. The predicted equity premium is denoted by $\hat{r}_{t+1}$. Because observations after $t+1$ are not used to estimate β, a real-time investor who forecasts the equity premium can implement this procedure. If no model-based priors are imposed, the parameters can be estimated via ordinary least squares (OLS). A common benchmark for a predictor in the equity premium literature is the historical average model, which forecasts that the equity premium will be next period what it has been on average in the past (β_1 in equation (1) is set to zero).

2.2 Model-based priors

An investor who wants to make use of the theoretical insights of a consumption-based asset pricing model can impose economic constraints on β derived from the asset pricing model. These model-based constraints are best imposed via Bayesian techniques. I assume that the investor's prior belief is that β and σ_ϵ^2 take the values implied by the asset pricing model. She then updates her belief through empirical data.

The prior distribution of the parameters in equation (1) — that is β and σ_ϵ^2 — is assumed to be Gamma-Normal (see, for example, Koop (2003) and Pettenuzzo et al. (2014)). The prior distribution is then given by

$$\beta \sim N\left(\underline{\beta}, \underline{V}\right), \ \sigma_\epsilon^{-2} \sim G\left(\sigma_\epsilon^{*-2}, \underline{v}(t-1)\right). \tag{2}$$

The mean and the variance of the Normal prior distribution are specified as

$$\underline{\beta} = \begin{bmatrix} \beta_0^* \\ \beta_1^* \end{bmatrix}, \ \underline{V} = \begin{bmatrix} \lambda^2 \sigma_{r,t}^2 & 0 \\ 0 & \lambda \sigma_{r,t}^2 / \sigma_{x,t}^2 \end{bmatrix}, \tag{3}$$

where β_0^* and β_1^* are the coefficient values implied by the consumption-based asset pricing model. The parameter λ is exogenously chosen and is weakly positive. If λ is large, the prior is loose. If λ is equal to zero, the prior is dogmatic. I set $\lambda = 1$ for the benchmark case. Section 5 discusses the robustness of my results for different choices of λ. The sample moments $\sigma_{r,t}^2$ and $\sigma_{x,t}^2$ are scaling factors, which ensures that the results are comparable for different predictors and forecast frequencies. Such scaling factors are commonly used in Bayesian macroeconometrics and date back to Litterman (1986). The sample moments are given by

$$\sigma_{r,t}^2 = \frac{1}{t-2} \sum_{\tau=2}^{t} (r_\tau - \bar{r}_t)^2, \ \bar{r}_t = \frac{1}{t-1} \sum_{\tau=2}^{t} r_\tau \tag{4}$$

and

$$\sigma_{x,t}^2 = \frac{1}{t-2}\sum_{\tau=1}^{t-1}(x_\tau - \bar{x}_t)^2,\ \overline{x}_t = \frac{1}{t-1}\sum_{\tau=1}^{t-1} x_\tau. \tag{5}$$

The Gamma distribution parametrization follows Koop (2003) by specifying the distribution with mean σ_ϵ^{*-2} and degrees of freedom $\underline{v}(t-1)$, where σ_ϵ^{*-2} is derived from the consumption-based asset pricing model. The tightness of the prior is controlled by $\underline{v}$, which is strictly positive. A large $\underline{v}$ corresponds to a tight prior, and a small $\underline{v}$ corresponds to a diffuse prior. The benchmark case sets $\underline{v}$ to 0.1, but my results are robust for a tighter or a more diffuse prior on σ_ϵ^{-2} (see Section 5).

2.3 Posterior distribution

The model-based prior distributions yield conditional posterior distributions for β and σ_ϵ^{-2}. I draw from these two conditional distributions through a Gibbs sampler. The conditional posterior distribution for β is

$$\beta|\sigma_\epsilon^{-2}, \mathcal{I}_t \sim N\left(\bar{\beta}, \overline{V}\right), \tag{6}$$

where

$$\overline{V} = (\underline{V}^{-1} + \sigma_\epsilon^{*-2}X'X)^{-1},\ \bar{\beta} = \overline{V}(\underline{V}^{-1}\underline{\beta} + \sigma_\epsilon^{*-2}X'R), \tag{7}$$

X is a $t-1 \times 2$ matrix with rows $[1\ x_\tau]$ for $\tau = 1, ..., t-1$, and R is a $t-1 \times 1$ vector with elements r_τ for $\tau = 2, ..., t$. The information set at time t is denoted by $\mathcal{I}_t$. The conditional posterior distribution for $\sigma_\epsilon{}^{-2}$ takes the form

$$\sigma_\epsilon^{-2}|\beta, \mathcal{I}_t \sim G\left(\bar{s}^{-2}, \bar{v}\right), \tag{8}$$

where

$$\bar{v} = \underline{v} + (t-1), \text{ and } \bar{s}^2 = \frac{\sum_{\tau=2}^{t}(r_\tau - \beta_0 - \beta_1 x_{\tau-1})^2 + \sigma_\epsilon^{*2}\underline{v}(t-1)}{\bar{v}}. \tag{9}$$

Through the Gibbs sampling algorithm with J iterations, we obtain a series of draws for each of the parameters denoted by $\{\beta^j\}$ and $\{\sigma_\epsilon^{-2,j}\}$ for $j = 1, ..., J$. These simulated series can then be used to draw from the predictive return distribution

$$p(r_{t+1}|\mathcal{I}_t) = \int_{\beta,\sigma_\epsilon^{-2}} p(r_{t+1}|\beta, \sigma_\epsilon^{-2}, \mathcal{I}_t)p(\beta, \sigma_\epsilon^{-2}|\mathcal{I}_t)d\beta d\sigma_\epsilon^2, \tag{10}$$

which yields $\{r_{t+1}^j\}$ for $j = 1, ..., J$. The point forecast for the equity premium in period $t+1$ is given by the mean of the sampled distribution

$$\hat{r}_{t+1}^m = \frac{1}{J}\sum_{j=1}^{J} r_{t+1}^j. \tag{11}$$

2.4 Deriving priors from asset pricing models

I next describe how the prior means $\beta^* = [\beta_0^*,\ \beta_1^*]'$ and σ_ϵ^{*-2} are derived from the three consumption-based asset pricing models: HF, LRR, and PT. All three models specify a log consumption and a log dividend growth process. By simulating random shocks, time series of consumption growth and dividend growth are generated, based on which I solve the models for the log equity premium, the dividend-price ratio, and the dividend yield. The dividend-price ratio is the difference between the log of dividends and the log of prices, and the dividend yield is the difference between the log of dividends and the log of prices lagged by one period.[2] (A more detailed description of the models and how to solve them is provided in Appendix A.) I denote the simulated period t^*+1 log equity premium r_{t+1}^*. I can then estimate the model given in equation (1) with simulated data, where the simulated predictor x_t^* is either the dividend-price ratio or the dividend yield:

$$r_{t+1}^* = \beta_{M,0} + \beta_{M,1}x_t^* + \epsilon_{t+1}^*, \text{ where } \epsilon_{t+1}^* \sim N(0, \sigma_{M,\epsilon}^2). \tag{12}$$

[2]The dividend-price ratio and the dividend yield are the only two predictors used by the equity premium prediction literature that can be simulated from the three asset pricing models.

The OLS estimates of $\beta_{M,0} = [\beta_{M,0},\ \beta_{M,1}]'$ and ${\sigma_{M,\epsilon}}^2$ are denoted by β^* and σ_ϵ^{*2}, which act as the prior means of the Gamma-Normal distribution described in Section 2.2.

The priors of my baseline results are based on data simulated from the asset pricing models when calibrating them as proposed the respective authors, that is, Campbell and Cochrane (1999), Barberis et al. (2001), and Bansal and Yaron (2004). However, in Section 5, I calibrate the models with a data sample that has no overlap with the OOS period and show that the results are robust. The calibration of the models is described in Appendix A.

For the HF model, the simulation is at a monthly frequency, and the quarterly (annual) data are constructed via time-averaging the monthly data. The same procedure is used by Campbell and Cochrane (1999). The log equity premium is summed across the quarter (year). For the dividend-price ratio and the dividend yield, consumption and dividends are summed across the quarter (year) and the end-of-quarter (year) price is used. I simulate 120,000 months, estimate β^* and σ_ϵ^{*2}, and average the estimates over 10 iterations. The HF model has two specifications, and I use both to generate priors. The first specification (HF 1) assumes a perfect positive correlation between the log consumption and log dividend growth, and the second specification (HF 2) assumes that the correlation is imperfect and positive.

Similar to the HF model, the PT model is specified by Barberis et al. (2001) with perfect positive correlation between the log consumption and log dividend growth processes and with imperfect positive correlation between the two processes. I only use the latter specification, as it more successfully matches the empirical data moments. The authors calibrate the model with a range of parameter values for the investor's sensitivity to financial wealth fluctuations ($b0$) and the effect of prior losses on risk aversion (k). I generate priors from the parameterizations that set $b0$ equal to 100 and k equal to 3 (PT 1) and 8 (PT 2). Of the specifications proposed by Barberis et al. (2001), setting $b0$ equal to 100 and k equal to 8 generates a log equity premium that is closest to the empirical data moment. For the $b0$ equal to 100 and k equal to 3, the generated log equity premium is lower, but the average loss aversion of the agent is 2.25, which is in line with experimental evidence. Following Barberis et al. (2001), I simulate the model at monthly, quarterly, and annual frequencies by adjusting the model parameters accordingly.[3]

The LRR model, like the HF model, is simulated at a monthly frequency, and quarterly (annual) values are time-averaged.[4] Bansal and Yaron (2004) use the same procedure to generate simulated data. Again, 120,000 months are simulated to estimate β^* and σ_ϵ^{*2}, and the estimates are averaged across 10 iterations. Bansal and Yaron (2004) present two specifications of their model: with and without time-varying volatility of consumption growth. Because the specification that accounts for time-varying volatility of consumption growth is substantially more successful at matching the empirical data moments, I generate priors only from this specification. However, as in Bansal and Yaron (2004), I consider two calibrations for the agent's risk aversion to simulate the model: a risk aversion of 7.5 (LRR 1) and a risk aversion of 10 (LRR 2).

Panels A and B of Table 1 show β^* and σ_ϵ^{*-2} estimated from simulated data of the three consumption-based asset pricing models. The table also reports the empirical estimates over the total sample from 1926 to 2014 for comparison. For all three asset pricing models, β_1^* is positive for the dividend-price ratio and the dividend yield. Thus, high valuation ratios predict higher subsequent returns, which is in line with the empirical estimates. For both predictors and across all return frequencies, the coefficients of the LRR model are substantially lower than for the HF and PT models. The implication is that in the LRR model, the predictive power of valuation ratios is weak. Of the three models, the PT model generates the highest β_0^* and β_1^* for the dividend-price ratio. For the dividend yield, the β_0^* and β_1^* of the HF

[3]For the monthly, quarterly, and annual frequencies, I simulate 120,000, 40,000, and 10,000 periods, and average the β^* and σ_ϵ^{*2} estimates over 10 iterations.

[4]I simulate the model based on the analytical solutions as done by Beeler and Campbell (2012).

model are greater than the estimates of the other two models and the empirical estimates. The pattern for σ_ϵ^{*-2} is more mixed. However, the values implied by the asset pricing models are also close to the empirical values.

The weak implied predictability of the LRR model can also be seen in Panel C. Panel C reports the R^2 for the single-variable predictive regression in equation (12). The R^2 values for the LRR model are lower than for the HF and PT models and the empirical data. The predictability of the equity premium is strongest for the HF model, for which the R^2 is higher than for the empirical data across all frequencies and both predictors. For the PT model, the dividend-price ratio has considerable predictive power, but the R^2 values for the dividend yield are lower — consistent with the higher β_1^* for the dividend-price ratio in Panel A. While the R^2 positively correlates with the magnitude of β_1^*, the β_1^* of the HF model is lower than for the PT model for the dividend-price ratio, despite the R^2 being higher for the HF model. The reason for the smaller β_1^* of the HF model is the more volatile simulated dividend-price ratio (shown in Appendix A).

2.5 Implied predictability of the asset pricing models

The reason for the weak implied predictability of the LRR prior relative to the HF and the PT priors lies in the mechanisms through which the three asset pricing models generate a higher equity premium. (A detailed outline of the asset pricing models and their implied predictability of the equity premium can be found in Appendix A.)

In the HF model, time-variation of the dividend-price ratio is driven by a surplus consumption ratio that increases (decreases) with positive (negative) shocks to consumption. A positive shock to consumption makes the agent less risk averse, which causes asset prices to rise. The increase in the asset prices results in a lower dividend-price ratio, which predicts lower expected returns as the agent requires less compensation for risk. Hence, time-variation in the dividend-price ratio is driven by changes in the risk aversion of the agent, and these changes also affect expected returns. However, the expected dividend growth remains constant and does not affect the time-variation of the dividend-price ratio.

As the HF model, the PT model generates a time-varying dividend-price ratio through time-varying risk aversion and not changes in expected cash flows. Barberis et al. (2001) incorporate utility from fluctuations in financial wealth into a standard power utility function. Gains (losses) in financial wealth make the agent less (more) risk averse. Thus, a positive shock to dividends will lower the risk aversion of the agent, which results in a higher asset price and a lower dividend-price ratio. As the expected dividend growth remains constant, the price increase leads to lower expected returns. Dividend-price ratios and future returns are therefore positively related.

In the LRR model, the agent is concerned about economic growth prospects and economic uncertainty. The key difference in terms of predictability of the LRR model compared with the HF and the PT models is that the time-variation of the dividend-price ratio is partly driven by changes to expected growth prospects. A positive shock to expected growth leads to a lower dividend-price ratio that is followed by higher cash flows. This mitigates the predictive power of the dividend-price ratio that is generated by the economic uncertainty channel of the LRR model: a negative shock to time-varying economic uncertainty results in higher asset prices and lower dividend-price ratios, which reduces subsequent returns.

3 Results

In this section, I describe the data and report the OOS results when imposing economic constraints derived from asset pricing models on equity premium forecasts.

Table 1: Model-implied parameters

Panel A reports the coefficient estimates of the single-variable predictive regression of the log equity premium given in equation (1) for two predictors: the log dividend-price ratio and the log dividend yield. The empirical estimates are for the data sample from 1926 to 2014. The model-based estimates from the three asset pricing models HF, LRR, and PT are obtained through the Monte Carlo simulation procedure described in Section 2.4. These coefficient estimates are used as means of the Normal prior in equation (2). Panel B shows the inverse of the variance of the return innovation used for the Gamma prior in equation (2). Panel C reports the R^2 (in percent) of the single-variable predictive regression.

Panel A: Coefficients (β)														
	Empirical		**HF 1**		**HF 2**		**LRR 1**		**LRR 2**		**PT 1**		**PT 2**	
	β_0	β_1	β_0^*	β_1^*	β_0^*	β_1^*	β_0^*	β_1^*	β_0^*	β_1^*	β_0^*	β_1^*	β_0^*	β_1^*
Log dividend-price ratio														
Annual returns	0.282	0.066	0.635	0.195	0.614	0.186	0.040	0.004	0.073	0.011	0.738	0.246	0.869	0.321
Quarterly returns	0.085	0.021	0.166	0.051	0.169	0.052	0.011	0.001	0.016	0.002	0.406	0.171	0.469	0.220
Monthly returns	0.023	0.005	0.057	0.018	0.057	0.018	0.004	0.000	0.007	0.001	0.255	0.132	0.270	0.158
Log dividend yield														
Annual returns	0.313	0.077	0.555	0.169	0.516	0.154	0.039	0.004	0.069	0.009	0.167	0.046	0.216	0.063
Quarterly returns	0.080	0.019	0.093	0.051	0.092	0.049	0.009	0.000	0.016	0.002	0.078	0.025	0.135	0.054
Monthly returns	0.027	0.007	0.056	0.017	0.056	0.017	0.003	0.000	0.007	0.001	0.022	0.005	0.046	0.019

	Empirical	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
Panel B: Inverse variance of return innovation (σ_ϵ^{-2})							
Log dividend-price ratio							
Annual returns	25.632	50.231	27.667	35.668	36.917	25.284	18.733
Quarterly returns	89.139	173.088	99.737	143.934	147.964	125.520	88.708
Monthly returns	333.634	507.683	293.951	429.903	441.194	499.265	394.022
Log dividend yield							
Annual returns	25.777	48.400	26.656	35.588	36.566	24.481	17.607
Quarterly returns	89.019	172.678	100.382	143.023	147.459	123.579	86.093
Monthly returns	333.969	502.700	294.971	432.105	440.879	498.053	391.489
Panel C: Predictability (R^2 in %)							
Log dividend-price ratio							
Annual	2.408	12.219	7.561	0.014	0.026	3.194	6.068
Quarterly	0.819	3.244	2.142	0.004	0.007	1.174	2.787
Monthly	0.200	1.100	0.733	0.001	0.001	0.423	0.925
Log dividend yield							
Annual	2.956	8.900	5.196	0.006	0.012	0.158	0.369
Quarterly	0.686	3.053	1.962	0.001	0.003	0.052	0.268
Monthly	0.300	1.062	0.703	0.000	0.002	0.002	0.038

3.1 Data

The empirical data on the equity premium and the predictors at a monthly, quarterly, and annual frequency are available on Amit Goyal's website.[5] The equity premium is computed as the log return on the S&P 500 index minus the log three-month U.S. Treasury bill rate. I set the start date of the time series at 1926, as high-quality return data on the S&P 500 from the Center of Research in Security Prices became available in 1926. The time series ends in 2014. The availability of predictor variables that can be used to assess the performance of the model-based priors is restricted by the three asset pricing models. The predictor variables that can be simulated from the three models are the dividend-price ratio and the dividend yield. Dividends on the S&P 500 index are 12-month moving sums from 1926 to 2014. As for the data simulated from the asset pricing models, the dividend-price ratio is defined as the difference between log dividends and log prices, and the dividend yield is defined as the difference between log dividends and log prices lagged by one period.

3.2 Measuring forecast accuracy

I assess the performance of the model-based priors via the OOS R^2 (see, for example, Campbell and Thompson (2008)):

$$R^2_{OOS} = 1 - \frac{\sum_{\tau=\underline{t}}^{T} (r_\tau - \hat{r}_\tau^m)^2}{\sum_{\tau=\underline{t}}^{T} (r_\tau - \hat{r}_\tau^h)^2}, \tag{13}$$

where $\hat{r}_\tau^m$ is the equity premium forecast when imposing the model-based prior as given in equation (11); $\hat{r}_\tau^h$ is the prediction of the historical average model; and $\underline{t}$ and T are the start and end dates, respectively, of the OOS forecast period. Thus, the R^2_{OOS} assesses the forecast performance of the model-based priors relative to the non-predictability model, which assumes that the best forecast of the equity

[5] Amit Goyal's website address is http://www.hec.unil.ch/agoyal/.

premium is its historical average. The historical average model corresponds to the model given in equation (1) with β_1 being set equal to zero.

3.3 Forecasting

I consider four sample periods for the OOS predictability exercise. First, I use the full sample from 1926 to 2014 and start the recursive OOS forecast in 1947. This starting point guarantees that a sufficient number of data points are available to estimate the predictive model. Next, I analyze the subsample stability by splitting the OOS forecast period (1947 to 2014) in half and consider forecasts up to 1980 and forecasts starting in 1981. Last, I only use the postwar sample from 1947 to 2014, and the forecasts start in 1968.

Figure 1 shows the quarterly OOS forecasts of the log equity premium from 1947 to 2014 in the top panel. The valuation ratios predict a substantial time variation of the equity premium. The lower panel depicts the corresponding coefficient estimates. Both predictors lost predictive power during the dot-com boom in the late 1990s, which leads to the sharp drop in the coefficient estimates.

Table 2 shows the R^2_{OOS} (in percent) results for all six model-based priors for three return frequencies: monthly, quarterly, and annual. The "no prior" column reports the R^2_{OOS} for the case in which the single-variable predictive model in equation (1) is estimated via OLS. If the model-based prior leads to an increase in the R^2_{OOS}, then the figure is in bold. The last two columns of the table show the best- and second-best-performing prior for the respective frequency, predictor, and time period. In Panel D, the performance of the model-based priors is summarized across the three frequencies.

For every case, there is at least one asset-pricing model that would help an investor improve the accuracy of her equity premium forecasts, with the exception of the log dividend-price ratio at a monthly frequency for the OOS period from 1947 to 1980. The gains in R^2_{OOS} are considerable compared with the literature

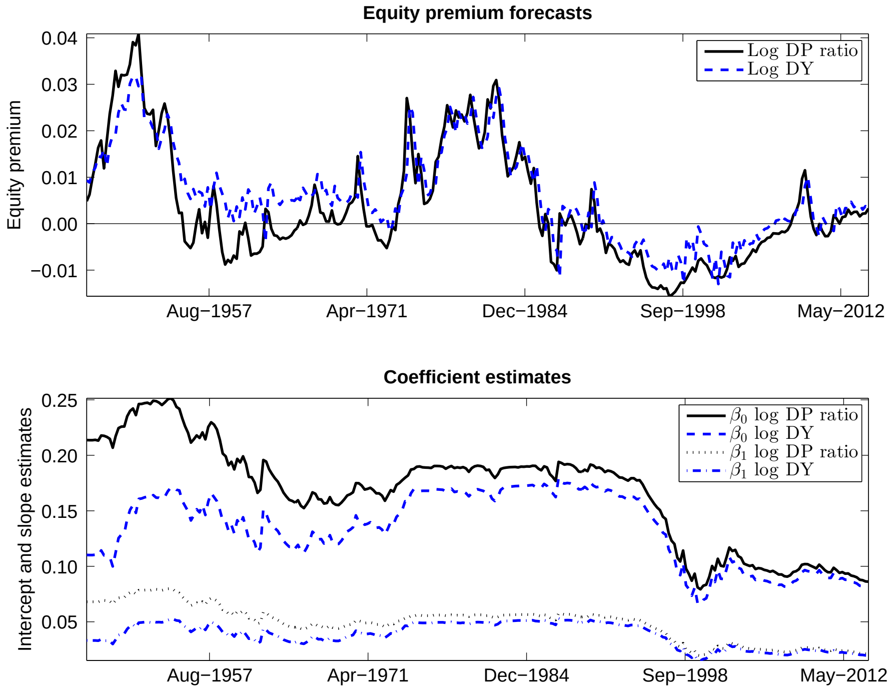

Figure 1: Empirical out-of-sample forecasts
The top panel shows the OOS quarterly log equity premium forecasts for two predictors: the log dividend-price ratio and the log dividend yield. The predictive model in equation (1) is estimated recursively via OLS. The data sample starts in 1926 and the OOS period is from 1947 to 2014. The lower panel depicts the corresponding OLS coefficient estimates.

(see, for example, Campbell and Thompson (2008)). In Section 4, I show that these R^2_{OOS} values correspond to sizable gains in utility for an investor with mean-variance preferences. For the log dividend-price ratio, the LRR priors are the best performing for three of the four sample periods. The only OOS period for which the LRR priors are never the best-performing priors is the period from 1947 to 1980. This finding is consistent across all return frequencies. For the OOS period from 1947 to 1980, the PT model-based priors lead to the greatest increase in R^2_{OOS} at the annual frequency, but the HF model-based priors result in more accurate forecasts at the monthly and quarterly frequencies. For the log dividend yield, the findings are similar: The LRR model-based priors are never superior to the priors derived from the other two asset pricing models for the 1947-1980 OOS period regardless of the frequency. However, for the other three sample periods, the LRR priors are the best performing for at least one return frequency. The summary in Panel D reports that the LRR 1 (LRR

2) model yielded the best-performing prior in 37.5 percent (20.8 percent) of all cases. Additionally, the LRR 1 (LRR 2) prior was the second-best prior 20.8 percent (41.7 percent) of the time. However, the HF 2 and PT 2 models were the best performing priors only in 12.5 percent of all cases (second best for 8.3 percent).

Table 3 reports the differences in the R^2_{OOS} (in percent) between the best-performing prior and the other priors for every return frequency, predictor, and sample period. To test whether the difference in forecast errors is statistically significant, I use a one-sided Diebold-Mariano test (see Diebold and Mariano (1995)). Despite the difficult task to statistically reject OOS forecasting models of the equity premium (see, for example, Campbell and Thompson (2008) and Welch and Goyal (2008)), the differences are statistically significant in several cases. For the log dividend-price ratio, the hypothesis of equal predictive power of the model estimated with the LRR priors and the PT priors can be rejected for the majority of data samples. The differences between the R^2_{OOS} of the LRR priors and the HF priors are generally smaller and, thus, significant in fewer cases. When the log dividend yield acts as the predictor, the results are not as pronounced as for the dividend-price ratio, but the hypothesis of equal predictive power can be rejected particularly at the monthly frequency, where more data points are available and the power of the test is increased. The analysis in Section 4 shows that even small differences in R^2_{OOS} can lead to substantial utility gains for an investor with mean-variance preferences.

The strong performance of the LRR prior can be explained by the low model-implied predictability. In Table 1, β_0^* and β_1^* are lower for all three frequencies and both predictors compared with the empirical estimates and β_0^* and β_1^* of the HF and PT models. Thus, imposing the LRR prior pushes the posterior estimates of β_0 and β_1 down. Figure 2 shows the OLS estimates — that is, no prior is imposed on the predictive regression — and the posterior estimates for the log dividend-price ratio and quarterly returns for the 1968-2014 OOS period. The LRR 1 posterior estimates

Table 2: Forecast performance of model-based priors

Panels A, B, and C show the OOS performance of the model-based priors derived from the three consumption-based asset pricing models: HF, LRR, and PT. The priors are imposed on the single variable predictive regression given in equation (1). Reported is the R^2_{OOS} (in percent) from equation (13), which measures the accuracy of OOS log equity premium forecasts of the single variable predictive regression relative to the historical average model. The predictors are the dividend-price ratio and the dividend yield. If the model-based prior leads to an increase in the R^2_{OOS} relative to OLS forecast in the "no prior" column, the figure is in bold. The last two columns denote which model-based prior leads to the greatest and second-greatest improvement in forecast accuracy. Panel D shows the number of times each model-based prior is the best or second best.

Panel A: Annual returns										
Sample start	**OOS period**	**No prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**	**1st prior**	**2nd prior**
Log dividend-price ratio										
1926	1947-2014	0.396	-1.209	-1.292	**1.137**	**0.607**	-4.045	-7.579	LRR 1	LRR 2
1926	1947-1980	11.895	**15.342**	**14.727**	3.927	5.083	**16.355**	**16.564**	PT 2	PT 1
1926	1981-2014	-12.242	-18.486	-18.681	**-3.027**	**-2.689**	-25.224	-34.249	LRR 2	LRR 1
1947	1968-2014	-3.521	**-1.831**	**-1.606**	**2.237**	**2.313**	-4.978	-9.168	LRR 2	LRR 1
Log dividend yield										
1926	1947-2014	-16.280	**-4.923**	**-3.060**	**0.309**	**0.443**	**0.837**	**0.788**	PT 1	PT 2
1926	1947-1980	-10.255	**11.488**	**11.530**	**4.417**	**5.449**	**6.698**	**8.389**	HF 2	HF 1
1926	1981-2014	-22.902	**-21.777**	**-19.763**	**-4.760**	**-4.257**	**-6.115**	**-8.073**	LRR 2	LRR 1
1947	1968-2014	-1.334	**-0.069**	**-0.001**	**0.785**	**1.070**	**1.899**	**2.345**	PT 2	PT 1

Panel B: Quarterly returns										
Sample start	**OOS period**	**No prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**	**1st prior**	**2nd prior**
Log dividend-price ratio										
1926	1947-2014	-0.815	**-0.607**	-1.020	**0.246**	**0.052**	-5.032	-7.551	LRR 1	LRR 2
1926	1947-1980	3.753	**4.110**	3.680	3.334	3.375	1.752	0.306	HF 1	HF 2
1926	1981-2014	-4.727	-4.995	-5.074	**-2.882**	**-2.448**	-11.334	-14.032	LRR 2	LRR 1
1947	1968-2014	-0.391	-0.541	-0.542	**0.640**	**0.619**	-5.060	-7.135	LRR 1	LRR 2
Log dividend yield										
1926	1947-2014	0.340	**0.877**	**0.642**	**0.834**	**0.859**	**0.740**	**0.416**	HF 1	LRR 2
1926	1947-1980	4.740	**4.889**	4.253	3.133	3.621	4.458	**5.035**	PT 2	HF 1
1926	1981-2014	-3.429	**-3.031**	**-2.646**	**-1.706**	**-1.892**	**-2.442**	**-3.225**	LRR 1	LRR 2
1947	1968-2014	-0.297	**0.495**	**0.439**	**0.702**	**0.830**	**0.606**	**0.075**	LRR 2	LRR 1

Table 2: Forecast performance of model-based priors (continued)

Panel C: Monthly returns										
Sample start	**OOS period**	**No prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**	**1st prior**	**2nd prior**
Log dividend-price ratio										
1926	1947-2014	-0.061	-0.254	-0.075	**0.092**	**0.063**	-1.680	-1.656	LRR 1	LRR 2
1926	1947-1980	1.264	1.109	1.204	1.080	0.777	0.740	0.182	HF 2	HF 1
1926	1981-2014	-1.142	-1.369	**-0.996**	**-0.937**	**-1.121**	-3.268	-3.450	LRR 1	HF 2
1947	1968-2014	-0.237	-0.357	-0.093	**0.126**	**0.014**	-2.483	-2.784	LRR 1	LRR 2
Log dividend yield										
1926	1947-2014	-0.393	**-0.300**	**-0.310**	**-0.211**	**-0.267**	**-0.356**	-0.435	LRR 1	LRR 2
1926	1947-1980	1.175	1.149	**1.238**	1.014	1.046	1.026	**1.217**	HF 2	PT 2
1926	1981-2014	-1.672	-1.944	-1.810	**-1.236**	**-1.272**	**-1.401**	-1.791	LRR 1	LRR 2
1947	1968-2014	-0.166	-0.253	**-0.129**	**-0.087**	**0.110**	**0.153**	**-0.038**	PT 1	LRR 2

Panel D: Summary of model-based prior performance

	Best-performing prior		**Second-best-performing prior**	
	#	**in %**	**#**	**in %**
HF 1	2	8.3	3	12.5
HF 2	3	12.5	2	8.3
LRR 1	9	37.5	5	20.8
LRR 2	5	20.8	10	41.7
PT 1	2	8.3	2	8.3
PT 2	3	12.5	2	8.3

Table 3: Comparison of the model-based prior forecast performance

Reported are the differences in the R^2_{OOS} (in percent), where the R^2_{OOS} is given in equation (13) and measures the accuracy of OOS log equity premium forecasts of a single-variable predictive regression, given in equation (1), relative to the historical average model. The predictive regression is estimated via model-based priors derived from the three consumption-based asset pricing models: HF, LRR, and PT. The predictors are the dividend-price ratio and the dividend yield. For each frequency, predictor, and sample period, the R^2_{OOS} of each prior is subtracted from the R^2_{OOS} of the best-performing prior. The statistical significance of the difference is tested with a one-sided Diebold-Mariano test (see Diebold and Mariano (1995)). A significant test statistic is denoted by ** at the 5 percent level and by * at the 10 percent level. The best performing-prior is reported in the third column.

Panel A: Annual returns								
Sample start	**OOS period**	**Best prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**
Log dividend-price ratio								
1926	1947-2014	LRR 1	2.346	2.429		0.530	5.181	8.716
1926	1947-1980	PT 2	1.222	1.837	12.637*	11.480*	0.209	
1926	1981-2014	LRR 2	15.797	15.992	0.338		22.535	31.560*
1947	1968-2014	LRR 2	4.144	3.918	0.076		7.291	11.480
Log dividend yield								
1926	1947-2014	PT 1	5.760	3.897	0.528	0.394		0.048
1926	1947-1980	HF 2	0.042		7.113	6.082	4.833	3.141
1926	1981-2014	LRR 2	17.521*	15.507*	0.503		1.859	3.817
1947	1968-2014	PT 2	2.415	2.346	1.560	1.275	0.446	

Table 3: Comparison of the model-based prior forecast performance (continued)

Panel B: Quarterly returns								
Sample start	**OOS period**	**Best prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**
Log dividend-price ratio								
1926	1947-2014	LRR 1	0.853	1.266		0.194	5.278**	7.797**
1926	1947-1980	HF 1		0.430	0.776	0.735	2.358	3.804
1926	1981-2014	LRR 2	2.547**	2.626**	0.434		8.886**	11.584**
1947	1968-2014	LRR 1	1.181	1.182		0.021	5.700*	7.774*
Log dividend yield								
1926	1947-2014	HF 1		0.235	0.331	0.436	0.137	0.461
1926	1947-1980	PT 2	0.146	0.782*	1.902*	1.414	0.577	
1926	1981-2014	LRR 1	1.325*	0.940		0.186	0.736	1.518
1947	1968-2014	LRR 2	0.334	0.391	0.127		0.224	0.755

Panel C: Monthly returns								
Sample start	**OOS period**	**Best prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**
Log dividend-price ratio								
1926	1947-2014	LRR 1	0.347*	0.167		0.029	1.773**	1.749**
1926	1947-1980	HF 2	0.096		0.124	0.427	0.464	1.023
1926	1981-2014	LRR 1	0.432*	0.059		0.184	2.331**	2.512**
1947	1968-2014	LRR 1	0.483**	0.218		0.112	2.608**	2.910**
Log dividend yield								
1926	1947-2014	LRR 1	0.089	0.099		0.057	0.146	0.224
1926	1947-1980	HF 2	0.088		0.224	0.192	0.212	0.021
1926	1981-2014	LRR 1	0.707**	0.574**		0.036	0.165	0.555**
1947	1968-2014	PT 1	0.406**	0.283*	0.240*	0.043		0.191*

are substantially lower than the OLS estimates and the posterior estimates of the HF 1 and PT 1 models. However, the model-based priors derived from the HF 1 model lead to posterior estimates that are similar to the OLS estimates. The model-based priors from the PT 1 model push the posterior estimates for both coefficients higher than they are when ignoring any prior and simply relying on OLS estimates.

The lower posterior estimates achieved through the LRR 1 prior are beneficial for an investor as shown in Figure 3. The top panel depicts the differences in the cumulative sum of squared errors (SSE) when forecasting with the historical average model compared with the single-variable predictive model estimated via OLS and via model-based priors. I subtract the cumulative SSE of the predictive model from the cumulative SSE of the historical average model. Hence, a positive value implies that the predictive model outperforms the historical average model. Until the beginning of the 1990s, the predictive regression performs better than the historical average model regardless of the estimation method. The highest cumulative SSE value is achieved for an investor who relies on the priors of the PT 1 model, which is due to the strong predictive power of the log dividend-price ratio implied by the PT 1 model. In the 1970s, valuation ratios had strong predictive power, and the PT 1 model makes the investor rely on this predictive power to a higher degree than an investor who uses the HF 1 or LRR 1 model to form her priors. The LRR 1 prior leads to the lowest cumulative SSE value until 1994. However, during the dot-com boom from 1994 to 1999, the predictive power of the log dividend-price ratio collapses and the cumulative SSE of the predictive model turns negative for all four estimation methods. The investor armed with the LRR 1 model is able to avoid poor forecasts to some extent, as her belief in the predictive ability of valuation ratios is qualified because of her prior. The lower panel of Figure 3 provides further detail. The equity premium forecasts for the OOS period from 1968 to 2014 are depicted. The posterior point forecasts given in equation (11) of the LRR 1 model are close to zero during the dot-com boom. The other two model-based priors and the OLS

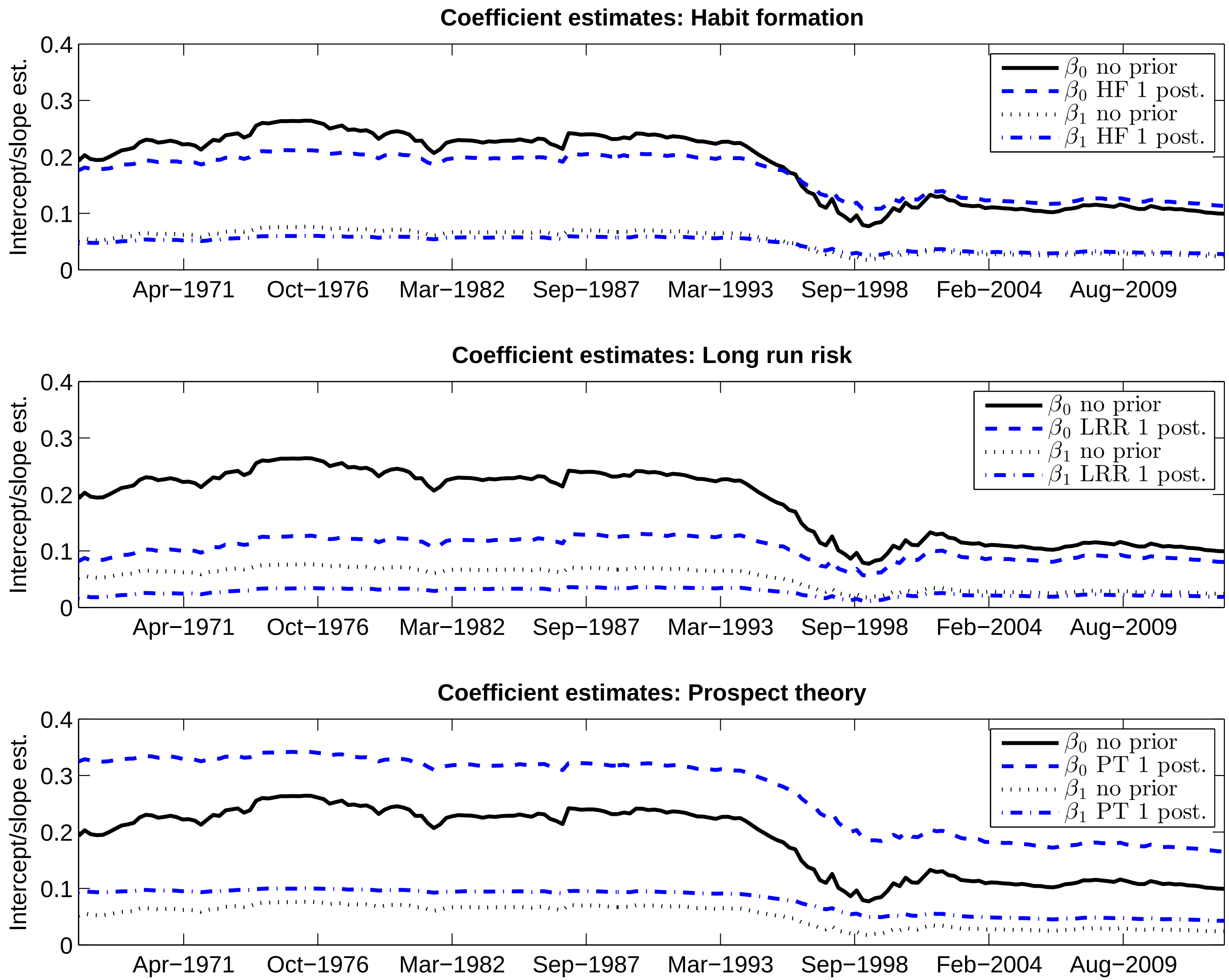

Figure 2: Posterior out-of-sample coefficient estimates
This figure shows the OLS estimates and the posterior estimates of β_0 and β_1 for the OOS period from 1968 to 2014 for the HF 1, LRR 1, and PT 1 priors. The predictor is the log dividend-price ratio, and the predictive model in equation (1) is estimated with quarterly data.

estimates result in strongly negative forecasts. Hence, an investor relying on these forecasts to time the market suffers losses during this bull market period.

Unlike the dot-com boom and its subsequent downturn, the financial crisis in 2008 does not have a substantial effect on the performance of the model-based priors, which can be explained by the different nature of these two episodes. Campbell, Giglio, and Polk (2013) find that during the dot-com boom, the discount rates of investors were at a historically low level. These low discount rates resulted in low valuation ratios, which were not followed by negative returns. Hence, the predictive power of the dividend-price ratio and the dividend yield decreased. However, according to Campbell et al. (2013), the downturn from 2007 to 2009 was driven by a decrease in rational expectations of future profits and the preceding boom caused by positive cash flow news. Consequently, valuation ratios performed better at predicting the equity premium compared with the late 1990s.

Figure 4 shows the simulated posterior density of the quarterly log equity premium prediction given in equation (10) for the third quarter in 1998. The predictor is the log dividend-price ratio, and the model-based priors are the same as in Figures 2 and 3. The densities are simulated with 10,000 draws. For all three model-based priors, the posterior densities are similarly shaped and approximate a Normal distribution. When imposing the LRR 1 prior, the density is furthest to the right, corresponding to an equity premium forecast that is greater than the forecast of the other two models. These posterior densities are in line with the predictions during the dot-com boom shown in the lower panel of Figure 3. Figure 5 shows the corresponding posterior densities of β_0 and β_1 given in equation (6). The densities are again similar across the three model-based priors. For both coefficients, the LRR 1 prior results in posterior densities that are centered to the left of the HF 1 and PT 1 priors, consistent with the higher posterior mean of the equity premium predictive density shown in Figure 4. Hence, in the third quarter of 1998 at the height of the dot-com boom, when the dividend-price ratio was low, an investor who believes in

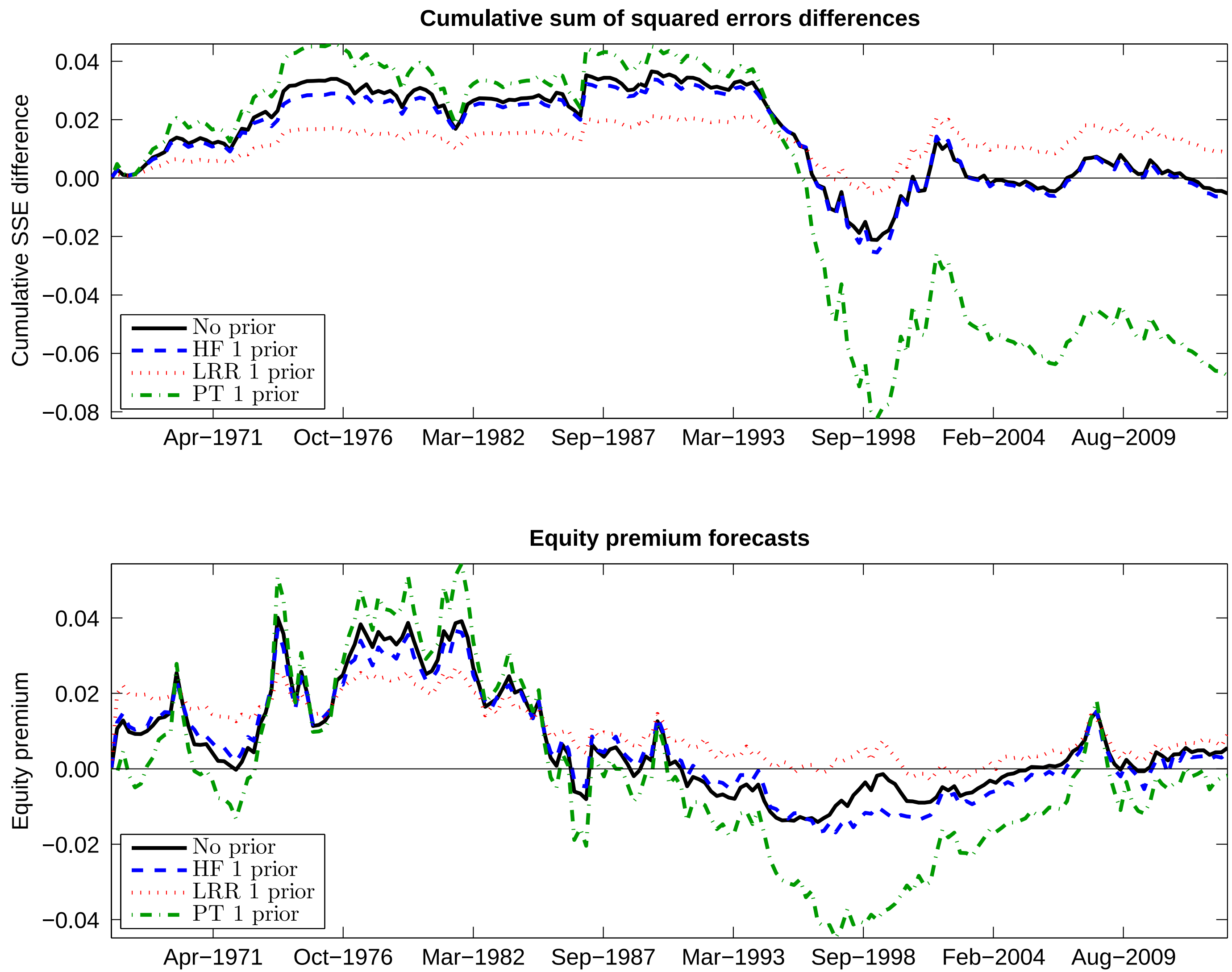

Figure 3: Forecast performance of model-based priors
The top panel shows the differences in cumulative sum of squared errors (SSE) between the log equity premium forecasts of the historical average model and the predictive model given in equation (1) estimated via OLS or model-based prior. The model-based priors are derived from the HF 1, LLR 1, and PT 1 models. The cumulative SSE of the predictive model are subtracted from the cumulative SSE of the historical average model. The OOS period is from 1968 to 2014 and the forecasts are at a quarterly frequency. The predictor is the log dividend-price ratio. The lower panel depicts the point forecasts of the equity premium of the predictive model for the OLS estimates and the model-based priors. For the model-based priors, the posterior point forecasts are given in equation (11).

the HF 1 or PT 1 model expects a negative equity premium to materialize in the next period. However, an investor whose prior beliefs are in line with the LRR 1 model is more hesitant to draw this conclusion.

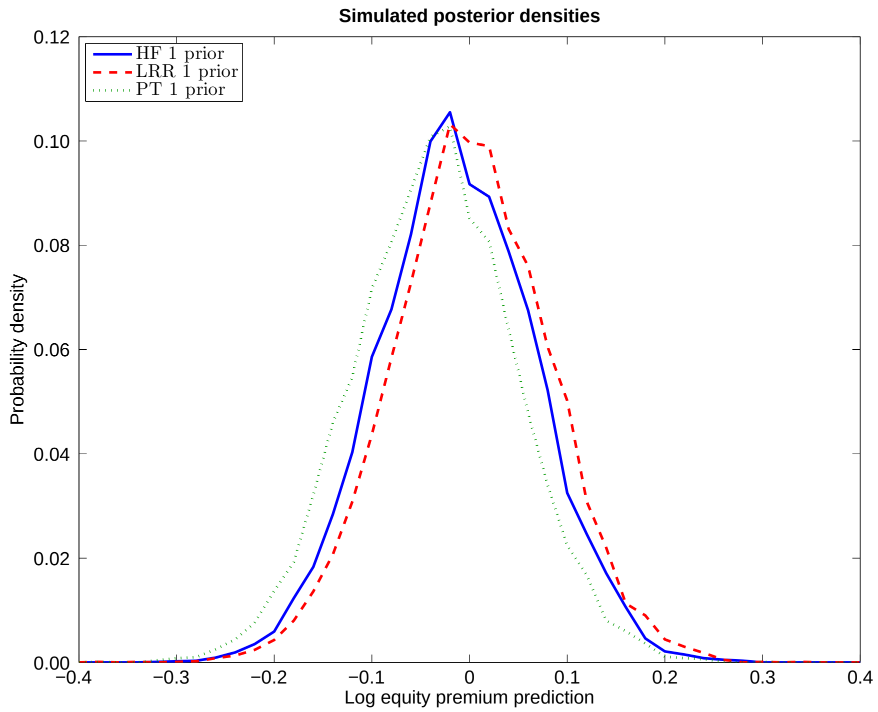

Figure 4: Posterior density of equity premium prediction
This figure shows the simulated posterior density of the quarterly log equity premium prediction given in equation (10) for the third quarter in 1998 for three model-based priors: HF 1, LRR 1, and PT 1. The predictor is the log dividend-price ratio. Data from the first quarter in 1947 to the second quarter in 1998 are used to estimate the predictive model. The density are simulated with 10,000 draws.

4 Utility of an investor

So far, I have analyzed how priors derived from the three consumption-based asset pricing models affect the forecast accuracy of single-variable predictive regressions. To investigate the economic significance of the changes in predictive performance, we need to compute the utility gains of an investor who uses the model-based priors to forecast the equity premium. Further, comparing the utility gains of investors who rely on different consumption-based asset pricing models accounts for the investors' risk preferences. The Bayesian technique that I use to impose the economic

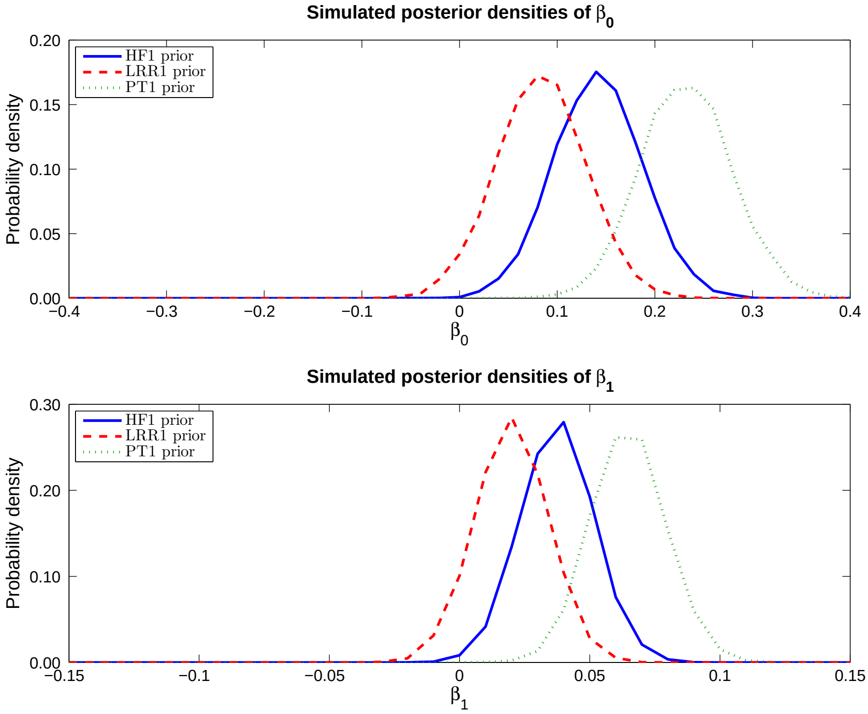

Figure 5: Posterior density of coefficients
This figure shows the simulated posterior density of the coefficients β_0 and β_1 given in equation (6) for the third quarter in 1998 for three model-based priors: HF 1, LRR 1, and PT 1. The predictor is the log dividend-price ratio. Quarterly data from the first quarter in 1947 to the second quarter in 1998 are used to estimate the predictive model. The densities are simulated with 10,000 draws.

constraints provides the full predictive density of the equity premium. Based on the mean and the variance of the predictive density, I can compute the portfolio allocation and utility gains of an investor with mean-variance preferences (see, for example, Campbell and Thompson (2008) and Wachter and Warusawitharana (2009)). The utility gains of an investor achieved through the model-based priors will also give us an estimate of how much an investor would be willing to pay to know the theory of one consumption-based asset pricing model over another.

4.1 Asset allocation

An investor is assumed to have mean-variance preferences, and she chooses portfolio weights for a risky asset and a risk-free asset. The return on the risky asset is the log equity premium r_{t+1} plus the log risk-free return $r_{f,t}$, and the risk-free asset yields

$r_{f,t}$. At time t, the investor solves the following maximization problem

$$\max_{\alpha_t} E_t\left[W_{t+1}\right] - \gamma \frac{1}{2} Var_t\left[W_{t+1}\right], \tag{14}$$

subject to

$$W_{t+1} = \alpha_t \exp(r_{t+1} + r_{f,t}) + (1 - \alpha_t)\exp(r_{f,t}), \tag{15}$$

where α_t is the portfolio share of the risky asset, and γ is the risk aversion of the investor. The solution to the maximization problem is

$$\alpha_t^* = \frac{E_t\left[\exp(r_{t+1} + r_{f,t}) - \exp(r_{f,t})\right]}{\gamma Var_t\left[\exp(r_{t+1})\right]}. \tag{16}$$

For an investor who imposes model-based priors to forecast the equity premium, we can use the mean and the variance of the sampled predictive density of r_{t+1} given in equation (10) to approximate α_t^* denoted $\hat{\alpha}_{t,m}^*$. Based on $\hat{\alpha}_{t,m}^*$ and the realized equity premium, the realized wealth can be obtained

$$\widehat{W}_{t+1,m} = \hat{\alpha}_{t,m}^* \exp(r_{t+1} + r_{f,t}) + (1 - \hat{\alpha}_{t,m}^*)\exp(r_{f,t}). \tag{17}$$

Solving for $\hat{\alpha}_{t,m}^*$ for $t = \underline{t} - 1, ..., T - 1$ results in a sequence of $\left\{\widehat{W}_{t,m}\right\}_{t=\underline{t}}^{T}$. The realized utility over the total OOS sample period is then given by

$$\widehat{U}_m = \overline{W}_m - \gamma \frac{1}{2} \frac{1}{T - \tau} \sum_{\tau=\underline{t}}^{T} (\widehat{W}_{\tau,m} - \overline{W}_m)^2, \tag{18}$$

where $\overline{W}_m = \frac{1}{T-(\tau-1)} \sum_{\tau=\underline{t}}^{T} \widehat{W}_{\tau,m}$.

When estimating the realized utility of portfolios N and A, a certainty equivalent return (CER) can be computed. The CER is defined as a constant return that, when added to the portfolio return of portfolio N, equates the realized utility of portfolios

N and A. The CER is given by

$$\text{CER} = \widehat{U}_A - \widehat{U}_N. \tag{19}$$

A more intuitive interpretation of the CER is a transaction cost or a management fee that the investor is willing to pay each period to have access to the equity premium forecasts used for portfolio A. For example, when portfolio N uses the model-based prior from the HF model and portfolio A uses the model-based prior from the LRR model, then the CER tells us how much the investor would be willing to pay to have access to the LRR model compared with the HF model.

4.2 Utility results

I compute the CER given in equation (19) for each return frequency, predictor, and OOS period. The share of the risky asset for portfolio A is computed based on the predictions of the predictive model when imposing the model-based prior which results in the highest utility for the investor. The share of the risky asset for portfolio N is computed based on the predictions when imposing one of the remaining five model-based priors, respectively. The results are shown in Table 4, which is structured like Table 3 but with the R^2_{OOS} figures replaced with the annualized CERs. The risk aversion parameter γ is set equal to 5.

The CER results are even more in favor of the LRR prior than the R^2_{OOS} results reported in Table 2: an investor who derives her prior belief about the predictability of the equity premium from the LRR model performs consistently the best for three out of the four sample periods across all frequencies and both predictors. The only OOS period during which the HF and PT priors dominate is from 1947 to 1980. The CERs are economically significant with the maximum value being 4.46%.

Panel D averages the CER for each prior pair across all frequencies, predictors, and sample periods. These values show how much on average an investor would be willing to pay to have access to the model-based priors derived from the consumption-based asset pricing model in the top row instead of any of the remaining five model-based priors. For the LRR 1 prior, all the CER values are positive, which implies that an investor who uses the LRR 1 prior to predict the equity premium and allocate her portfolio according to these predictions achieves the highest average utility. The LRR 2 model is a close second with positive average CER values against all model-based priors except the LRR 1 prior. The average CER values are substantial. Investors who rely on the HF 1 or HF 2 priors would pay between 20 and 30 basis points per year to have access to the LRR priors. The investors who derive their prior beliefs about the predictability of the equity premium from the PT model would on average need an additional 60 and 75 basis points per year to achieve the utility level of the investor who uses the LRR priors.

Figure 6 shows the risky asset share of the portfolio given in equation (16) for the HF 2, LRR 2, and PT 2 priors. The forecasts are at an annual frequency, and the OOS period is from 1947 to 2014. The top panel shows the risky asset share when the dividend-price ratio is used as the predictor. The bottom panel replaces the dividend-price ratio with the dividend-yield. Generally, the LRR prior leads to a more stable portfolio share of the risky asset, which corresponds to the low predictability implied by the model. The greatest difference between the priors is again during the bull market of the late 1990s. For the dividend-price ratio, the HF 2 and the PT 2 investors short the risky asset during this period, because they expect low valuation ratios to predict strongly negative returns. However, the investor with prior beliefs derived from the LRR 2 model is skeptical about the predictive power of the low valuation ratios and maintains a positive weight on the risky asset. The bottom panel is similar to the top panel with the difference being that the PT 2 investor is more bullish during the bull market of the late 1990s when predicting with the dividend yield. This difference is explained by the prior means of the predictive regression coefficients reported in Panel A of Table 1. The model-implied

Table 4: Economic performance of the model-based priors

Panel A, B, and C report the annualized CER (in percent) given in equation (19). The CER can be interpreted as a management fee that an investor with mean-variance utility and a risk aversion of 5 is willing to pay each year to have access to the equity premium forecasts which result from imposing the best-performing prior on the single-variable predictive regression given in equation (1) instead of the equity premium forecasts based on one of the remaining five priors. The best-performing prior is reported in the "best prior" column. The predictors are the dividend-price ratio and the dividend yield. Panel D reports the average CER across all frequencies, predictors, and sample periods, for each prior pair.

Panel A: Annual returns

Sample start	OOS period	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
Log dividend-price ratio								
1926	1947-2014	LRR 1	0.375	0.349		0.011	0.536	0.768
1926	1947-1980	HF 1		0.049	0.286	0.219	0.001	0.019
1926	1981-2014	LRR 1	1.028	0.937		0.062	1.315	1.716
1947	1968-2014	LRR 2	0.857	0.818	0.012		1.235	1.675
Log dividend yield								
1926	1947-2014	LRR 2	0.402	0.361	0.012		0.078	0.122
1926	1947-1980	HF 2	0.022		0.242	0.223	0.176	0.173
1926	1981-2014	LRR 1	1.144	0.902		0.059	0.175	0.255
1947	1968-2014	LRR 2	0.465	0.338	0.079		0.020	0.026

Panel B: Quarterly returns

Sample start	OOS period	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
Log dividend-price ratio								
1926	1947-2014	LRR 1	0.205	0.172		0.001	0.671	0.862
1926	1947-1980	PT 1	0.015	0.049	0.177	0.196		0.044
1926	1981-2014	LRR 2	0.474	0.485	0.027		1.538	1.803
1947	1968-2014	LRR 1	0.702	0.592		0.110	3.170	3.791
Log dividend yield								
1926	1947-2014	LRR 1	0.134	0.096		0.051	0.051	0.129
1926	1947-1980	HF 1		0.102	0.216	0.171	0.063	0.053
1926	1981-2014	LRR 2	0.187	0.221	0.011		0.030	0.281
1947	1968-2014	LRR 1	0.427	0.203		0.025	0.221	0.325

Table 4: Economic performance of the model-based priors (continued)

Panel C: Monthly returns								
Sample start	**OOS period**	**Best prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**
Log dividend-price ratio								
1926	1947-2014	LRR 1	0.138	0.097		0.027	0.619	0.771
1926	1947-1980	PT 1	0.104	0.051	0.084	0.212		0.034
1926	1981-2014	LRR 2	0.297	0.188	0.159		1.374	1.382
1947	1968-2014	LRR 1	0.477	0.136		0.072	4.443	4.459
Log dividend yield								
1926	1947-2014	LRR 1	0.153	0.166		0.159	0.074	0.101
1926	1947-1980	PT 2	0.016	0.121	0.070	0.170	0.015	
1926	1981-2014	LRR 2	0.150	0.410	0.075		0.197	0.374
1947	1968-2014	PT 2	0.616	0.419	0.044	0.003	0.266	

Panel D: Average CER						
Comparison prior	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**
HF 1		-0.047	0.287	0.275	-0.323	-0.443
HF 2	-0.047		0.240	0.229	-0.370	-0.490
LRR 1	-0.287	-0.240		-0.012	-0.611	-0.731
LRR 2	-0.275	-0.229	0.012		-0.599	-0.719
PT 1	0.323	0.370	0.611	0.599		-0.120
PT 2	0.443	0.490	0.731	0.719	0.120	

parameters of PT 2 for the dividend yield are smaller than for the dividend-price ratio, which makes the PT 2 investor more hesitant to believe that the low dividend yield will cause and immediate downturn of the stock market.

5 Robustness

For the baseline results shown previously, the consumption-based asset pricing models used to derive the priors are calibrated as proposed by the respective authors, that is, Campbell and Cochrane (1999), Barberis, et al. (2001), and Bansal and Yaron (2004). To test whether the results are robust to calibrating the asset pricing models with data from a time period that has no overlap with the OOS period, I calibrate the parameters of the asset pricing models with data from 1926 to 1967. In the OOS forecast exercise above, 1926 is the beginning of the return sample, and 1967 is the end of the burn-in period for the postwar data sample. All three asset pricing models are calibrated with annual data. Hence, using a shorter sample for the calibration makes the task of matching empirical moments too challenging for the models, as the empirical moments are likely distorted by outliers. Details regarding the calibration of the models can be found in Appendix A.

Table 5 reports the results for the priors derived from the asset pricing models calibrated with data from 1926 to 1967. The OOS forecasts start in 1968. Panel A shows the R^2_{OOS} for each prior, predictor, and return frequency. The priors from the LRR model perform consistently the best and improve the R^2_{OOS} relative to the "no prior" forecast in every case. In Panel B, the difference between the R^2_{OOS} of the respective prior and the best performing prior is shown. The significance of the difference in forecast errors is tested with a Diebold-Mariano test (see Diebold and Mariano (1995)). The differences in forecast performance between the LRR model-based priors and priors derived form the HF and the PT model are generally significant at a monthly and a quarterly frequency, where the higher number of

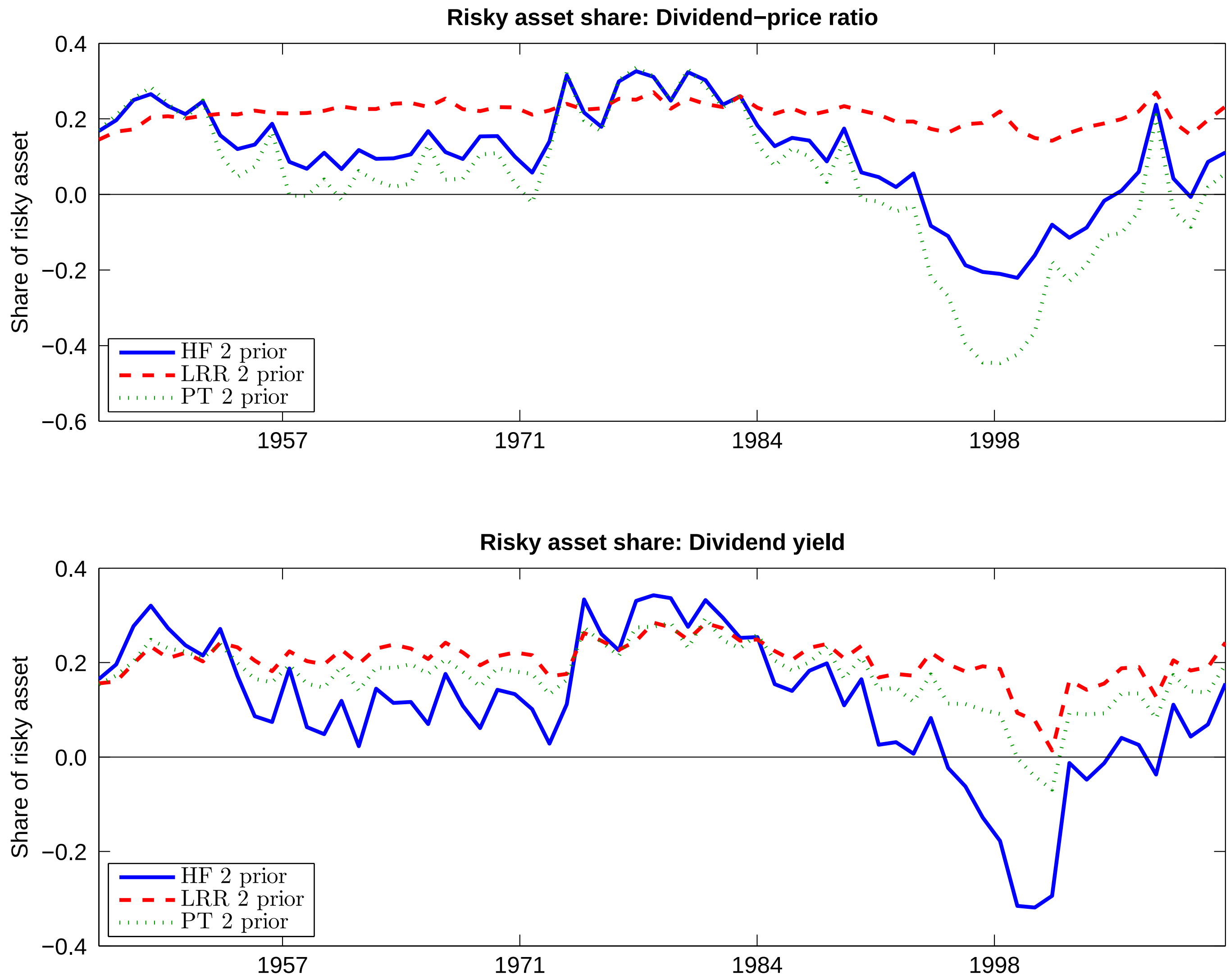

Figure 6: Portfolio share of risky asset
This figure shows the risky asset share of the portfolio given in equation (16) for mean-variance investors with γ equal to 5 who rely on the HF 2, LRR 2, and PT 2 priors, respectively. In the top (bottom) panel, the return on the risky asset is forecast with the dividend-price ratio (dividend yield). The returns are at an annual frequency, and the OOS period is from 1947 to 2014.

observations leads to more power compared to the annual returns.

For the baseline results discussed previously, the tightness parameters of the Gamma-Normal prior, λ and $\underline{v}$, are set equal to 1 and 0.1, respectively, as described in Section 2.2. This section analyzes whether the results and the conclusions drawn in this paper are robust to tightening or loosening the model-based priors.

Table 6 reports the results when tightening the prior by a factor of two — that is, $\lambda = 0.5$ and $\underline{v} = 0.2$. As in Table 2, the LRR priors excel for three of the four sample periods. Across both predictors, the OOS period from 1947 to 1980 is the only period for which the LRR priors are never the best-performing prior. This finding is consistent across all frequencies. Panel D shows that the LRR 1 and LRR 2 priors are the best-performing priors 25.0 percent and 33.3 percent of the time, respectively. In comparison, the HF 2 prior only achieves 16.7 percent, and the remaining three priors are even lower. Additionally, the LRR 1 and LRR 2 priors are the second-best-performing priors in 29.2 percent and 25.0 percent, respectively, of all cases. These values are greater than for the other four priors.

We obtain a similar picture when loosening the priors by a factor of two, i.e. $\lambda = 2$ and $\underline{v} = 0.05$. The results are shown in Table 7. The LRR priors generally perform worse than priors derived form the other asset pricing models for the OOS period from 1947 to 1980. However, for the other sample periods, the LRR priors dominate, with LRR 1 being the best performing prior 33.3 percent of the time. The prior from the LRR 2 model is the best performing prior in 25.0 percent of all cases but the second best performing prior 41.7 percent of the time. These results are consistent with the benchmark parametrization that assumes $\lambda = 1$ and $\underline{v} = 0.1$.

6 Conclusion

Different theories have been proposed to resolve the equity premium puzzle (Mehra and Prescott (1985)). Three prominent consumption-based asset pricing models

Table 5: Priors derived from asset pricing models calibrated with data from 1926 to 1967

Panels A shows the OOS performance of the model-based priors derived from the three consumption-based asset pricing models: HF, LRR, and PT. The consumption-based asset pricing models are calibrated with annual data from 1926 to 1967. Appendix A contains details of the calibration. The priors are imposed on the single-variable predictive regression given in equation (1). Reported is the R^2_{OOS} (in percent) from equation (13), which measures the accuracy of OOS log equity premium forecasts of the single variable predictive regression relative to the historical average model. The predictors are the dividend-price ratio and the dividend yield. If the model-based prior leads to an increase in the R^2_{OOS} relative to OLS forecast in the "no prior" column, the figure is in bold. Panel B shows the difference between the R^2_{OOS} of the best-performing prior and each of the remaining five priors. The best performing-prior is reported in the second column. In both panels, the statistical significance of the difference in forecast errors is tested with a Diebold-Mariano test (see Diebold and Mariano (1995)). The test is one-sided for Panel B. A significant test statistic is denoted by ** at the 5 percent level and by * at the 10 percent level.

Panel A: Forecast performance

Frequency / predictor	**No Prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**
				Sample start: 1926 / OOS period: 1968-2014			
Annual returns							
Log dividend-price ratio	-1.158	-11.836	-11.181	**1.247**	**1.472**	-9.665	-23.733
Log dividend yield	-5.528	-10.248	-10.558	**0.370**	**0.809**	**-0.066**	**0.550**
Quarterly returns							
Log dividend-price ratio	-0.899	-2.220	-1.902	**0.149**	**0.024**	-5.778	-10.389
Log dividend yield	-0.163	**-0.128**	-0.227	**0.214**	**0.243**	-0.171	-0.927
Monthly returns							
Log dividend-price ratio	-0.165	-0.455	-0.244	**0.119**	**-0.082**	-1.289	-2.028
Log dividend yield	-0.379	**-0.284**	-0.588	**-0.106**	**-0.251**	-0.395	-0.698

Panel B: Prior comparison

Frequency / predictor	**Best prior**	**HF 1**	**HF 2**	**LRR 1**	**LRR 2**	**PT 1**	**PT 2**
				Sample start: 1926 / OOS period: 1968-2014			
Annual returns							
Log dividend-price ratio	LRR 2	13.309	12.654	0.225		11.137	25.205
Log dividend yield	LRR 2	11.057	11.368	0.439		0.875	0.260
Quarterly returns							
Log dividend-price ratio	LRR 1	2.370*	2.052		0.125	5.928**	10.539**
Log dividend yield	LRR 2	0.370	0.470	0.029		0.413	1.170
Monthly returns							
Log dividend-price ratio	LRR 1	0.573**	0.363		0.201	1.408*	2.147**
Log dividend yield	LRR 1	0.177	0.481*		0.145	0.288*	0.591**

that provide different explanations for the existence of the equity premium puzzle are the Habit Formation (HF), the Long Run Risk (LRR), and the Prospect Theory (PT) models. I analyze whether these asset pricing models can profitably guide the investment decisions of investors who try to time the equity market. I propose a simple Bayesian framework in which prior distributions on the parameters of a single-variable predictive regression are derived from the three asset pricing models. The investors update their prior beliefs with empirical data and predict the equity premium with valuation ratios — that is, the dividend-price ratio and the dividend yield.

The priors derived from the LRR model perform particularly well during the dot-com boom in the late 1990s. During that period, low valuation ratios predicted negative returns that failed to materialize for several years. The key to the strong performance of the LRR priors is the low implied predictive power of valuation ratios for the equity premium. Hence, an investor who uses the LRR model to guide her investment choices is hesitant to conclude that low valuation ratios result in an immediate fall in stock prices. The stronger predictability implied by the HF and PT models helps to improve forecast accuracy up to the 1980s. However, the performance deteriorates quickly during the dot-com boom, as the investors who believe in the strong predictive power of valuation ratios anticipate a sharp price decline much earlier than it materializes. Because the performance during the dot-com boom dominates, an investor whose prior beliefs are anchored in the LRR model would have outperformed investors whose prior beliefs stem from the HF and PT models in most sample periods. These differences in forecast accuracy are not only shown by differences in the R^2_{OOS}, but also translate into considerable utility gains for an investor with mean-variance preferences.

By imposing model-based priors derived from consumption-based asset pricing models on predictive regressions and showing how the performances of these priors differ, this paper makes novel contributions to the equity premium prediction literature and also adds to our understanding of consumption-based asset pricing models.

Table 6: Forecast performance of tighter model-based priors

Panels A, B, and C show the OOS performance of the model-based priors derived from the three consumption-based asset pricing models: HF, LRR, and PT. The priors are imposed on the single-variable predictive regression given in equation (1). Reported is the R^2_{OOS} (in percent) from equation (13), which measures the accuracy of OOS log equity premium forecasts of the single variable predictive regression relative to the historical average model. The predictors are the dividend-price ratio and the dividend yield. If the model-based prior leads to an increase in the R^2_{OOS} relative to OLS forecast in the "no prior" column, the figure is in bold. The last two columns denote which model-based prior leads to the greatest and second-greatest improvement in forecast accuracy. Panel D shows the number of times each model-based prior is the best or second best. The tightness of the priors is increased relative to the benchmark of $\lambda = 1$ and $\underline{v} = 0.1$ by setting $\lambda = 0.5$ and $\underline{v} = 0.2$.

Panel A: Annual returns

Sample start	OOS period	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	1st prior	2nd prior
Log dividend-price ratio										
1926	1947-2014	0.396	-3.973	-3.896	**-0.535**	**-0.056**	-9.778	-17.885	LRR 2	LRR 1
1926	1947-1980	11.895	**15.790**	**15.310**	-0.002	1.002	**15.473**	**15.641**	HF 1	PT 2
1926	1981-2014	-12.242	-25.786	-24.813	**-0.981**	**-0.882**	-37.621	-55.485	LRR 2	LRR 1
1947	1968-2014	-3.521	**-2.131**	**-1.271**	**0.281**	**1.370**	-6.049	-13.556	LRR 2	LRR 1
Log dividend yield										
1926	1947-2014	-16.280	**-5.320**	**-3.164**	**-1.088**	**0.216**	**1.529**	**1.338**	PT 1	PT 2
1926	1947-1980	-10.255	**11.874**	**12.018**	**0.723**	**1.492**	**5.116**	**6.521**	HF 2	HF 1
1926	1981-2014	-22.902	**-22.605**	**-20.132**	**-1.682**	**-1.358**	**-2.274**	**-3.112**	LRR 2	LRR 1
1947	1968-2014	-1.334	**-0.716**	**0.134**	**-0.177**	**0.601**	**2.537**	**3.283**	PT 2	PT 1

Panel B: Quarterly returns

Sample start	OOS period	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	1st prior	2nd prior
Log dividend-price ratio										
1926	1947-2014	-0.815	-1.065	-1.382	**0.740**	**1.028**	-15.619	-22.461	LRR 2	LRR 1
1926	1947-1980	3.753	**4.546**	**3.878**	2.562	3.084	-1.404	-5.016	HF 1	HF 2
1926	1981-2014	-4.727	-5.654	-6.536	**-1.078**	**-2.201**	-28.044	-38.214	LRR 1	LRR 2
1947	1968-2014	-0.391	-0.458	-0.760	**0.547**	**0.765**	-14.340	-21.854	LRR 2	LRR 1
Log dividend yield										
1926	1947-2014	0.340	**0.893**	**0.977**	**0.493**	**0.475**	**0.834**	**0.347**	HF 2	HF 1
1926	1947-1980	4.740	4.716	4.229	1.058	1.548	3.027	**5.423**	PT 2	HF 1
1926	1981-2014	-3.429	**-2.232**	**-2.103**	**-0.132**	**-0.505**	**-0.897**	-3.513	LRR 1	LRR 2
1947	1968-2014	-0.297	**1.017**	**1.003**	**0.732**	**0.743**	**1.061**	**0.675**	PT 1	HF 1

Table 6: Forecast performance of tighter model-based priors (continued)

Panel C: Monthly returns

Sample start	OOS period	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	1st prior	2nd prior
Log dividend-price ratio										
1926	1947-2014	-0.061	-0.347	-0.347	**0.175**	-0.124	-9.205	-10.724	LRR1	LRR2
1926	1947-1980	1.264	1.227	**1.343**	0.824	0.978	-5.103	-6.015	HF2	HF1
1926	1981-2014	-1.142	-1.453	-1.756	**-0.482**	**-0.445**	-12.645	-14.710	LRR2	LRR1
1947	1968-2014	-0.237	-0.320	-0.263	**0.217**	**0.118**	-11.671	-13.943	LRR1	LRR2
Log dividend yield										
1926	1947-2014	-0.393	**-0.083**	-0.523	**0.192**	**0.060**	**0.003**	**-0.275**	LRR1	LRR2
1926	1947-1980	1.175	1.088	**1.558**	0.641	**1.279**	**1.379**	**1.414**	HF2	PT2
1926	1981-2014	-1.672	-1.776	-2.141	**-1.145**	**-0.764**	**-0.984**	**-1.446**	LRR2	PT1
1947	1968-2014	-0.166	-0.399	-0.208	**0.282**	**0.135**	**0.109**	**-0.154**	LRR1	LRR2

Panel D: Summary of model-based prior performance

	Best-performing prior		Second-best-performing prior	
	#	in %	#	in %
HF 1	2	8.3	5	20.8
HF 2	4	16.7	1	4.2
LRR 1	6	25.0	7	29.2
LRR 2	8	33.3	6	25.0
PT 1	2	8.3	2	8.3
PT 2	2	8.3	3	12.5

Table 7: Forecast performance of looser model-based priors

Panels A, B, and C show the OOS performance of the model-based priors derived from the three consumption-based asset pricing models: HF, LRR, and PT. The priors are imposed on the single-variable predictive regression given in equation (1). Reported is the R^2_{OOS} (in percent) from equation (13), which measures the accuracy of OOS log equity premium forecasts of the single-variable predictive regression relative to the historical average model. The predictors are the dividend-price ratio and the dividend yield. If the model-based prior leads to an increase in the R^2_{OOS} relative to OLS forecast in the "no prior" column, the figure is in bold. The last two columns denote which model-based prior leads to the greatest and second-greatest improvement in forecast accuracy. Panel D shows the number of times each model-based prior is the best or second best. The tightness of the priors is reduced relative to the benchmark of $\lambda = 1$ and $\underline{v} = 0.1$ by setting $\lambda = 2$ and $\underline{v} = 0.05$.

Panel A: Annual returns

Sample start	OOS period	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	1st prior	2nd prior
Log dividend-price ratio										
1926	1947-2014	0.396	**0.493**	-0.074	**0.836**	**0.985**	-0.167	-1.591	LRR 2	LRR 1
1926	1947-1980	11.895	**14.002**	**14.637**	7.895	8.507	**13.798**	**14.978**	PT 2	HF 2
1926	1981-2014	-12.242	-15.678	-13.389	**-6.518**	**-7.753**	-15.967	-19.843	LRR 1	LRR 2
1947	1968-2014	-3.521	**-2.139**	**-1.633**	**2.109**	**1.424**	-3.716	-5.555	LRR 1	LRR 2
Log dividend yield										
1926	1947-2014	-16.280	**-5.358**	**-5.479**	**-0.889**	**-1.537**	**-1.666**	**-2.849**	LRR 1	LRR 2
1926	1947-1980	-10.255	**7.847**	**8.435**	**8.860**	**9.109**	**9.173**	**9.395**	PT 2	PT 1
1926	1981-2014	-22.902	**-20.556**	**-20.715**	**-12.604**	**-11.729**	**-13.492**	**-14.987**	LRR 2	LRR 1
1947	1968-2014	-1.334	**-0.462**	**-0.816**	**0.691**	**1.494**	**1.194**	**1.103**	LRR 2	PT 1

Panel B: Quarterly returns

Sample start	OOS period	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	1st prior	2nd prior
Log dividend-price ratio										
1926	1947-2014	-0.815	**-0.716**	-0.947	**-0.382**	**-0.713**	-1.912	-2.298	LRR 1	LRR 2
1926	1947-1980	3.753	**4.085**	**4.163**	**4.405**	**4.132**	3.535	3.194	LRR 1	HF 2
1926	1981-2014	-4.727	-5.068	-5.308	**-4.119**	**-4.315**	-6.708	-6.833	LRR 1	LRR 2
1947	1968-2014	-0.391	**-0.346**	**-0.244**	**0.197**	**0.084**	-1.445	-1.895	LRR 1	LRR 2
Log dividend yield										
1926	1947-2014	0.340	**0.342**	0.337	**0.399**	**0.666**	**0.490**	0.144	LRR 2	PT 1
1926	1947-1980	4.740	**4.765**	**4.865**	4.628	4.607	**4.824**	**4.814**	HF 2	PT 1
1926	1981-2014	-3.429	**-3.296**	**-3.059**	**-3.290**	**-2.525**	**-3.129**	-4.064	LRR 2	HF 2
1947	1968-2014	-0.297	**0.039**	**0.227**	**0.168**	**0.142**	**0.132**	**-0.050**	HF 2	LRR 1

Table 7: Forecast performance of looser model-based priors (continued)

Panel C: Monthly returns

Sample start	OOS period	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	1st prior	2nd prior
Log dividend-price ratio										
1926	1947-2014	-0.061	-0.087	**0.112**	-0.172	-0.087	-0.394	-0.343	HF2	LRR2
1926	1947-1980	1.264	1.048	1.250	1.006	1.156	**1.308**	1.113	PT1	HF2
1926	1981-2014	-1.142	-1.273	-1.190	-1.247	**-1.027**	-1.727	-1.659	LRR2	HF2
1947	1968-2014	-0.237	-0.342	**-0.064**	**-0.235**	**-0.201**	-0.729	-0.584	HF2	LRR2
Log dividend yield										
1926	1947-2014	-0.393	-0.603	**-0.361**	**-0.388**	-0.433	**-0.323**	-0.586	PT1	HF2
1926	1947-1980	1.175	0.836	1.122	**1.287**	**1.252**	**1.489**	**1.225**	PT1	LRR1
1926	1981-2014	-1.672	-1.869	-2.028	-1.683	**-1.587**	**-1.580**	-1.676	PT1	LRR2
1947	1968-2014	-0.166	**-0.146**	-0.208	**-0.053**	**-0.133**	**-0.136**	-0.184	LRR1	LRR2

Panel D: Summary of model-based prior performance

	Best-performing prior		Second-best-performing prior	
	#	in %	#	in %
HF 1	0	0.0	0	0.0
HF 2	4	16.7	6	25.0
LRR 1	8	33.3	4	16.7
LRR 2	6	25.0	10	41.7
PT 1	4	16.7	4	16.7
PT 2	2	8.3	0	0.0

Appendix A Asset pricing models

A.1 By force of habit: A consumption-based explanation of aggregate stock market behavior

Campbell and Cochrane (1999) use a standard representative-agent consumption-based asset pricing model but add a slow-moving habit to the basic power utility function. This slow-moving habit leads to a time-varying risk premium that is higher at business cycle troughs than at peaks.

The agents are identical and maximize their utility given by

$$E\left[\sum_{t=0}^{\infty}\delta^t\frac{(C_t-X_t)^{(1-\gamma)}-1}{1-\gamma}\right], \tag{20}$$

where C_t is the consumption level, X_t is the level of habit, δ is the time discount factor, and γ is the risk aversion. A surplus consumption ratio $S_t \equiv (C_t - X_t)/C_t$ is defined — a small value of S_t indicates that the economy is in a bad state. The local curvature of this utility function is given by

$$\eta_t \equiv -\frac{C_t u_{cc}(C_t,X_t)}{u_c(C_t,X_t)} = \frac{\gamma}{S_t}. \tag{21}$$

A process is specified for $s_t = \ln(S_t)$, which ensures that C_t is always above X_t:

$$s_{t+1} = (1-\phi)\bar{s} + \phi s_t + \lambda(s_t)(c_{t+1} - c_t - g), \tag{22}$$

with ϕ reflecting habit persistence. The function $\lambda(s_t)$ takes the form

$$\lambda(s_t) = \begin{cases} \frac{1}{\bar{S}}\sqrt{1-2(s_t-\bar{s})}-1, & s_t \leq s_{max} \\ 0, & s_t > s_{max}, \end{cases} \tag{23}$$

with the parameter s_{max} set equal to $\bar{s} + \frac{1}{2}(1-\bar{S}^2)$. The steady state value $\bar{s}$ is given by $\ln(\sigma\sqrt{\gamma/(1-\phi)})$. The evolution of s_{t+1} is based on consumption growth being an i.i.d. lognormal process

$$\Delta c_{t+1} = g + v_{t+1}, \text{ where } v_{t+1} \overset{i.i.d.}{\sim} N(0, \sigma_v^2). \tag{24}$$

Stocks represent a claim to the consumption stream. The price-consumption ratio for a consumption claim satisfies

$$\frac{P_t}{C_t}(s_t) = E_t\left[M_{t+1}\frac{C_{t+1}}{C_t}\left[1 + \frac{P_{t+1}}{C_{t+1}}(s_{t+1})\right]\right]. \tag{25}$$

The underlying assumption is that dividend growth is perfectly correlated with consumption growth in equation (24). Above, I denote this specification the HF 1 model.[6] The intertemporal marginal rate of substitution (IMRS) M_{t+1} takes the form

$$M_{t+1} \equiv \delta\left(\frac{S_{t+1}}{S_t}\frac{C_{t+1}}{C_t}\right)^{-\gamma}. \tag{26}$$

Because the term $(S_{t+1}/S_t)^{-\gamma}$ correlates positively with asset returns, the HF model generates a higher equity premium compared with the standard power utility model. The log risk-free rate is given by

$$r_t^f = -ln(\delta) + \gamma g - \gamma(1-\phi)(s_t - \bar{s}) - \frac{\gamma^2\sigma^2}{2}[1 + \lambda(s_t)]^2. \tag{27}$$

The price-consumption ratio is correlated with the business cycles, as it depends on s_t. The ratio is high at business cycle peaks and low at troughs. Why is the price-consumption ratio procyclical? Suppose there is a positive shock to consumption in period t. Higher consumption raises s_t and consequently $E_t[M_{t+1}]$, which results in a higher asset price and price-consumption ratio. (Equation (21) shows how an increase in s_t lowers the the local curvature of the utility function and makes

[6] The solution for the model specification which assumes imperfectly correlated consumption and dividend processes (HF 2) is given in Campbell and Cochrane (1999).

the agent less risk averse.) Because expected future cash flows remain constant, the higher asset prices will lead to lower expected returns. Hence, the price-consumption ratio and subsequent returns are inversely correlated.

A.1.1 Calibration and simulation of model

For my baseline results, priors from the HF model are based on the parameter values proposed by Campbell and Cochrane (1999). These parameter values are reported in Table A.1. in the "original value" column. My recalibration of the model with data from 1926 to 1967 results in parameter values reported in the "1926-1967 value" column. For the recalibration, consumption data are real per capita consumption of nondurables and services from the Bureau of Economic Analysis (BEA). The standard deviation of log consumption growth is chosen such that annual log consumption growth simulated from the model matches the empirical counterpart of 3.02%. The risk-free rate time series is from Amit Goyal's website and deflated with inflation data from Federal Reserve Economic Data. Dividends are computed using CRSP New York Stock Exchange (NYSE) data. The persistence of the log price-dividend ratio is 0.82. Following Campbell and Cochrane (1999), I chose γ to match the NYSE equity premium sharpe ratio, which is 0.33 for the 1926-1967 period, with the HF 1 specification. The discount factor δ is selected such that the annualized log risk-free rate matches the empirical value of 0.31.

I apply the fixed-point method to solve for the price-consumption and the price-dividend ratio (see Wachter (2005)). The model is simulated at a monthly frequency and time-aggregated to lower frequencies. Summary statistics of the simulation for the model specification with perfectly (HF 1) and imperfectly (HF 2) correlated log consumption and log dividend growth are given in Panel A of Table A.2. The simulated moments match the moments obtained by Campbell and Cochrane (1999) and Wachter (2005). The simulated moments based on the model recalibrated with data from 1926-1967 can be found in Panel B of Table A.2.

Table A.1: Habit Formation model parameter values

The parameter values from Campbell and Cochrane (1999) are reported in the "original value" column. The parameter values chosen for the calibration of the model based on data from 1926 to 1967 are reported in the "1926-1967 value" column. A * denotes that the value is annualized.

Description	**Variable**	**Original value**	**1926-1967 value**
Mean log consumption growth*	g	1.89%	1.77%
Std. dev. log consumption growth*	σ	1.50%	3.75%
Log risk-free rate*	r^f	0.94%	0.31%
Persistence coefficient*	ϕ	0.87	0.82
Utility curvature	γ	2.00	1.00
Std. dev. log dividend growth*	σ_ω	11.2%	14.3%
Corr. log cons. and log div. growth	ρ	0.20	0.57
Subjective discount factor*	δ	0.89	0.92

Table A.2: Habit Formation model simulated moments

Simulated moments at monthly, quarterly, and annual frequencies that are reported for the specifications of the HF model that assume perfect (HF 1) and imperfect correlation (HF 2) between log consumption and log dividend growth. For Panel A, the parameter values of Campbell and Cochrane (1999) are used. For Panel B, the parameter values are estimated based on a sample with data from 1926 to 1967. The price-dividend ratio moments are annualized.

Panel A: Original parameter values						
Model	**Freq.**	**P/D Mean**	**Log P/D Std. dev.**	**Log equity prem. Mean**	**Log equity prem. Std. dev.**	**Log Sharpe ratio**
HF 1	Annual	18.55	0.27	6.60%	15.06%	0.44
HF 2	Annual	19.00	0.30	6.52%	19.91%	0.33
HF 1	Quarterly	18.43	0.27	1.65%	7.73%	0.21
HF 2	Quarterly	18.92	0.28	1.63%	10.08%	0.16
HF 1	Monthly	18.39	0.27	0.55%	4.49%	0.12
HF 2	Monthly	18.89	0.28	0.54%	5.84%	0.09
Panel B: 1926-1967 parameter values						
HF 1	Annual	17.32	0.35	7.76%	23.71%	0.33
HF 2	Annual	17.14	0.40	7.92%	31.59%	0.25
HF 1	Quarterly	17.20	0.35	1.95%	12.26%	0.16
HF 2	Quarterly	17.01	0.37	1.97%	16.17%	0.12
HF 1	Monthly	17.12	0.34	0.65%	7.12%	0.09
HF 2	Monthly	17.03	0.37	0.66%	9.38%	0.07

A.2 Prospect theory and asset prices

In the model of Barberis et al. (2001), the agent not only derives utility from consumption but also from financial wealth fluctuations. There are two important aspects in the way financial wealth fluctuations affect the utility of an economic agent. First, the agent is loss averse. Second, the degree of loss aversion depends on prior investment outcomes. Prior gains lead to less loss aversion, and prior losses lead to more loss aversion. Hence, the risk aversion of the agent varies over time.

Aggregate consumption growth and dividend growth follow the i.i.d. lognormal processes given by

$$\Delta c_{t+1} = g_c + \sigma_c \epsilon_{c,t+1}, \text{ where } \epsilon_{c,t+1} \overset{i.i.d.}{\sim} N(0,1) \tag{28}$$

and

$$\Delta d_{t+1} = g_d + \sigma_d \epsilon_{d,t+1}, \text{ where } \epsilon_{d,t+1} \overset{i.i.d.}{\sim} N(0,1), \tag{29}$$

with the correlation between $\epsilon_{c,t+1}$ and $\epsilon_{d,t+1}$ being denoted by ω.[7]

The agent's maximization problem is set up as

$$E\left[\sum_{t=0}^{\infty}\left(\rho^t \frac{C_t^{1-\gamma}}{1-\gamma} + b_0 \bar{C}_t^{-\gamma} \rho^{t+1} v(X_{t+1}, S_t, z_t)\right)\right]. \tag{30}$$

The second term captures the fact that the agent's utility is affected by fluctuations in financial wealth. The variable X_{t+1} denotes the change of the financial wealth between time t and $t+1$ and is defined as

$$X_{t+1} \equiv S_t R_{t+1} - S_t R_{f,t}. \tag{31}$$

The variable S_t measures the value of the agent's risky assets at time t. The variable

[7]Barberis et al. (2001) consider two different specifications: Economy I, in which dividends equal consumption, and Economy II, in which consumption and dividends follow separate but positively correlated processes. The simulated moments of Economy II are much more successful in matching the empirical moments; hence, I do not consider Economy I.

z_t accounts for prior gains and losses up to time t and is defined as Z_t/S_t, where Z_t is a historical benchmark level for the value of the risky asset. If z_t is smaller than one, the agent has prior gains; if z_t is greater than one, the agent faces prior losses. The time discount factor is ρ, and $b_0\bar{C}_t^{-\gamma}$ is a scaling term, with γ being the risk aversion over consumption. The form of the utility function over financial wealth $v(.)$ is different conditional on prior gains or prior losses.

The dynamics of z_t are given by the process

$$z_{t+1} = \eta\left(z_t\frac{\bar{R}}{R_{t+1}}\right) + (1-\eta). \tag{32}$$

This process ensures that the benchmark level Z_t reacts sluggishly to changes in the stock price. The parameter $\bar{R}$ is chosen such that the median value of z_t is around one.

The price-dividend ratio is assumed to be a function of the state variable z_t:

$$f_t \equiv P_t/D_t = f(z_t). \tag{33}$$

The real stock returns are thus given as

$$R_{t+1} = \frac{1+f(z_{t+1})}{f(z_t)}e^{g_d+\sigma_d\epsilon_{d,t+1}}. \tag{34}$$

Barberis et al. (2001) show that the equilibrium is characterized by a constant real risk-free rate,

$$R_f = \delta^{-1}e^{\gamma g_c-\gamma^2\sigma_c^2/2}, \tag{35}$$

and a price-dividend ratio determined by the equation

$$\begin{aligned}1 =&\delta e^{g_d-\gamma g_c+\gamma^2\sigma_c^2(1-\omega^2)/2}E_t\left[\frac{1+f(z_{t+1})}{f(z_t)}e^{(\sigma_d-\gamma\omega\sigma_c)\epsilon_{d,t+1}}\right]\\ &+ b_0\delta E_t\left[\hat{v}\left(\frac{1+f(z_{t+1})}{f(z_t)}e^{g_d+\sigma_d\epsilon_{d,t+1}}, z_t\right)\right],\end{aligned} \tag{36}$$

where the utility function $\hat{v}(R_{t+1}, z_t)$ is equal to $v(X_{t+1}, S_t, z_t)/S_t$ and specified for $z_t \leq 1$ as

$$\hat{v}(R_{t+1}, z_t) = \begin{cases} R_{t+1} - R_{f,t}, & R_{t+1} \geq z_t R_{f,t} \\ (z_t R_{f,t} - R_{f,t}) + \lambda(R_{t+1} - z_t R_{f,t}), & R_{t+1} < z_t R_{f,t} \end{cases} \tag{37}$$

and for $z_t > 1$ as

$$\hat{v}(R_{t+1}, z_t) = \begin{cases} R_{t+1} - R_{f,t}, & R_{t+1} \geq R_{f,t} \\ \lambda(z_t)(R_{t+1} - R_{f,t}), & R_{t+1} < R_{f,t}, \end{cases} \tag{38}$$

where $\lambda(z_t) = \lambda + k(z_t - 1)$ with $k > 0$.

The PT model generates an equity premium that is predictable by the dividend-price ratio. The mechanism works through time-varying risk aversion. A positive period t shock to dividends in equation (29) increases the return of the asset and leads to a lower z_t through equation (32). A lower z_t implies that the agent is less loss averse as shown in equations (37) and (38). Hence, the price of the asset will increase, which reduces the agent's loss aversion further, leading to a higher price-dividend ratio. Because of the higher prices and unchanged cash flow expectations, the expected returns are lower. Price-dividend ratios and future returns are therefore negatively related.

A.2.1 Calibration and simulation of model

The baseline parameter values from Barberis et al. (2001) are reported in Table A.3. in the "original value" column. My calibration of the model with data from 1926 to 1967 uses the parameter values in the "1926-1967 value" column. For some parameters, two values are given. In these cases, the first value corresponds to the PT 1 model. The PT 1 model is calibrated such that the average effective loss aversion of the model is 2.25. The second value corresponds to the PT 2 model, which is calibrated such that the log equity premium of the model matches the empirical moment. When calibrating the model with the 1926-1967 data sample, I use the the same consumption, dividend, and return data as for the calibration of the HF model, described previously.[8] The parameters γ, ρ, and δ are chosen to bring the risk-free rate close to the empirical value of 0.31%. The prior outcome parameter k and the time discount factor ρ are set to 4 and 0.98, respectively, for the PT 1 model such that the annual average effective loss aversion is 2.25. For the PT 2 model, the parameter values are chosen to be 18 and 0.99, respectively, to bring the annual simulated equity premium close to 7.42%. The persistence parameter η is set such that the persistence of the log price-dividend ratio is close to the empirical value of 0.82. The remaining parameters are not estimated with empirical data and set equal to the values of Barberis et al. (2001).

I solve the model by following the process laid out by Barberis et al. (2001). The moments in Panel A of Table A.4. are generated by simulating the model with the parameter values proposed by the authors, particularly $b0 = 100$ and $k = 3$ for PT 1 and $b0 = 100$ and $k = 8$ for PT 2. The moments match the moments obtained by Barberis et al. (2001). Panel B reports the simulated moments based on my recalibration of the parameter values with data from 1926 to 1967.

[8] I set σ_D equal to 12% for the 1926-1967 parameter values, as in Barberis et al. (2001), instead of 14.2% as in the HF model, as a convergence of the numerical solution was not achieved with a more volatile log dividend growth process.

Table A.3: Prospect Theory model parameter values

The parameter values from Barberis et al. (2001) are reported in column "original value". The parameter values chosen for the calibration of the model based on data from 1926 to 1967 are reported in the "1926-1967 value" column. When two values are given for the same parameter, then the first value stands for the PT 1 model and the second value for the PT 2 model. All values are annualized.

Description	**Variable**	**Original value**	**1926-1967 value**
Mean log consumption growth	g_c	1.84%	1.77%
Mean log dividend growth	g_d	1.89%	1.77%
Std. dev. log consumption growth	σ_c	3.79%	3.02%
Std. dev. log dividend growth	σ_d	12.0%	12.0%
Corr log cons. and log div. growth	ω	0.15	0.57
Utility curvature	γ	1.00	1.00
Time discount factor	ρ	0.98	0.98 / 0.99
Loss aversion	λ	2.25	2.25
Prior outcome parameter	k	3 / 8	4 / 18
Prospect utility weight	$b0$	100	100
Persistence factor	η	0.90	0.90

Table A.4: Prospect Theory model simulated moments

Simulated moments at monthly, quarterly, and annual frequencies are reported. In Panel A, the parameter values of Barberis et al. (2001) are used, particularly $b0 = 100$ and $k = 3$ for the PT 1 specification and $b0 = 100$ and $k = 8$ for the PT 2 specification. For Panel B, the parameter values are estimated based on a sample with data from 1926 to 1967, particularly $b0 = 100$ and $k = 4$ for the PT 1 specification and $b0 = 100$ and $k = 18$ for the PT 2 specification. The price-dividend ratio moments are annualized.

Model	**Freq.**	**Price-dividend ratio**		**Log equity prem.**		**Log Sharpe ratio**
		Mean	**Std. dev.**	**Mean**	**Std. dev.**	
Panel A: Original parameter values						
PT 1	Annual	17.30	2.38	3.74%	20.23%	0.19
PT 2	Annual	12.73	2.21	5.87%	23.87%	0.25
PT 1	Quarterly	9.46	0.54	2.13%	9.00%	0.24
PT 2	Quarterly	7.45	0.60	2.84%	10.79%	0.26
PT 1	Monthly	6.30	0.14	1.15%	4.48%	0.26
PT 2	Monthly	5.05	0.16	1.47%	5.05%	0.29
Panel B: 1926-1967 parameter values						
PT 1	Annual	16.99	2.48	3.90%	21.18%	0.18
PT 2	Annual	12.30	2.54	7.47%	28.54%	0.26
PT 1	Quarterly	9.45	0.57	2.12%	9.34%	0.23
PT 2	Quarterly	6.73	0.67	3.51%	12.99%	0.27
PT 1	Monthly	6.32	0.16	1.15%	4.65%	0.25
PT 2	Monthly	4.35	0.18	1.85%	5.85%	0.32

A.3 Risks for the long run: a potential resolution of asset pricing puzzles

Bansal and Yaron (2004) propose a solution to the equity premium puzzle through a consumption-based asset pricing model with Epstein and Zin (1989) preferences. Their model differs from other consumption-based asset pricing models in two ways. First, they include a small persistent expected growth rate component in the consumption and dividend growth rate processes. This component causes consumption and the return on the market portfolio to covary positively, and hence, the economic agents require a higher risk premium. Second, they allow for time-varying volatility, which accounts for fluctuating economic uncertainty, in both processes: this additional source of systematic risk increases the risk premium further.

The asset pricing restriction for the real return on the market portfolio $R_{m,t+1}$, according to the Epstein and Zin (1989) preferences, is

$$E_t\left[\delta^\theta G_{c,t+1}^{-\frac{\theta}{\psi}} R_{c,t+1}^{-(1-\theta)} R_{m,t+1}\right] = E_t\left[M_{t+1} R_{m,t+1}\right] = 1, \tag{39}$$

where $G_{c,t+1}$ is the aggregate gross growth rate of consumption, $R_{c,t+1}$ denotes the real return on an asset that pays aggregate consumption as dividends, δ is the time discount factor, and M_{t+1} is the IMRS. The parameter θ is defined as $(1-\gamma)/(1-\frac{1}{\psi})$, where γ is the risk aversion parameter, and ψ accounts for the intertemporal elasticity of substitution (IES). To derive the real returns, the authors use the standard approximation of Campbell and Shiller (1988). The real log return for the claim to aggregate consumption is

$$r_{c,t+1} = \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{c,t+1}, \tag{40}$$

where $g_{c,t+1}$ is the log consumption growth, and z_t denotes the log price-consumption ratio. The specification for the real log return on the market portfolio is

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m} z_{m,t+1} - z_{m,t} + g_{d,t+1}, \tag{41}$$

where $g_{d,t+1}$ is the log dividend growth rate, and $z_{m,t}$ denotes the log price-dividend ratio. The values for κ_0, $\kappa_{0,m}$, κ_1, and $\kappa_{1,m}$ are constants that are derived through the approximation of Campbell and Shiller (1988).[9]

The dynamics of log consumption growth and log dividend growth — which incorporate a small persistent predictable component x_t, the long run risk component, and a time-varying volatility component σ_t, reflecting fluctuating economic uncertainty — are

$$\begin{aligned} x_{t+1} =& \rho x_t + \varphi_e \sigma_t e_{t+1} \\ g_{c,t+1} =& \mu_c + x_t + \sigma_t \eta_{t+1} \\ g_{d,t+1} =& \mu_d + \phi x_t + \varphi_d \sigma_t u_{t+1} \\ \sigma^2_{t+1} =& \sigma^2 + v_1(\sigma^2_t - \sigma^2) + \sigma_w w_{t+1}, \end{aligned} \tag{42}$$

with e_{t+1}, u_{t+1}, η_{t+1}, and w_{t+1} having i.i.d. standard Normal distributions.[10] The state variables, which determine the price-consumption and price-dividend ratios, are x_t and σ_t. The solutions for z_t and $z_{m,t}$ are

$$\begin{aligned} z_t =& A_0 + A_1 x_t + A_2 \sigma^2_t \\ z_{m,t} =& A_{0,m} + A_{1,m} x_t + A_{2,m} \sigma^2_t. \end{aligned} \tag{43}$$

The derivation of A and A_m can be found in Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2010).

The model generates excess returns that are predictable by the price-dividend

[9]Bansal, Kiku, and Yaron (2010) show that κ_1 is equal to $exp(\bar{z})/(1+exp(\bar{z}))$, and κ_0 is equal to $ln(1+exp(\bar{z})) - \kappa_1 \bar{z}$, where $\bar{z}$ is the mean log price-consumption ratio. Accordingly, $\kappa_{1,m}$ is given by $exp(\bar{z}_m)/(1+exp(\bar{z}_m))$, and $\kappa_{0,m}$ is equal to $ln(1+exp(\bar{z}_m)) - \kappa_{1,m}\bar{z}_m$, with $\bar{z}_m$ being the mean log price-dividend ratio.

[10]Bansal and Yaron (2004) also simulate a version of their model without time-varying volatility of consumption growth, which is less successful in matching empirical data moments.

ratio, but the predictability is weak. The predictability is affected by the two state variables σ_t^2 and x_t. A negative shock to σ_t^2 results in a lower $E_t[R_{c,t+1}]$, which causes $E_t[M_{t+1}]$ to increase. Consequently, asset prices and price-dividend ratios are both higher. The higher prices cause a decrease in expected returns, and thus, a negative correlation between the price-dividend ratio and future returns. A positive shock to x_t also causes an increase in $E_t[M_{t+1}]$ as $E_t[G_{t+1}]$ goes up: asset prices and price-dividend ratios increase. However, dividends in subsequent periods will be higher because of the positive shock to the growth rate. Thus, high price-dividend ratios are followed by higher cash flows which weakens the predictive power of the price-dividend ratio for returns.

A.3.1 Calibration and simulation of model

The baseline parameter values used by Bansal and Yaron (2004) are reported in Table A.5. in the "original value" column. My calibration of the model over the 1926-1967 sample uses the parameter values in the "1926-1967 value" column. For the risk aversion parameter γ two values are given. The first value corresponds to the LRR 1 model. The LRR 1 model yields a simulated price-dividend ratio that is close to the empirical moment. The second value corresponds to the LRR 2 model, which matches the empirical log equity premium closely. For the calibration with the 1926-1967 sample, I use the same consumption, dividend, and return data as for the calibration of the HF model, described previously.[11] Following Bansal and Yaron (2004), the parameters μ, μ_d, ρ, φ_e, ϕ, φ_d, and σ, are chosen such that the model can replicate the log consumption growth and log dividend growth dynamics of the annual empirical data, as well as producing a price-dividend ratio (LRR 1) and an equity premium (LRR 2) that are close to their empirical counterparts of 22.34 and 7.42%, respectively. For the 1926-1967 sample, log consumption growth has a mean of 1.80% and a standard deviation of 3.08% with an autocorrelation of 0.32. The

[11] Bansal and Yaron (2004) assume consumption takes place at the end of a period. I assume the same timing convention.

variance ratios at the 2, 5, and 10 year horizon are 1.35, 1.32, and 1.37, respectively. The log dividend growth has a standard deviation of 14.27% and an autocorrelation of -0.03. The correlation between log consumption and log dividend growth is 0.57. The parameters of the economic uncertainty process v_1 and σ_w are selected such that predictable variation of consumption volatility with the log price-dividend ratio is 3% as in the empirical data.

Panel A of Table A.6. reports the moments of the simulated data from the LRR model for $\gamma = 7.5$ (LRR 1) and $\gamma = 10$ (LRR 2) when the Bansal and Yaron (2004) parameter values are used. The simulation is based on the analytical solutions of the model. The analytical solutions are considered more reliable than the numerical solutions (see, for example, Bansal et al. (2007) and Beeler and Campbell (2012)). The model is simulated at a monthly frequency and time-aggregated to lower frequencies. The obtained data moments match the data moments in Bansal and Yaron (2004) and Beeler and Campbell (2012). Panel B of Table A.6. reports the simulated moments based on my recalibration of the model with data from 1926 to 1967.

Table A.5: Long Run Risk model parameter values

The parameter values from Bansal and Yaron (2004) are reported in the "original value" column. The parameter values chosen for the calibration of the model based on data from 1926 to 1967 are reported in the "1926-1967 value" column. When two values are given for the same parameter, then the first value stands for the LRR 1 model and the second value for the LRR 2 model. A * denotes that the value is at a monthly frequency.

Description	Variable	Original value	1926-1967 value
Mean log consumption growth*	μ_c	0.0015	0.0015
Mean log dividend growth*	μ_d	0.0015	0.0015
Persistence of x_t*	ρ	0.979	0.977
Volatility multiple of x_t*	φ_e	0.044	0.049
Dividend leverage*	ϕ	3.00	3.70
Dividend volatility multiple*	φ_d	4.50	4.80
Unconditional mean of σ_t*	σ	0.0078	0.0083
Persistence of σ_t*	v_1	0.987	0.987
Baseline volatility*	σ_w	0.23×10^{-5}	0.23×10^{-5}
Risk aversion	γ	7.5 / 10	7.5 / 10
IES	ψ	1.50	1.50
Time discount factor*	δ	0.9880	0.9885

Table A.6: Long Run Risk model simulated moments

Simulated moments at monthly, quarterly, and annual frequencies are reported for the specifications of the LRR model with $\gamma = 7.5$ (LRR 1) and $\gamma = 10$ (LRR 2). For Panel A, the parameter values of Bansal and Yaron (2004) are used. For Panel B, the parameter values are estimated based on a sample with data from 1926 to 1967. The price-dividend ratio moments are annualized.

Model	Freq.	P/D Mean	Log P/D Std. dev.	Log equity prem. Mean	Log equity prem. Std. dev.	Log Sharpe ratio
Panel A: Original parameter values						
LRR 1	Annual	26.86	0.20	2.70%	16.75%	0.16
LRR 2	Annual	20.61	0.20	4.08%	16.46%	0.25
LRR 1	Quarterly	26.68	0.17	0.67%	8.32%	0.08
LRR 2	Quarterly	20.43	0.17	1.03%	8.22%	0.13
LRR 1	Monthly	26.65	0.16	0.23%	4.81%	0.05
LRR 2	Monthly	20.44	0.16	0.35%	4.76%	0.07
Panel A: 1926-1967 parameter values						
LRR 1	Annual	23.10	0.27	4.13%	21.10%	0.20
LRR 2	Annual	16.46	0.26	6.17%	20.50%	0.30
LRR 1	Quarterly	22.79	0.23	1.04%	10.56%	0.10
LRR 2	Quarterly	16.31	0.22	1.58%	10.30%	0.15
LRR 1	Monthly	22.72	0.22	0.35%	6.09%	0.06
LRR 2	Monthly	16.27	0.21	0.52%	5.95%	0.09

References

Avramov, D., S. Cederburg, and K. Lucivjanska. 2016. Are stocks riskier over the long run? Taking cues form economic theory. Unpublished working paper..

Baker, M., and J. Wurgler. 2000. The equity share in new issues and aggregate stock returns. *Journal of Financial Economics* 55:2219-2257.

Bansal, R., D. Kiku, and A. Yaron. 2010. Risks for the long run: estimation and inference. Unpublished working paper.

Bansal, R., and A. Yaron. 2004. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance* 59:1481-1509.

Barberis, N. 2000. Investing for the long run when returns are predictable. *Journal of Finance* 55:225-264.

Barberis, N., M. Huang, and T. Santos. 2001. Prospect theory and asset prices. *Quarterly Journal of Economics* 116:1-53.

Beeler, J., and J. Y. Campbell. 2012. The long run risks model and aggregate asset prices: An empirical assessment. *Critical Finance Review* 1:141-182.

Brandt, M. W., A. Goyal, P. Santa-Clara, and J. R. Stroud. 2005. A simulation approach to dynamic portfolio choice with an application to learning about return predictability. *Review of Financial Studies* 18:831-873.

Campbell, J. Y. 1987. Stock returns and the term structure. *Journal of Financial Economics* 18:373-399.

Campbell, J. Y., and R. J. Shiller. 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies* 1:195-228.

Campbell, J. Y., and J. H. Cochrane. 1998. By force of habit: A consumption-based explanation of aggregate stock market behavior. Center for Research in Security Prices Working Paper No. 412.

Campbell, J. Y., and J. H. Cochrane. 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107:205-251.

Campbell, J. Y., and S. B. Thompson. 2008. Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies* 21:1509-1531.

Campbell, J. Y., S. Giglio, and C. Polk. 2013. Hard times. *Review of Asset Pricing Studies* 3:95-132.

Del Negro, M., and F. Schorfheide. 2011. Bayesian macroeconometrics. The Oxford Handbook of Bayesian Econometrics.

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.

Fama, E. F., and K. French. 1988. Dividend yields and expected stock returns. *Journal of Financial Economics* 22:3-25.

Fama, E. F., and K. French. 1989. Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics* 25:23-49.

Koop, G. 2003. Bayesian econometrics. John Wiley & Sons, Ltd, England.

Kruttli, M. S., A. J. Patton, and T. Ramadorai. 2015. The impact of hedge funds on asset markets. *Review of Asset Pricing Studies* 5:185-226.

Lettau, M., and S. Ludvigson. 2001. Consumption, aggregate wealth, and expected stock returns. *Journal of Finance* 56:815-849.

Li, Y., D. T. Ng, and B. Swaminathan. 2013. Predicting market returns using aggregate implied cost of capital. *Journal of Financial Economics* 110:419-436.

Litterman, R. B. 1986. Forecasting with Bayesian vector autoregressions: Five years of experience. *Journal of Business & Economic Statistics* 4:25-38.

Mehra, R., and E. C. Prescott. 1985. The equity premium: A puzzle. *Journal of Monetary Economics* 15:145-161.

Pastor, L., and R. F. Stambaugh. 2009. Predictive systems: Living with imperfect predictors. *Journal of Finance* 64:1583-1628.

Pastor, L., and R. F. Stambaugh. 2012. Are stocks really less volatile in the long run? *Journal of Finance* 67:431-478.

Pettenuzzo, D., A. Timmermann, and R. Valkanov. 2014. Forecasting stock returns under economic constraints. *Journal of Financial Economics* 114:517-553.

Polk, C., S. Thompson, and T. Vuolteenaho. 2006. Cross-sectional forecasts of the equity premium. *Journal of Financial Economics* 81:101-141.

Shanken, J. A., and A. Tamayo. 2012. Payout yield, risk, and mispricing: A Bayesian analysis. *Journal of Financial Economics* 105:131-152.

Stambaugh, R. F. 1999. Predictive regressions. *Journal of Financial Economics* 54:375-421.

Wachter, J. A. 2005. Solving models with external habit. NBER Working Paper No. 11559.

Wachter, J. A., and M. Warusawitharana. 2009. Predictable returns and asset allocation: should a skeptical investor time the market? *Journal of Econometrics* 148:162-178.

Wachter, J. A., and M. Warusawitharana. 2015. What is the chance that the equity premium varies over time? Evidence from regressions on the dividend-price ratio. *Journal of Econometrics* 186:74-93.

Welch, I., and A. Goyal. 2008. A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies* 21:1455-1508.