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

Mathias S. Kruttli

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: September 26, 2016

Abstract

This paper compares consumption-based asset pricing models based on the forecasting performance of investors who use economic constraints derived from the models to predict the equity premium. Three prominent asset pricing models are considered: Habit Formation, Long Run Risk, and Prospect Theory. I propose a simple Bayesian framework through which the investors impose the economic constraints as model-based priors on the parameters of their predictive regressions. 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

---

[†]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 academics and investors alike. For academics, the predictability of stock returns is important for testing market efficiency. For investors, 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 (1987); Campbell and Shiller (1988); 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, fluctuations in the consumption-wealth ratio, 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 an investor could have used these predictors to forecast the equity premium OOS. Campbell and Thompson (2008) further investigate these findings by imposing economic constraints when estimating the single-variable predictive regressions. They apply sign restrictions on the parameter estimates of the predictive regression and a non-negativity restriction on the forecast of the equity premium. These restrictions help the investor to reduce uncertainty about the regression parameters. 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 on the parameters of single-variable predictive regressions that are typically used in the equity premium prediction literature. I propose a simple Bayesian econometric framework to implement these economic constraints as prior

distributions on the parameters. These prior distributions are named model-based priors. My approach relates to the macroeconometric literature, in which prior distributions from dynamic stochastic general equilibrium models (DSGE) are imposed on vector autoregressions (VAR) (see, for example, Ingram and Whiteman (1994) and Del Negro and Schorfheide (2011)). Instead of macroeconomists who use priors from DSGE models to improve their forecasts of macroeconomic variables, I consider investors who use model-based priors from consumption-based asset pricing models to forecast the equity premium. The three consumption-based asset pricing models that act as sources for the model-based priors are the Habit Formation (HF) model of Campbell and Cochrane (1999), the Prospect Theory (PT) model of Barberis, Huang, and Santos (2001), and the Long Run Risk (LRR) model of Bansal and Yaron (2004). I chose these three models, as they propose different theories that can explain the equity premium puzzle (Mehra and Prescott (1985)). Also, the respective authors emphasize the equity premium predictability implied by their models and calibrate their models with similar US data.[1]

The model-based priors allow me to assess whether an investor could have profited from knowing the theories and the theories' 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 regression that stems from one of the asset pricing models. The investor then updates her beliefs about the predictive regression parameters with empirical data and forecasts 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

---

[1]Barro (2006) explains the equity premium puzzle through a disaster risk model, but unlike the three asset pricing models used in this paper, the calibration is based on international data on large economic declines, and the model-implied predictability of the equity premium is not analyzed.

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 asset pricing models' descriptions of the macro-finance world 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 (see, for example, Bansal, Gallant, and Tauchen (2007) for a comparison of the HF and the LRR models and Ludvigson (2012) for a survey of the literature).[2]

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 type of Bayesian framework on the predictive regressions. 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 the risk-return tradeoff and on the extent to which mispricing drives predictability. 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 initially 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 regression parameters through uninformative priors (see, for example,

---

[2]The analytical solutions and the empirically observable state variables of the LRR model allow for additional model evaluations: the in-sample estimation proposed by Bansal, Kiku, and Yaron (2010) and Constantinides and Ghosh (2012) and the OOS fit proposed by Ferson, Nallareddy, and Xie (2013).

Stambaugh (1999); Barberis (2000); Brandt, Goyal, Santa-Clara, and Stroud (2005); Penasse (2016)) 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 accuracy of hypothetical investors who had access to the three asset pricing models from 1926 onward and use the model-based priors to reduce the uncertainty about the parameters of the predictive regression. The investors try to time the market by forecasting the equity premium with either the dividend-price ratio or the dividend yield.[3] For my benchmark analysis, the calibration of the asset pricing models is the same as presented by the authors in the respective publications of the models.[4] However, the 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 performances 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 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 (see, for example, Lettau and Ludvigson (2005)). The LRR model implies a

---

[3]As in Welch and Goyal (2008) and Pettenuzzo et al. (2014), the dividend-price ratio is defined as dividends divided by price, and the dividend yield is defined as dividends divided by price lagged by one period.

[4]Because the authors of the asset pricing models use almost identical data sets for the calibration of their respective models, the comparison of the model-based priors' forecasting performances should not be distorted.

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 more skeptical to conclude that low valuation ratios imply an immediate decline in stock prices. This skepticism 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 as, 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. Relative to the PT model-based priors, the investor would on average be willing to pay 67 basis points per year to have access to the LRR model-based priors.

The implied predictability of the equity premium differs across the asset pricing models due to the model-specific mechanisms that lead to time-variation in valuation ratios. In the HF and PT models, changes in the valuation ratios are driven by time-varying discount rates, and this mechanism leads to a predictable equity premium. In the LRR model, the discount rate channel similarly leads to fluctuations in the valuation ratios and return predictability. However, the LRR model also incorporates a predictable component in the dividend growth rate, the long-run risk component, which drives valuation ratios and mitigates their predictive power for the equity premium. There exists considerable debate about whether the empirically observed changes in valuation ratios are driven by time-variation of discount rates or time-variation in the forecasts of dividend growth (see, for example, Campbell and Shiller (1988); Lettau and Ludvigson (2005); Bansal, Kiku, and Yaron (2007 and 2012); Cochrane (2008)). This paper shows that from the perspective of an investor who tries to time the market OOS over the 1926-2014 sample, model-based

priors derived from an asset pricing model that accounts for changes in valuation ratios due to time-varying dividend growth forecasts are preferred because of the lower equity premium predictability that the model implies.

The rest 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 log 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:

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

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 fore-casts 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 derived from the asset pricing model on β and σ_ϵ^2. These model-based constraints reduce uncertainty about the predictive regression parameters and 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)), which has the advantage of being a tractable prior distribution. The prior distribution is 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

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

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 analysis. Section 5 reports results for different values of λ and shows that these are in line with the benchmark case. The sample moments $\sigma_{r,t}^2$ and $\sigma_{x,t}^2$ are scaling factors, which ensure 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 to 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 σ_ϵ^{-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. These two processes drive the state variables of the models, and the state variables determine the dividend-price ratio. By simulating random shocks, time series of consumption growth and dividend growth are generated, based on which I solve the models in each period for the log equity premium, the log dividend-price ratio, and the log

dividend yield. The log dividend-price ratio is the difference between the log of dividends and the log of prices, and the log dividend yield is the difference between the log of dividends and the log of prices lagged by one period.[5] (A more detailed description of the models and how to solve them is provided in Appendix A.) I denote the simulated log equity premium in period $t_M + 1$ as $r_{M,t+1}$. I can then estimate the single-variable predictive regression that is generally used in the equity premium prediction literature and given in equation (1) with simulated data, where the simulated predictor $x_{M,t}$ is either the log dividend-price ratio or the log dividend yield:

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

The OLS estimates of $\beta_M = [\beta_{M,0}, \ \beta_{M,1}]'$ 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 predictive regression with simulated data in equation (12) is also estimated by the respective authors of the consumption-based asset pricing models, that is, Campbell and Cochrane (1999), Barberis et al. (2001), and Bansal and Yaron (2004), to assess the predictability of the equity premium implied by their proposed theories.

The model-based priors for my benchmark results are based on data simulated from the asset pricing models when using the same calibration as proposed by the authors in the respective published papers. The authors use almost identical calibration data sets, and thus, the comparison of the model-based priors' performances should not be distorted. I assume that the investors have no uncertainty about the parameters of the asset pricing models, since the focus of this paper (as in, for example, Campbell and Thompson (2008) and Pettenuzzo et al. (2014)) is the investors' uncertainty about the parameters of the predictive regression given in equation (1). There is a concern that the results are affected by an overlap of the OOS period

---

[5]The dividend-price ratio and the dividend yield are the only two predictors from the equity premium prediction literature that can be simulated from the three asset pricing models. Also, valuation ratios are the most prominent predictors of the equity premium prediction literature.

with the sample period used by the authors to calibrate the asset pricing models. To address this concern, I show the robustness of the results in Section 5, by re-calibrating the models with an empirical data sample that has no overlap with the OOS period. The calibration of the models is described in Appendix A.

For the HF model, the simulation is at a monthly frequency, and the quarterly and annual data are constructed by 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.[6]

The LRR model, like the HF model, is simulated at a monthly frequency, and the quarterly and annual values are time-averaged.[7] 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.[8] 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 model are greater than the estimates of the other two models and the empirical

---

[6]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.

[7]I simulate the model based on the analytical solutions as done by Bansal et al. (2010 and 2012) and Beeler and Campbell (2012).

[8]Adjusting the β_1 empirical estimates for the bias discussed by Stambaugh (1999) leads to annual, quarterly, and monthly estimates of 0.034, 0.011, and 0.002, respectively, for the log dividend-price ratio. For the dividend yield, the adjusted annual, quarterly, and monthly estimates are 0.079, 0.020, and 0.006, respectively. For the model-implied parameter estimates, a bias-adjustment is not necessary, because the data is simulated, and thus, finite sample issues do not apply.

estimates. The pattern for σ_ϵ^{*-2} is more mixed. However, the values implied by the asset pricing models are 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 different mechanisms of the three asset pricing models.

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.

Similar to the HF model, the PT model generates a time-varying dividend-price

14

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 between the LRR and the other two models is that the time-variation of the dividend-price ratio is partly driven by changes to expected dividend growth prospects. A positive shock to expected dividend growth leads to a lower dividend-price ratio that is followed by higher cash flows. This mechanism mitigates the predictive power of the dividend-price ratio that is generated by the economic uncertainty channel of the LRR model: because of the Epstein-Zin (see Epstein-Zin (1989)) preferences of the agents, a negative shock to time-varying economic uncertainty results in higher asset prices and lower dividend-price ratios, which reduces subsequent returns.

Whether the changes in valuation ratios are driven by time-variation in the forecasts of dividend growth or time-variation in discount rates is a source of considerable debate (see, for example, Lettau and Ludvigson (2005), Bansal et al. (2007 and 2012), and Cochrane (2008)). The former leads to weak predictability of returns, while the latter implies strong predictability of returns. This paper uses the asset pricing models and the implied mechanisms for the time-variation in valuation ratios dogmatically and investigates which mechanism leads to more accurate forecasts through the model-based priors. The investors who use the model-based priors to forecast the equity premium have no uncertainty about the asset pricing parameters. The investors' uncertainty is about the parameters of the predictive

15

regression in equation (1).

3   Results

In this section, I describe the data and report the OOS results when imposing
model-bases priors derived from asset pricing models on the single-variable predictive
regressions.

3.1   Data

The empirical data on the equity premium and the predictors at a monthly, quar-
terly, and annual frequency are available on Amit Goyal's website.[9] 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 performances 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 divi-
dend 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 div-
idend yield is defined as the difference between log dividends and log prices lagged
by one period.

---

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

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 DP 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 DY														
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

Panel B: Inverse variance of return innovation (σ_ϵ^{-2})

	Empirical	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
Log DP 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 DY							
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 %)

	Empirical	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
Log DP 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 DY							
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

3.2 Measuring forecast accuracy

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

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

where $\hat{r}^m_\tau$ is the equity premium forecast when imposing the model-based prior as given in equation (11); $\hat{r}^h_\tau$ 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 prior relative to the non-predictability model, which assumes that the best forecast of the equity premium is its historical average, that is, β_1 being set equal to zero in equation (1).

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 regression. Next, I analyze the subsample stability by splitting the 1947-2014 OOS forecast period 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, when the predictive regression in equation (1) is estimated via OLS. The valuation ratios predict a substantial time variation of the equity premium. The lower panel depicts the corresponding OLS 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 model-based priors for three re-
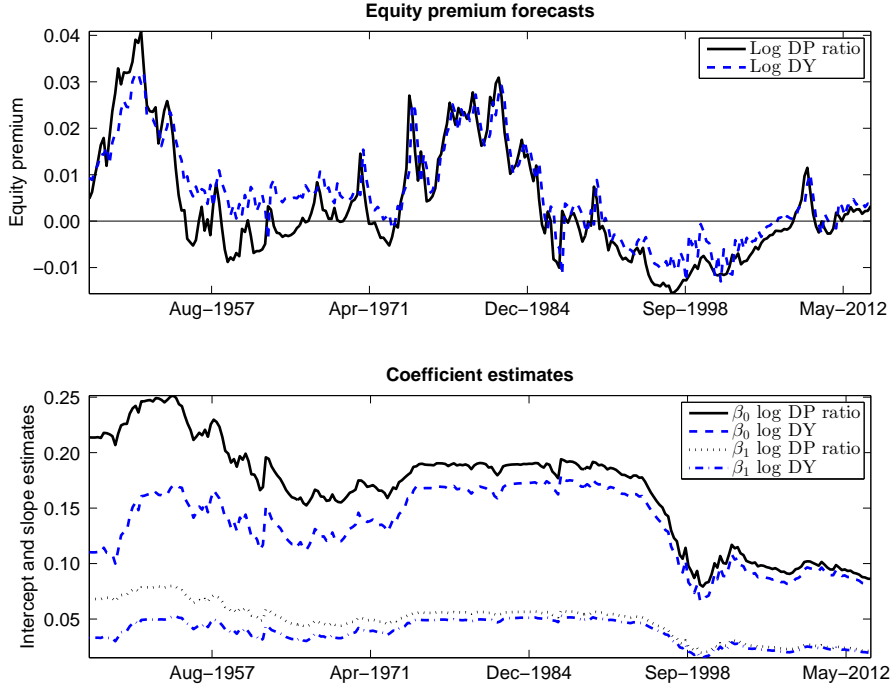
Figure 1: Empirical out-of-sample forecasts
The top panel shows the OOS quarterly log equity premium forecasts for two predictors: the log dividend-price ratio and the log dividend yield. The predictive regression in equation (1) is estimated recursively via OLS. The data sample starts in 1926 and the OOS period is from 1947 to 2014. The lower panel depicts the corresponding OLS coefficient estimates.

turn frequencies. Following the equity premium OOS forecasting literature, monthly, quarterly, and annual return frequencies are used. The "no prior" column reports the R^2_{OOS} for the case in which the single-variable predictive regression 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 column of the table show the best-performing prior for the respective frequency, predictor, and time period. Whether the differences in forecast errors between the predictive regression, estimated via OLS or the model-based priors, and the historical average model are significant is tested with a Diebold-Mariano test (see Diebold and Mariano (1995)).

Overall, the model-based priors help to improve the forecast accuracy of the single-variable predictive regression relative to the OLS estimates. The gains in R^2_{OOS} are considerable compared with the literature (see, for example, Campbell and Thompson (2008)). Out of the three asset pricing models, the priors derived

from the LRR model perform consistently better than the priors derived from the HF and PT models for three out of the four OOS periods. The LRR model-based priors yield less accurate forecasts than the other model-based priors only for the 1947-1980 OOS period. In most cases, the LRR model-based priors outperform the historical average model, which is shown by the positive R^2_{OOS} values. The exception is the 1981-2014 OOS period, for which the LRR model-based priors improve the forecast accuracy the most relative to the OLS estimates but fail to beat the historical average model. The Diebold-Mariano test leads to statistically significant results only for the 1947-1980 OOS period, for which we can reject the historical average model at the quarterly and annual frequency. The difficulty of statistically rejecting the historical average model when predicting the equity premium OOS is emphasized by Welch and Goyal (2008).

Table 3 compares the model-based priors by assessing their forecast errors against each other instead of comparing them to the forecast errors of the historical average model. The differences in the R^2_{OOS} (in percent) between the best-performing prior and the other priors are reported 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. 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 log 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

Table 2: Model-based prior forecast performance

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 log dividend-price ratio and the log 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 column denotes which model-based prior leads to the greatest improvement in forecast accuracy. The statistical significance of the difference between the forecast errors of the historical average model and the predictive regression is tested with a 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.

Panel A: Annual returns

Sample start	OOS period	Predictor	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	Best prior
1926	1947-2014	Log DP ratio	0.396	-1.209	-1.292	**1.137**	**0.607**	-4.045	-7.579	LRR 1
		Log DY	-16.280	**-4.923**	**-3.060**	**0.309**	**0.443**	**0.837**	**0.788**	PT 1
		Average	-7.942	**-3.066**	**-2.176**	**0.723**	**0.525**	-1.604	-3.396	LRR 1
1926	1947-1980	Log DP ratio	11.895*	15.342*	14.727**	3.927*	5.083*	16.355*	16.564	PT 2
		Log DY	-10.255	**11.488**	**11.530**	**4.417**	**5.449**	**6.698**	**8.389**	HF 2
		Average	0.820	**13.415**	**13.129**	**4.172**	**5.266**	**11.526**	**12.476**	HF 1
1926	1981-2014	Log DP ratio	-12.242	-18.486	-18.681	**-3.027**	**-2.689**	-25.224	-34.249	LRR 2
		Log DY	-22.902	-21.777	-19.763	**-4.760**	**-4.257**	**-6.115**	**-8.073**	LRR 2
		Average	-17.572	-20.132	-19.222	**-3.893**	**-3.473**	-15.670	-21.161	LRR 2
1947	1968-2014	Log DP ratio	-3.521	**-1.831**	**-1.606**	**2.237**	**2.313**	-4.978	-9.168	LRR 2
		Log DY	-1.334	**-0.069**	**-0.001**	**0.785**	**1.070**	**1.899**	**2.345**	PT 2
		Average	-2.428	**-0.950**	**-0.803**	**1.511**	**1.691**	-1.539	-3.411	LRR 2

Table 2: Model-based prior forecast performance (continued)

Panel B: Quarterly returns

Sample start	OOS period	Predictor	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	Best prior
1926	1947-2014	Log DP ratio	-0.815	**-0.607**	-1.020	**0.246**	**0.052**	-5.032	-7.551	LRR 1
		Log DY	0.340	**0.877**	**0.642**	**0.834**	**0.859**	**0.740**	**0.416**	HF 1
		Average	-0.238	**0.135**	**-0.189**	**0.540**	**0.456**	-2.146	-3.567	LRR 1
1926	1947-1980	Log DP ratio	3.753	**4.110**	3.680	3.334	3.375	1.752	0.306	HF 1
		Log DY	4.740*	**4.889****	4.253*	3.133**	3.621**	4.458**	**5.035***	PT 2
		Average	4.247	**4.499**	3.967	3.233	3.498	3.105	2.670	HF 1
1926	1981-2014	Log DP ratio	-4.727	-4.995	-5.074	**-2.882**	**-2.448**	-11.334	-14.032*	LRR 2
		Log DY	-3.429	**-3.031**	**-2.646**	**-1.706**	**-1.892**	**-2.442**	**-3.225**	LRR 1
		Average	-4.078	-4.013	**-3.860**	**-2.294**	**-2.170**	-6.888	-8.628	LRR 2
1947	1968-2014	Log DP ratio	-0.391	-0.541	-0.542	**0.640**	**0.619**	-5.060	-7.135	LRR 1
		Log DY	-0.297	**0.495**	**0.439**	**0.702**	**0.830**	**0.606**	**0.075**	LRR 2
		Average	-0.344	-0.023	-0.051	**0.671**	**0.724**	-2.227	-3.530	LRR 2

Panel C: Monthly returns

Sample start	OOS period	Predictor	No prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2	Best prior
1926	1947-2014	Log DP ratio	-0.061	-0.254	-0.075	**0.092**	**0.063**	-1.680	-1.656	LRR 1
		Log DY	-0.393	**-0.300**	**-0.310**	**-0.211**	**-0.267**	**-0.356**	-0.435	LRR 1
		Average	-0.227	-0.277	**-0.192**	**-0.059**	**-0.102**	-1.018	-1.046	LRR 1
1926	1947-1980	Log DP ratio	1.264	1.109	1.204	1.080	0.777	0.740	0.182	HF 2
		Log DY	1.175	1.149	**1.238**	1.014	1.046	1.026	1.217	HF 2
		Average	1.219	1.129	**1.221**	1.047	0.911	0.883	0.699	HF 2
1926	1981-2014	Log DP ratio	-1.142	-1.369	-0.996	**-0.937**	**-1.121**	-3.268	-3.450	LRR 1
		Log DY	-1.672	-1.944	-1.810	**-1.236**	**-1.272**	**-1.401**	-1.791	LRR 1
		Average	-1.407	-1.657	**-1.403**	**-1.087**	**-1.197**	-2.335	-2.620	LRR 1
1947	1968-2014	Log DP ratio	-0.237	-0.357	-0.093	**0.126**	**0.014**	-2.483	-2.784	LRR 1
		Log DY	-0.166	-0.253	-0.129	**-0.087**	**0.110**	**0.153**	**-0.038**	PT 1
		Average	-0.202	-0.305	-0.111	**0.019**	**0.062**	-1.165	-1.411	LRR 2

Section 4 shows that even small and statistically insignificant 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 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.

Figure 3 shows that the lower posterior estimates achieved through the LRR 1 prior are beneficial for an investor. The top panel depicts the difference between the cumulative sum of squared errors (SSE) of the historical average model and the single-variable predictive regression estimated via OLS or via model-based priors. I subtract the cumulative SSE of the predictive regression from the cumulative SSE of the historical average model. Hence, a positive value implies that the predictive regression 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

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

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 log dividend-price ratio and the log 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 in forecast errors 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 fourth column.

Panel A: Annual returns

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

24

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

Panel B: Quarterly returns

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

Panel C: Monthly returns

| Sample start	OOS period	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
1926	1947-2014	Log DP ratio	LRR 1	0.347*	0.167		0.029	1.773**	1.749**
		Log DY	LRR 1	0.089	0.099		0.057	0.146	0.224
1926	1947-1980	Log DP ratio	HF 2	0.096		0.124	0.427	0.464	1.023
		Log DY	HF 2	0.088		0.224	0.192	0.212	0.021
1926	1981-2014	Log DP ratio	LRR 1	0.432*	0.059		0.184	2.331**	2.512**
		Log DY	LRR 1	0.707**	0.574**		0.036	0.165	0.555**
1947	1968-2014	Log DP ratio	LRR 1	0.483**	0.218		0.112	2.608**	2.910**
		Log DY	PT 1	0.406**	0.283*	0.240*	0.043		0.191*
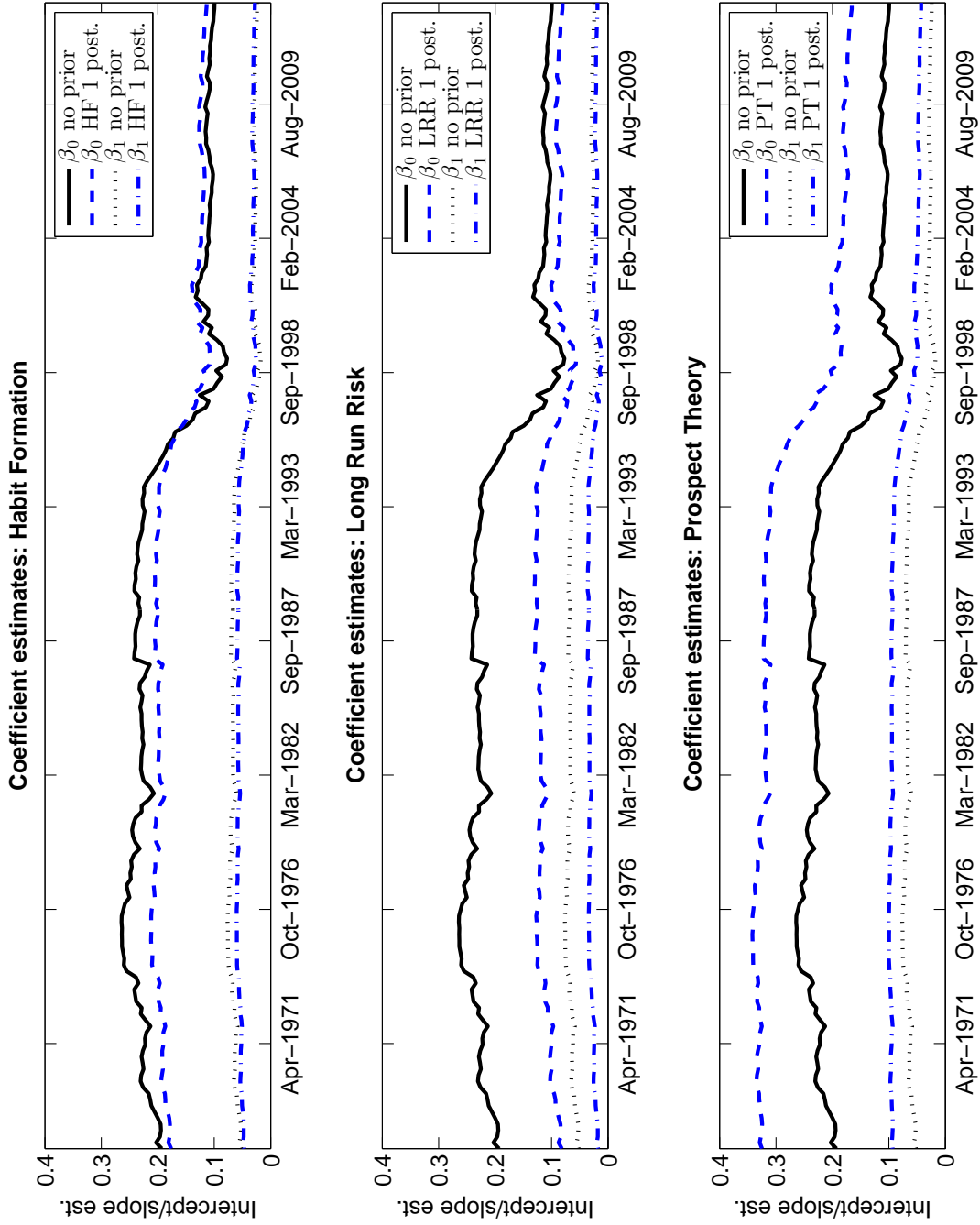
Figure 2: Posterior out-of-sample coefficient estimates
This figure shows the OLS estimates and the posterior estimates of β_0 and β_1 for the OOS period from 1968 to 2014 for the HF 1, LRR 1, and PT 1 priors. The predictor is the log dividend-price ratio, and the predictive regression in equation (1) is estimated with quarterly data.

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 regression turns negative for all four estimation methods. The investor armed with the LRR 1 model is able to limit poor forecasts, 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 estimates result in strongly negative forecasts. Hence, an investor relying on these forecasts to time the market suffers losses during this bull market period.

Some papers in the equity premium prediction literature restrict the model estimates to yield only non-negative predictions of the equity premium (see, for example, Campbell and Thompson (2008) and Pettenuzzo et al. (2014)). Such a restriction will lead to a result that is similar to imposing the LRR prior. However, an investor who derives a prior belief about the predictability of the equity premium through data simulated from the asset pricing models would conclude that valuation ratios can forecast a negative equity premium. Particularly for the HF and PT models, negative forecasts occur frequently when estimating the predictive regression solely with simulated data (as in equation (12)). For simulated data from the LRR model, negative forecasts of the equity premium are not as frequent, as less weight is placed on the predictor variable: β_1 is small.

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. Hence, there are no substantial differences in terms of the risk that the

27
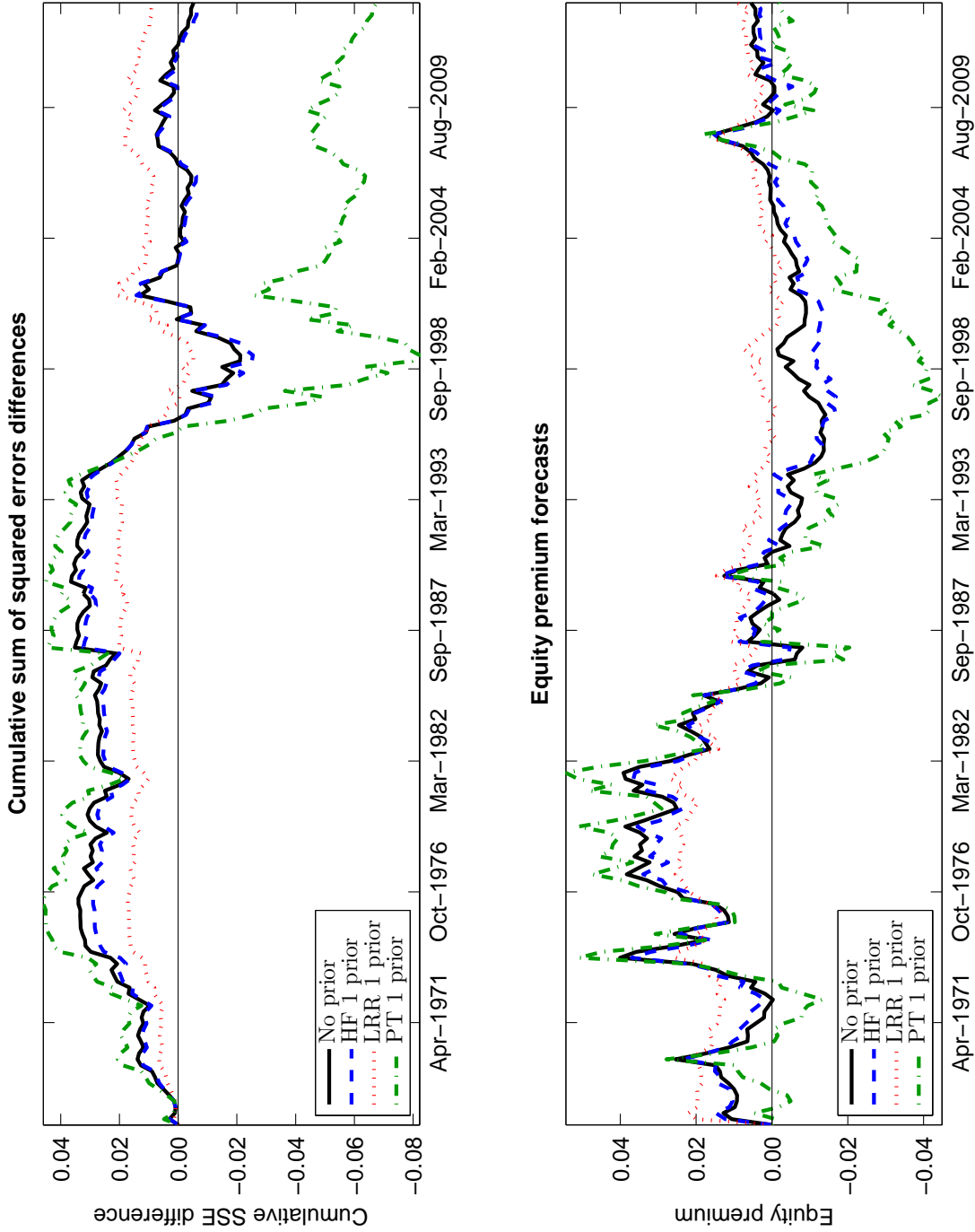
Figure 3: Forecasts of model-based priors

The top panel shows the differences in the cumulative sum of squared errors (SSE) between the log equity premium forecasts of the historical average model and the predictive regression 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 regression 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 regression for the OLS estimates and the model-based priors. For the model-based priors, the posterior point forecasts are given in equation (11).

28

predictive densities imply. However, 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 model-based priors. 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). As the predictive density of the log equity premium, the coefficient densities approximate Normal distributions. 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 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. This finding is related to Wachter and Warusawitharana (2009), who show that an investor who is skeptical about the predictive power of the dividend-price ratio and the yield spread performs better when forecasting the equity premium OOS.

My paper focuses on how model-based priors from different consumption-based asset pricing models perform without investigating the reasons behind empirical fluctuations in valuation ratios. However, several papers analyze drivers of the high equity prices during the 1990s. Some researchers propose that an increase in stock market participation and diversification is at least partially responsible for the higher equity prices (see, for example, Heaton and Lucas (1999)). Lettau, Ludvigson, and Wachter (2008) estimate through a regime-switching model that the low dividend-price ratios during 1990s are driven by a shift to substantially lower consumption volatility. This decrease in macroeconomic risk led to a lower expected equity premium. Lettau and Van Nieuwerburgh (2008) allow for shifts in the steady state of the economy and find a structural break in the 1990s for the dividend-price
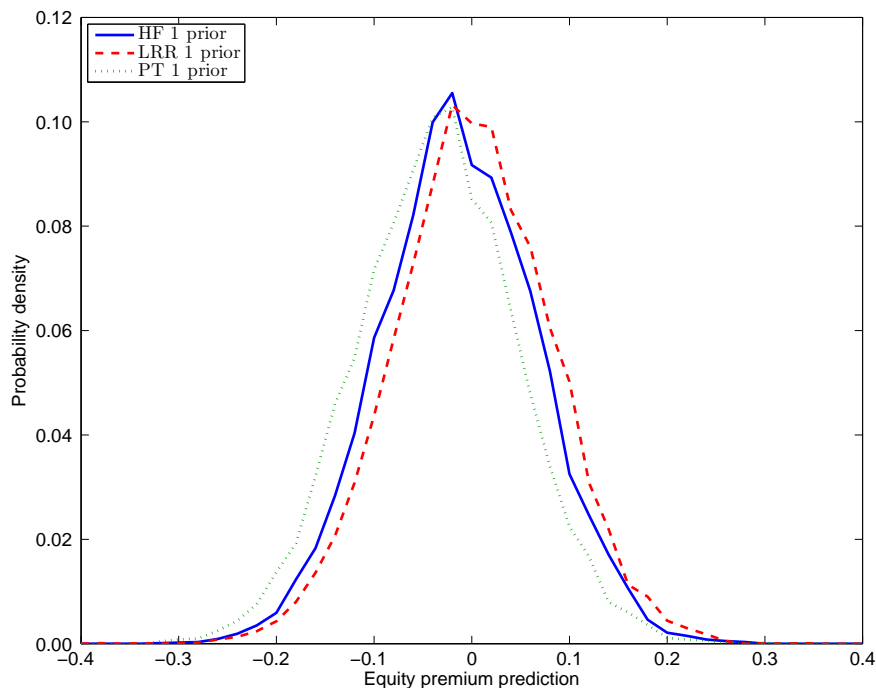
Figure 4: Predictive posterior density of the equity premium
This figure shows the simulated posterior density of the quarterly log equity premium prediction given in equation (10) for the third quarter in 1998 for three model-based priors: HF 1, LRR 1, and PT 1. The predictor is the log dividend-price ratio. Data from the first quarter in 1947 to the second quarter in 1998 are used to estimate the predictive regression. The densities are simulated with 10,000 draws.

ratio. The authors show that while this structural break can be detected in-sample, an investor could not have exploited it OOS. Based on a VAR framework, Campbell, Giglio, and Polk (2013) find that during the dot-com boom, the discount rates of investors were at a historically low level, but the boom preceding the financial crisis of 2007-2009 was caused by positive cash flow news.

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. However, investors are ultimately concerned about utility, and thus, we need to compute differences in utility gains when comparing the model-based priors. Further, comparing utility gains takes the investors' risk aversion into account.
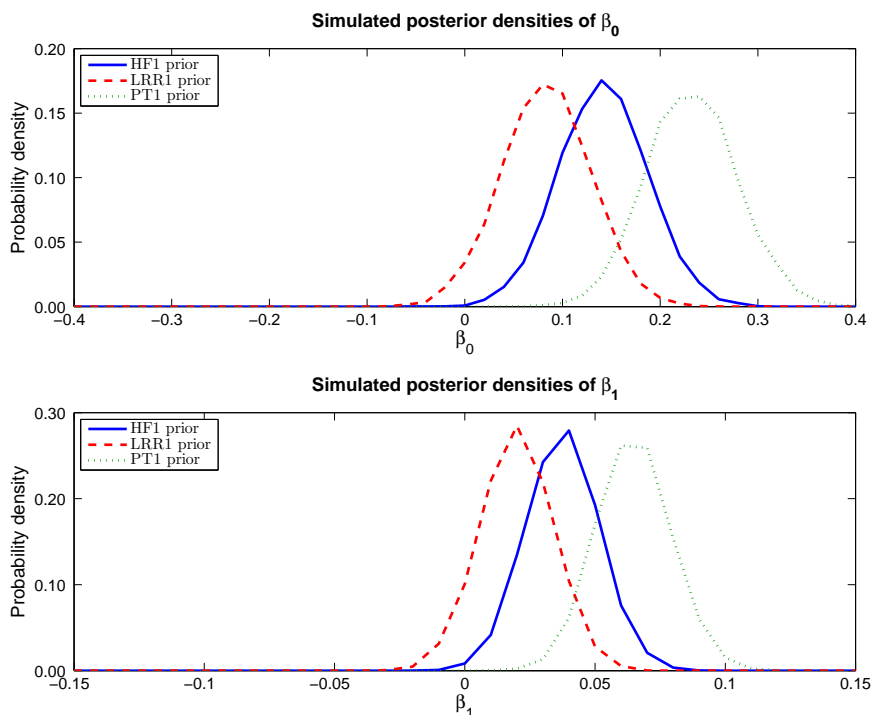
Figure 5: Posterior density of coefficients
This figure shows the simulated posterior density of the coefficients β_0 and β_1 given in equation (6) for the third quarter in 1998 for three model-based priors: HF 1, LRR 1, and PT 1. The predictor is the log dividend-price ratio. Quarterly data from the first quarter in 1947 to the second quarter in 1998 are used to estimate the predictive regression. The densities are simulated with 10,000 draws.

The Bayesian technique that I use to impose the economic 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

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

subject to

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

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

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

For an investor who imposes model-based priors on the predictive regression 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^*. The optimal risky asset portfolio share based on the model-based prior forecasts is denoted $\hat{\alpha}_{t,m}^*$. Based on $\hat{\alpha}_{t,m}^*$ and the realized equity premium, the realized wealth can be computed

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

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

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

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 each period to have access to the LRR model instead of 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 regression 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 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 favorable to the LRR prior than the R^2_{OOS} results reported in Table 2 and 3: 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, predic-

tors, 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 economically meaningful. 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 log dividend-price ratio is used as the predictor. For the bottom panel, the log dividend-yield is the predictor. Generally, the LRR prior leads to a more stable portfolio share of the risky asset, which is due to the low predictability implied by the LRR 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

34

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

Panel B: Quarterly returns

Sample start	OOS period	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
1926	1947-2014	Log DP ratio	LRR 1	0.205	0.172		0.001	0.671	0.862
		Log DY	LRR 1	0.134	0.096		0.051	0.051	0.129
1926	1947-1980	Log DP ratio	PT 1	0.015	0.049	0.177	0.196		0.044
		Log DY	HF 1		0.102	0.216	0.171	0.063	0.053
1926	1981-2014	Log DP ratio	LRR 2	0.474	0.485	0.027		1.538	1.803
		Log DY	LRR 2	0.187	0.221	0.011		0.030	0.281
1947	1968-2014	Log DP ratio	LRR 1	0.702	0.592		0.110	3.170	3.791
		Log DY	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	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
1926	1947-2014	Log DP ratio	LRR 1	0.138	0.097		0.027	0.619	0.771
		Log DY	LRR 1	0.153	0.166		0.159	0.074	0.101
1926	1947-1980	Log DP ratio	PT 1	0.104	0.051	0.084	0.212	0.015	
		Log DY	PT 2	0.016	0.121	0.070	0.170		0.034
1926	1981-2014	Log DP ratio	LRR 2	0.297	0.188	0.159		1.374	1.382
		Log DY	LRR 2	0.150	0.410	0.075		0.197	0.374
1947	1968-2014	Log DP ratio	LRR 1	0.477	0.136	0.044	0.072	4.443	4.459
		Log DY	PT 2	0.616	0.419		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	

predictive regression coefficients reported in Panel A of Table 1. The model-implied parameters of PT 2 for the dividend yield are smaller than for the dividend-price ratio, which makes the investor more hesitant to believe that the low dividend yield will predict an immediate downturn of the stock market.

5   Robustness

For the benchmark 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. I follow the calibration methodology proposed by the respective authors: some parameters are set equal to their empirical counterparts and others are chosen such that the model simulated moments match the empirical moments, as, for example, the mean and standard deviation of the dividend-price ratio or the equity premium. 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
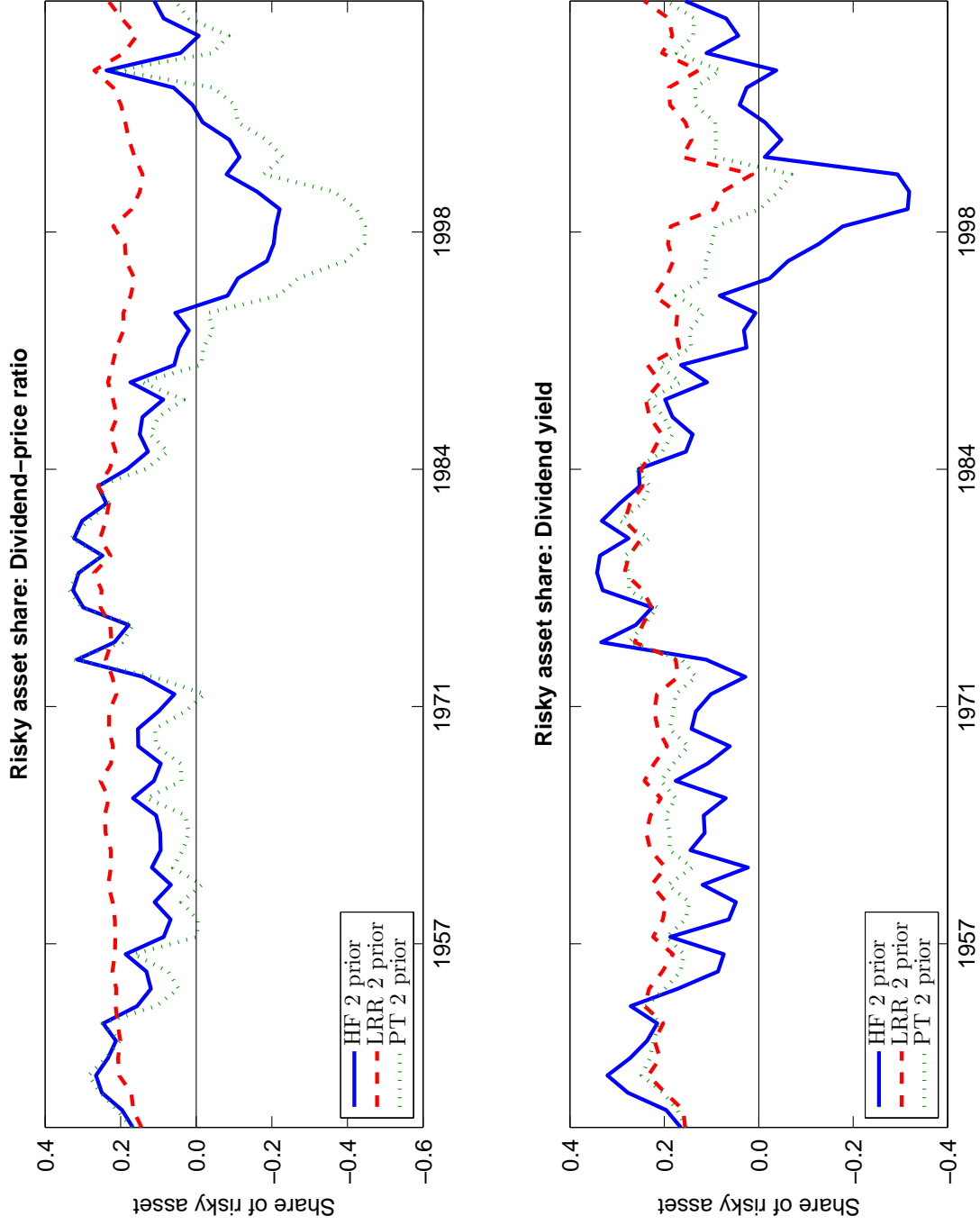
Figure 6: Portfolio share of risky asset
This figure shows the risky asset share of the portfolio given in equation (16) for mean-variance investors with γ equal to 5 who rely on the HF 2, LRR 2, and PT 2 priors, respectively. In the top (bottom) panel, the return on the risky asset is forecast with the log dividend-price ratio (log dividend yield). The returns are at an annual frequency, and the OOS period is from 1947 to 2014.

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 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 observations leads to more power compared to the annual returns.

For the benchmark analysis, 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. However, the results and the conclusions drawn in this paper are robust to tightening or loosening the model-based priors.

Table 6 reports the results for the total sample, that is, the 1947-2014 OOS period. Relative to the benchmark, the model-based priors are tightened and loosened by a factor of 2 and 4. Tightening the priors further does not alter the conclusion, and loosening the priors by more than factor of 4 leads to forecast results that are not substantially different from the OLS estimates. Tightening (loosening) the prior by a factor of 4 results in $\lambda = 0.25$ ($\lambda = 4$) and $\underline{v} = 0.4$ ($\underline{v} = 0.025$). The LRR model-based priors yield the most accurate forecasts across the range of λ and $\underline{v}$ values. For the dividend-price ratio, a LRR model-based prior is the best performing prior for all the hyperparameter values and return frequencies, with the exception of the monthly return frequency when $\lambda = 4$ and $\underline{v} = 0.4$. For the dividend-yield, the LRR model-based priors outperform the other priors in half of the cases. As in Table 3, the differences in OOS forecast errors are statistically significant based on the Diebold-Mariano test (see Diebold and Mariano (1995)) in several cases at a monthly and quarterly frequency.

The sensitivity to the prior tightness of the economic performance of a mean-variance investor are shown in Table 7, which is stuctured similarly to Table 4 and reports the CER given in equation (19) for each return predictor, prior tightness, and return frequency, for the total sample OOS period, that is, from 1947 to 2014. An

investor who forecasts the equity premium with an LRR model-based prior achieves the highest utility in 23 out of the 24 cases. Panel D shows the average CER for each model-based prior pair. Compared to an investor who times the market based on forecasts from HF priors, an investor using the LRR priors generates a CER of close to 20 basis point a year. The average CER of the LRR priors compared to the PT priors is around 50 basis points per year. These results confirm that the strong forecasting performance of an investor who derives her prior beliefs about the equity premium predictability from the LRR model are robust to changes in the prior tightness.

6   Conclusion

Different theories have been proposed to resolve the equity premium puzzle of Mehra and Prescott (1985). Three prominent consumption-based asset pricing models 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 compare these asset pricing models based on whether they can profitably guide the investment decisions of investors who try to time the stock market. I propose a simple Bayesian framework in which investors reduce the uncertainty about predictive regression parameters by imposing economic constraints derived from the three asset pricing models. The predictor of the single-variable predictive regression is a valuation ratio — that is, the log dividend-price ratio or the log 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 weak implied predictive power of valuation ratios for the equity premium. The weak predictive power is caused by the LRR

Table 5: Model-based priors from asset pricing models calibrated with 1926-1967 data

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 log dividend-price ratio and the log 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**

model's time-variation in the dividend growth forecasts, the long-run risk component. 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 the 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 over the total sample period, that is, from 1926 to 2014, 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. These differences in forecast accuracy are not only shown by 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 forecast performances of these priors differ, this paper makes a novel contribution to the equity premium prediction literature. This paper also adds to our understanding of consumption-based asset pricing models. The paper shows that over the 1926-2014 sample, an investor whose beliefs had been rooted in an asset pricing model that implies weak equity premium predictability would have outperformed investors who relied on priors from models in which the equity premium is strongly predictable.

Table 6: Hyperparameter sensitivity of the 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 prior tightness is varied, and the respective values for the hyperparameters λ and $\underline{v}$ are shown in the first two columns. The predictors are the log dividend-price ratio and the log dividend yield. For each frequency, predictor, and prior tightness, 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 in forecast errors 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 fourth column.

Panel A: Annual returns

λ	$\underline{v}$	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
					Sample start: 1926 / OOS period: 1947-2014				
0.25	0.4	Log DP ratio	LRR 2	6.538	5.067	1.819		12.914	24.394
0.25	0.4	Log DY	PT 2	7.871	5.621	2.849	2.590	1.493	
0.5	0.2	Log DP ratio	LRR 2	4.141	4.460	0.403		9.955	18.746
0.5	0.2	Log DY	PT 2	7.181	6.049	2.033	1.838	1.120	
2	0.05	Log DP ratio	LRR 1	1.171	1.095		0.561	1.780	2.815
2	0.05	Log DY	LRR 1	5.688*	4.559		0.053	1.420	1.244
4	0.025	Log DP ratio	LRR 1	0.631	0.268		0.618**	0.319	0.742
4	0.025	Log DY	LRR 2	3.605**	2.878**	0.731**		0.501	1.226**

Table 6: Hyperparameter sensitivity of the forecast performance (continued)

Panel B: Quarterly returns

λ	v	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
						Sample start: 1926 / OOS period: 1947-2014			
0.25	0.4	Log DP ratio	LRR 2	2.338	2.591	0.374		30.207**	43.069**
0.25	0.4	Log DY	PT 1	0.263	0.080	1.387	1.028		1.258
0.5	0.2	Log DP ratio	LRR 1	1.888	1.778		0.009	16.058**	23.426**
0.5	0.2	Log DY	PT 1	0.338	0.198	0.906	0.612		1.030
2	0.05	Log DP ratio	LRR 2	0.538	0.313	0.085		1.526*	1.919*
2	0.05	Log DY	HF 1		0.068	0.248	0.458*	0.119	0.159
4	0.025	Log DP ratio	LRR 2	0.864**	0.555**	0.330*		0.708**	0.616**
4	0.025	Log DY	HF 1		0.268	0.343	0.012	0.403*	0.338

Panel C: Monthly returns

λ	v	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
						Sample start: 1926 / OOS period: 1947-2014			
0.25	0.4	Log DP ratio	LRR 2	0.845	0.868	0.124		31.993**	37.229**
0.25	0.4	Log DY	LRR 2	1.184	1.163	0.384		0.161	0.856
0.5	0.2	Log DP ratio	LRR 2	0.499	0.484	0.237*		9.283**	10.978**
0.5	0.2	Log DY	LRR 1	0.360	0.793*		0.171	0.259	0.226
2	0.05	Log DP ratio	LRR 2	0.047	0.212*	0.145		0.205	0.215
2	0.05	Log DY	LRR 2	0.106	0.305**	0.355**		0.305**	0.218
4	0.025	Log DP ratio	HF 2	0.037		0.122	0.083	0.054	0.238*
4	0.025	Log DY	LRR 1	0.390**	0.183		0.295**	0.303**	0.167

Table 7: Hyperparameter sensitivity of the economic performance

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 that result from imposing the best-performing prior on the single-variable predictive regression, given in equation (1), instead of one of the remaining five priors. The best-performing prior is reported in the "best prior" column. The prior tightness is varied, and the respective values for the hyperparameters λ and $\underline{v}$ are shown in the first two columns. The predictors are the log dividend-price ratio and the log dividend yield. Panel D reports the average CER across all frequencies, hyperparameters, and predictors, for each prior pair.

Panel A: Annual returns

λ	$\underline{v}$	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
						Sample start: 1926 / OOS period: 1947-2014			
0.25	0.4	Log DP ratio	LRR 2	0.462	0.448	0.012		0.636	0.909
0.25	0.4	Log DY	LRR 2	0.357	0.303	0.004		0.021	0.118
0.5	0.2	Log DP ratio	LRR 2	0.471	0.436	0.035		0.649	0.935
0.5	0.2	Log DY	LRR 2	0.463	0.433	0.049		0.107	0.175
2	0.05	Log DP ratio	LRR 2	0.295	0.267	0.007		0.355	0.494
2	0.05	Log DY	LRR 1	0.335	0.329		0.021	0.072	0.103
4	0.025	Log DP ratio	LRR 1	0.164	0.164		0.010	0.214	0.275
4	0.025	Log DY	LRR 1	0.257	0.235		0.006	0.015	0.076

Panel B: Quarterly returns

λ	$\underline{v}$	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
						Sample start: 1926 / OOS period: 1947-2014			
0.25	0.4	Log DP ratio	LRR 2	0.356	0.369	0.075		1.719	2.056
0.25	0.4	Log DY	LRR 1	0.096	0.099		0.002	0.031	0.207
0.5	0.2	Log DP ratio	LRR 1	0.278	0.205		0.017	1.157	1.384
0.5	0.2	Log DY	LRR 1	0.114	0.091	0.026	0.026	0.016	0.151
2	0.05	Log DP ratio	LRR 1	0.151	0.098		0.004	0.369	0.394
2	0.05	Log DY	LRR 2	0.024	0.047			0.007	0.008
4	0.025	Log DP ratio	LRR 2	0.001	0.019	0.020		0.192	0.241
4	0.025	Log DY	LRR 2	0.052	0.122	0.072		0.110	0.100

45

Table 7: Hyperparameter sensitivity of the economic performance (continued)

Panel C: Monthly returns

λ	v	Predictor	Best prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
					Sample start: 1926 / OOS period: 1947-2014				
0.25	0.4	Log DP ratio	LRR 1	0.268	0.395			2.786	3.083
0.25	0.4	Log DY	LRR 2	0.281	0.193	0.009	0.089	0.123	0.223
0.5	0.2	Log DP ratio	LRR 2	0.047	0.140	0.082		1.534	1.520
0.5	0.2	Log DY	LRR 1	0.173	0.230		0.097	0.126	0.091
2	0.05	Log DP ratio	HF 2	0.002		0.011	0.010	0.349	0.408
2	0.05	Log DY	LRR 2	0.107	0.032	0.080		0.055	0.127
4	0.025	Log DP ratio	LRR 2	0.101	0.074	0.065		0.180	0.275
4	0.025	Log DY	LRR 1	0.150	0.120		0.143	0.038	0.183

Panel D: Average CER

Comparison prior	HF 1	HF 2	LRR 1	LRR 2	PT 1	PT 2
HF 1		0.007	0.186	0.191	-0.242	-0.353
HF 2	-0.007		0.179	0.184	-0.248	-0.360
LRR 1	-0.186	-0.179		0.005	-0.428	-0.539
LRR 2	-0.191	-0.184	-0.005		-0.433	-0.544
PT 1	0.242	0.248	0.428	0.433		-0.111
PT 2	0.353	0.360	0.539	0.544	0.111	

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

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

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

$$\eta_t\equiv-\frac{C_t u_{cc}(C_t,X_t)}{u_c(C_t,X_t)}=\frac{\gamma}{S_t}.\tag{A.2}$$

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

$$s_{t+1}=(1-\phi)\bar{s}+\phi s_t+\lambda(s_t)(c_{t+1}-c_t-g),\tag{A.3}$$

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

$$\lambda(s_t)=\begin{cases}\frac{1}{\bar{S}}\sqrt{1-2(s_t-\bar{s})}-1,&s_t\leq s_{max}\\[2mm]0,&s_t>s_{max},\end{cases}\tag{A.4}$$

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

$$\Delta c_{t+1} = g + v_{t+1}, \text{ where } v_{t+1} \stackrel{i.i.d.}{\sim} N(0, \sigma_v^2). \tag{A.5}$$

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

$$\frac{P_t}{C_t}(s_t) = E_t \left[M_{t+1} \frac{C_{t+1}}{C_t} \left[1 + \frac{P_{t+1}}{C_{t+1}}(s_{t+1}) \right] \right]. \tag{A.6}$$

The underlying assumption is that dividend growth is perfectly correlated with consumption growth in equation (A.5). Above, I denote this specification the HF 1 model.[10] The intertemporal marginal rate of substitution (IMRS) M_{t+1} takes the form

$$M_{t+1} \equiv \delta \left(\frac{S_{t+1}}{S_t} \frac{C_{t+1}}{C_t} \right)^{-\gamma}. \tag{A.7}$$

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

$$r_t^f = -ln(\delta) + \gamma g - \gamma(1-\phi)(s_t - \bar{s}) - \frac{\gamma^2 \sigma^2}{2}[1 + \lambda(s_t)]^2. \tag{A.8}$$

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 (A.2) shows how an increase in s_t lowers the the local curvature of the utility function and makes

---

[10]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 benchmark analysis, 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, I follow the methodology of Campbell and Cochrane (1999). Consumption data are real per capita consumption of non-durables 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.

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

$$\Delta c_{t+1} = g_c + \sigma_c \epsilon_{c,t+1}, \text{ where } \epsilon_{c,t+1} \overset{i.i.d.}{\sim} N(0,1) \tag{A.9}$$

and

$$\Delta d_{t+1} = g_d + \sigma_d \epsilon_{d,t+1}, \text{ where } \epsilon_{d,t+1} \overset{i.i.d.}{\sim} N(0,1), \tag{A.10}$$

with the correlation between $\epsilon_{c,t+1}$ and $\epsilon_{d,t+1}$ being denoted by ω.[11]

The agent's maximization problem is set up as

$$E\left[\sum_{t=0}^{\infty} \left(\rho^t \frac{C_t^{1-\gamma}}{1-\gamma} + b_0 \bar{C}_t^{-\gamma} \rho^{t+1} v(X_{t+1}, S_t, z_t)\right)\right]. \tag{A.11}$$

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

$$X_{t+1} \equiv S_t R_{t+1} - S_t R_{f,t}. \tag{A.12}$$

---

[11]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.

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 calibrated based on a sample with data from 1926 to 1967. The price-dividend ratio moments are annualized.

Panel A: Based on original parameter values						
Model	Freq.	P/D Mean	Log P/D Std. dev.	Log equity prem. Mean	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: Based on 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

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

$$z_{t+1} = \eta\left(z_t\frac{\bar{R}}{R_{t+1}}\right) + (1-\eta). \tag{A.13}$$

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:

$$f_t \equiv P_t/D_t = f(z_t). \tag{A.14}$$

The real stock returns are thus given as

$$R_{t+1} = \frac{1+f(z_{t+1})}{f(z_t)}e^{g_d+\sigma_d\epsilon_{d,t+1}}. \tag{A.15}$$

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

$$R_f = \delta^{-1}e^{\gamma g_c - \gamma^2\sigma_c^2/2}, \tag{A.16}$$

and a price-dividend ratio determined by the equation

$$
\begin{aligned}
1 =& \delta e^{g_d - \gamma g_c + \gamma^2 \sigma_c^2 (1-\omega^2)/2} E_t \left[\frac{1 + f(z_{t+1})}{f(z_t)} e^{(\sigma_d - \gamma \omega \sigma_c) \epsilon_{d,t+1}} \right] \\
& + b_0 \delta E_t \left[\hat{v} \left(\frac{1 + f(z_{t+1})}{f(z_t)} e^{g_d + \sigma_d \epsilon_{d,t+1}}, z_t \right) \right],
\end{aligned}
\tag{A.17}
$$

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

$$
\hat{v}(R_{t+1}, z_t) = \begin{cases} R_{t+1} - R_{f,t}, & R_{t+1} \geq z_t R_{f,t} \\ (z_t R_{f,t} - R_{f,t}) + \lambda(R_{t+1} - z_t R_{f,t}), & R_{t+1} < z_t R_{f,t} \end{cases}
\tag{A.18}
$$

and for $z_t > 1$ as

$$
\hat{v}(R_{t+1}, z_t) = \begin{cases} R_{t+1} - R_{f,t}, & R_{t+1} \geq R_{f,t} \\ \lambda(z_t)(R_{t+1} - R_{f,t}), & R_{t+1} < R_{f,t}, \end{cases}
\tag{A.19}
$$

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 (A.10) increases the return of the asset and leads to a lower z_t through equation (A.13). A lower z_t implies that the agent is less loss averse as shown in equations (A.18) and (A.19). 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 parameter values from Barberis et al. (2001) are reported in Table A.3. in the "original value" column. My recalibration 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 recalibration follows the methodology of Barberis et al. (2001). The PT 1 model is calibrated such that the average effective loss aversion of the model is 2.25.[12] 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.[13] 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

---

[12]This value is chosen by Barberis et al. (2001) based on experimental evidence.

[13]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.

recalibration of the parameter values with data from 1926 to 1967.

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

$$E_t \left[\delta^\theta G_{c,t+1}^{-\frac{\theta}{\psi}} R_{c,t+1}^{-(1-\theta)} R_{m,t+1} \right] = E_t \left[M_{t+1} R_{m,t+1} \right] = 1, \tag{A.20}$$

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

$$r_{c,t+1} = \kappa_0 + \kappa_1 z_{t+1} - z_t + g_{c,t+1}, \tag{A.21}$$

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.

Panel A: Based on original parameter values						
Model	Freq.	Price-dividend ratio		Log equity prem.		Log Sharpe
		Mean	Std. dev.	Mean	Std. dev.	ratio
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: Based on 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

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

$$r_{m,t+1} = \kappa_{0,m} + \kappa_{1,m}z_{m,t+1} - z_{m,t} + g_{d,t+1,} \tag{A.22}$$

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).[14]

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

$$
\begin{aligned}
x_{t+1} &= \rho x_t + \varphi_e \sigma_t e_{t+1} \\
g_{c,t+1} &= \mu_c + x_t + \sigma_t \eta_{t+1} \\
g_{d,t+1} &= \mu_d + \phi x_t + \varphi_d \sigma_t u_{t+1} \\
\sigma_{t+1}^2 &= \sigma^2 + v_1(\sigma_t^2 - \sigma^2) + \sigma_w w_{t+1},
\end{aligned}
\tag{A.23}
$$

with e_{t+1}, u_{t+1}, η_{t+1}, and w_{t+1} having i.i.d. standard Normal distributions.[15] 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

$$
\begin{aligned}
z_t &= A_0 + A_1 x_t + A_2 \sigma_t^2 \\
z_{m,t} &= A_{0,m} + A_{1,m} x_t + A_{2,m} \sigma_t^2.
\end{aligned}
\tag{A.24}
$$

The derivation of A and A_m can be found in Bansal and Yaron (2004) and Bansal et al. (2010 and 2012).

---

[14]Bansal et al. (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.

[15]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.

The model generates excess returns that are predictable by the price-dividend 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 negative correlation of price-dividend ratios and subsequent returns.

A.3.1 Calibration and simulation of model

The 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.[16] 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

---

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

of 1.80% and a standard deviation of 3.08% with an autocorrelation of 0.32. The 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. (2010 and 2012) 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: Based on 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: Based on 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 from 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., A. R. Gallant, and G. Tauchen. 2007. Rational pessimism, rational exuberance, and asset pricing models. *Review of Economic Studies* 74:1005-1033.

Bansal, R., D. Kiku, and A. Yaron. 2007. A note on the economics and statistics of predictability: A long run risks perspective. Unpublished working paper.

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

Bansal, R., D. Kiku, and A. Yaron. 2012. An empirical evaluation of the long-run risks model for asset prices. *Critical Finance Review* 1:183-221.

Bansal, R., 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.

Barro, R. J. 2006. Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics* 121:823-866.

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

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

Cochrane, J. H. 2008. The dog that did not bark: A defense of return predictability. *Review of Financial Studies* 21:1534-1575.

Constantinides, G. M., and A. Ghosh. 2012. Asset pricing tests with long-run risks in consumption growth. *Review of Asset Pricing Studies* 1:96-136.

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.

Ferson, W., S. Nallareddy, and B. Xie. 2013. The "out-of-sample" performance of long run risk models. *Journal of Financial Economics* 107:537-556.

Heaton, J., and D. Lucas. 1999. Stock prices and fundamentals, in O. J. Blanchard, and S. Fischer (eds), NBER Macroeconomics Annual: 1999. Cambridge, MA: MIT Press.

Ingram, B. F., and C. H. Whiteman. 1994. Supplanting the 'Minnesota' prior: Forecasting macroeconomic time series using real business cycle model priors. *Journal of Monetary Economics* 34:497-510.

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

Lettau, M., and S. C. Ludvigson. 2005. Expected returns and expected dividend growth. *Journal of Financial Economics* 76:583-626.

Lettau, M., S. C. Ludvigson, and J. A. Wachter. 2008. The declining equity premium: What role does macroeconomic risk play. *Review of Financial Studies* 21:1653-1687.

Lettau, M., and S. Van Nieuwerburgh. 2008. Reconciling the return predictability evidence. *Review of Financial Studies* 21:1607-1652.

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.

Ludvigson, S. C. 2013. Advances in consumption-based asset pricing: Empirical tests, in G. M. Constantinides, M. Harris, and R. M. Stulz (eds), Handbook of the Economics of Finance: 2013. Netherlands, Amsterdam: Elsevier B.V.

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.

Penasse, J. 2016. Return predictability: Learning from the cross-section. Unpublished working paper.

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.