A Likelihood-Based Comparison of Macro Asset Pricing Models

Andrew Y. Chen
Federal Reserve Board

Rebecca Wasyk
Federal Reserve Board

Fabian Winkler
Federal Reserve Board

December 2016*

Abstract

We develop a model in which asset prices depend on long run growth, long run volatility, habit, and a persistent residual. We estimate the model using Bayesian methods which account for the entire likelihood of the data on consumption growth, dividend growth, and the price-dividend ratio. The residual is dominant, accounting for more than 80% of the variance of the price-dividend ratio across a variety of priors and specifications. Moreover, the filtered residual tracks most of the recognizable features of the U.S. stock market, such as the late 1990's boom and bust. Long run volatility, long run growth, and habit contribute in crises, but overall have a low correlation with the price-dividend ratio between 1929 and 2014. These results show that while long run risks and habit play a non negligible role, something else is driving the bulk of stock market fluctuations. We categorize and discuss theories which are consistent with our results.

*We thank Francesco Bianchi, Ed Herbst, Francisco Palomino, and Missaka Warusawitharana for advice and comments. Our emails are andrew.y.chen, rebecca.d.wasyk, and fabian.winkler, all @frb.gov. The views expressed herein are those of the authors and do not necessarily reflect the position of the Board of Governors of the Federal Reserve or the Federal Reserve System.

1. Introduction

Models of asset prices have come a long way since Mehra and Prescott (1985). We now have several explanations of aggregate stock market fluctuations. Arguably the most prominent are habit formation, long run risks, and rare disasters. But there are more, including limited participation, intermediary-based models, and learning.[1]

In this paper, we evaluate the relative importance of these explanations. Our evaluation focuses on a model in which the price dividend ratio depends on four sources of market volatility: habit, long run growth, long run volatility, and a persistent residual. Habit and long run risks are related to consumption and dividends in the usual way (Campbell and Cochrane (1999), Bansal, Kiku, and Yaron (2012a)). The residual accounts for all other sources of stock price movements: disaster probability movements, shifts in beliefs about returns, etc.

We estimate the model using Bayesian methods and data on consumption growth, dividend growth, and the price dividend ratio. The estimated parameters and latent states allow us to decompose the variance of the price dividend ratio into contributions from each source of market volatility.

We find that the residual is the most important source of market volatility, accounting for the vast majority of the variance of the price dividend ratio. The residual accounts for more than 80% of the variance across a variety of priors and specifications. Moreover, the filtered residual tracks most of the recognizable features of the U.S. stock market's history, such as the booms and busts of the 1960s and late 1990s. Long run volatility, long run growth, and habit have large effects in the Great Depression and 2008 Financial Crisis, but overall they display a low correlation with asset prices between 1929 and 2014. These results show that, while long run risks and habit have a non-negligible effect, something else is the key driver of market volatility.

Importantly, the dominance of the residual is independent of our choice of target moments, as our Bayesian estimation accounts for the entire likelihood

[1] Here we list just a couple references for each literature. For habit formation see Constantinides (1990) and Campbell and Cochrane (1999). For long run growth and volatility risks see Bansal and Yaron (2004) and Bansal, Kiku, and Yaron (2012a). For rare disasters see Rietz (1988), Barro (2006), Gabaix (2012), Wachter (2013). For limited participation, Mankiw and Zeldes (1991) and Guvenen (2009). For intermediary-based models see He and Krishnamurthy (2013) and Brunnermeier and Sannikov (2014). For learning models see Adam, Marcet, and Beutel (2015) and Adam, Marcet, and Nicolini (2016).

of consumption, dividends, and the price dividend ratio. This methodology cuts through the problem of weighing disparate pieces of evidence that arise from the moment matching literature. How much should we discount long run risks since it predicts counterfactually strong dividend predictability? How important is the fact that habit implies a counterfactual link between asset prices and lagged consumption growth? What do we make of models that are not evaluated against these particular moments? By accounting for all moments, our Bayesian approach and variance decomposition provides a succinct answer to these questions.

Fluctuations in the residual are a kind of excess market volatility: The residual moves closely with asset prices, but is unconnected to real economic growth and real economic volatility. This description matches several theories in the literature which fit into two broad categories: tractable models with hard-to-observe shocks to risk (such as variable disaster risk) and more complex models which directly link expected returns to observables other than aggregate consumption (such as intermediary-based models). We discuss these theories and avenues for future research, but we cannot distinguish among these theories in this paper.

Models with hard-to-observe risk lead to several observationally equivalent structural models. This equivalence motivates us to focus on a semi-structural model— that is, we simply assume that the log price-dividend ratio is linear in the four state variables rather than derive the coefficients from assumptions about preferences and market structure. But there are additional considerations which compel us to deviate from the standard approach of looking for equilibrium among optimizing agents.

The semi-structural model lets the estimator speak freely. It ensures that the estimation results are due to properties of the data rather than functional form restrictions imposed by our choice of model economy. Similarly, the reduced form is much less costly for the reader to work through. This is especially important as our model includes several sources of risk. Lastly, an agnostic model seems appropriate considering the vast disagreement in the literature about the economic structure underlying stock prices (see, for example, Gabaix (2012) and Cochrane (2016) for some contrasting perspectives).

Our estimator uses the entire likelihood, but the low correlation between price/dividends and real growth or real volatility drives the results. To demonstrate this, we replicate our main finding with a simplified version of our estima-

tor. Specifically, we use the Bayesian estimator to construct state paths using only consumption and dividend data. We then use OLS to regress the price-dividend ratio on the states. The variance decomposition from these procedure uses only correlations between price/dividends and the estimated states, and this information leads to OLS estimator to conclude that the residual explains the vast majority of market volatility.

Though our approach does not require choosing moment targets, we do need to take a stand on a few modeling and econometric issues. In every case we make choices that favor simplicity for its various scientific virtues. Simple formulations are easier to dissect, communicate, replicate, and extend. Indeed, the lack of replicability of economic research has been recently highlighted by Chang and Li (2015).

Simplicity has costs, however. Specifically, our desire for simplicity requires that we offer two caveats regarding our conclusion that residual is dominant.

The first caveat is that our approach may favor the residual because of our use of an annual model. We argue that an annual model is ideal, not only because it is the simplest approach, but also because the striking seasonality in sub-annual data suggests that risk is best understood at an annual frequency. The fact that we recover similar parameters to the cash flow only estimates of Schorfheide, Song, and Yaron (2016)'s mixed frequency model is reassuring. Nevertheless, some studies suggest that a monthly model is critical for matching asset prices with both long run risks (Bansal, Kiku, and Yaron (2012a)) and habit (Campbell and Cochrane (2000)), and adding this layer of complexity may decrease the role of the residual.

The second caveat is that our approach may favor the residual because we use relatively simple formalizations of habit and long run risks. More subtle formalizations, such as the use of several volatility processes for long run risk (Schorfheide, Song, and Yaron (2016)) or the incorporation of additional shocks to habit (Bekaert, Engstrom, and Xing (2009)) may also decrease the residual contribution.

Relation to the Literature Our paper fits into the growing literature that compares the empirical performance of macro asset pricing models. Bansal, Gallant, and Tauchen (2007), Beeler and Campbell (2012), Bansal, Kiku, and Yaron (2012a), and Barro and Jin (2016) use moment matching methods to compare the

empirical performance of habit, long run risks, and rare disasters. The picture that emerges from this approach is somewhat muddled, as the preferred model depends on which moments one considers important. For example, habit is preferred if one places a large weight on accounting for the Shiller (1981) volatility puzzle. On the other hand, long run risks are preferred if one is particularly concerned with matching time-varying consumption volatility.

Aldrich and Gallant (2011) attempt to clarify the picture by comparing habit, long run risks, and prospect theory in a likelihood-based Bayesian framework. Our results echo theirs: long run risks is critical for addressing the volatile 1930s, but less important for other time periods. We differ from Aldrich and Gallant, however, by allowing for a residual / missing risk factors to drive asset prices.

The importance of including a residual is seen in more recent papers which find that neither long run risks nor habit formation is capable of matching some interesting stylized facts. Van Binsbergen, Brandt, and Koijen (2012) examine dividend strips and equity options, Dew-Becker et al. (2015) examine variance swaps, and Muir (2015) examines international wars and financial crises. We complement these papers by showing that one does not need to introduce derivative markets nor international data to empirically challenge long run risks and habit formation. The time series of U.S. consumption and stock prices is sufficient for showing that something outside these two kinds of risks is critical for understanding stock market volatility.

In terms of econometrics, our paper owes a large intellectual debt to Schorfheide, Song, and Yaron (2016). They also use a particle filter and Bayesian MCMC methods (Herbst and Schorfheide (2014)) to estimate a model with long run risks. We follow their approach closely, adopting their elegant state space system and filtering procedure. Our results complement theirs in that we also find strong evidence of long run risks in consumption and dividends, and indeed, similar posterior estimates, using an annual model and annual data. We deviate from Schorfheide, Song, and Yaron (2016), however, by allowing for a persistent residual in the price dividend ratio. Thus, our estimation is a much more stringent test of the long run risks model. We also assume a simpler version of long run risk, which highlights the importance of the multiple volatility states in their model.

2. Model, Estimation, and Main Results

This section begins with the model and ends with the main results: decompositions of the price-dividend ratio (Section 2.6). Along the way, we cover the model frequency, data, and parameter estimates.

2.1. Semi-Structural Model with Multiple Sources of Risk

Our key variable of interest is the log price-dividend ratio pd_t. pd_t is linear in four state variables

$$pd_t = \mu_{pd} + A_x x_t + A_V \tilde{\sigma}_t^2 + A_s \tilde{s}_t + A_e e_t. \tag{1}$$

where x_t and $\tilde{\sigma}_t^2$ correspond to the long run growth and volatility risk, $\tilde{s}_t$ is the habit (*s*urplus consumption), and e_t is a residual.

Residuals are not usually called "state variables," but our residual is persistent, plays an important role in accounting for the data, and can be interpreted to through several economic models (Section 5). The tildes over $\tilde{\sigma}_t$ and $\tilde{s}_t$ indicate that they're demeaned ($\tilde{\sigma}_t^2 = \sigma_t^2 - \mathbb{E}(\sigma_t^2)$), which implies that μ_{pd} is the mean log price-dividend ratio.

Our goal is to estimate the coefficients A_x, A_V, A_s, A_e and filter out the historical paths of the states $x_t, \tilde{\sigma}_t^2, \tilde{s}_t$ and e_t. The coefficients and state paths provide a simple description of the importance of each source of market volatility.

We do not derive (1) from an equilibrium model in order to let the estimator speak freely. However, there are several ways to derive (1). For example, one can extend Yang (2015)'s Epstein-Zin habit model to include time-varying disaster probability.

Most of the states are identified by their linkages with the other observables: consumption and dividends. The long run risk states $x_t, \tilde{\sigma}_t^2$ are identified by their relationship with consumption and dividend growth

$$\begin{aligned} \Delta c_t &= \mu_c + x_{t-1} + \sigma_{t-1}\eta_{c,t} \\ \Delta d_t &= \mu_d + \phi_x x_{t-1} + \phi_{\eta c}\sigma_{t-1}\eta_{c,t} + \varphi_d \sigma_{t-1}\eta_{d,t} \\ \eta_{c,t}, \eta_{d,t} &\sim N(0,1) \text{ i.i.d.}, \end{aligned} \tag{2}$$

where long run growth x_t evolves according to the standard heteroskedastic AR1

$$\begin{aligned} x_t &= \rho_x x_{t-1} + \varphi_x \sqrt{1-\rho_x^2}\sigma_{t-1}\eta_{x,t} \\ \eta_{x,t} &\sim N(0,1) \quad \text{i.i.d.}, \end{aligned} \tag{3}$$

and long run volatility σ_t evolves according to

$$\begin{aligned} h_t &= \rho_h h_{t-1} + \sigma_h \sqrt{1-\rho_h^2}\eta_{h,t} \\ \sigma_t &= \bar{\sigma}\exp(h_t). \\ \eta_{h,t} &\sim N(0,1) \quad \text{i.i.d.} \end{aligned} \tag{4}$$

These specifications borrow some technical fixes from Schorfheide, Song, and Yaron (2016), but otherwise the above consumption and dividends are identical to that in Bansal, Kiku, and Yaron (2012a).[2] This specification ensures that volatility is always positive, and also help for specifying a good prior.

Importantly, our specification does not include the multiple volatility processes of Schorfheide et al. This choice keeps things simple and closer to the bulk of the long run risk literature, but is restrictive in some ways. Specifically, our simpler specification assumes that the impact of volatility on the price-dividend ratio can be identified with realized consumption growth.

The demeaned habit state $\tilde{s}_t$ is also identified by consumption growth. This link comes from the transition equation

$$\begin{aligned} \tilde{s}_t &= \rho_s \tilde{s}_{t-1} + \lambda(\tilde{s}_{t-1})(\Delta c_t - \mathbb{E}_{t-1}\Delta c_t) \\ \lambda(\tilde{s}_{t-1}) &= \begin{cases} \frac{1}{\bar{S}}\sqrt{1-2\tilde{s}_{t-1}} - 1, & \tilde{s}_t \leq 1/2(1-\bar{S}^2) \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{5}$$

which is equivalent to the formulation in Campbell and Cochrane (1999). Section 4, shows that having surplus consumption respond to consumption growth itself (rather than consumption innovations) does not affect the main results.

The residual, however, is not identified by either consumption or dividends.

[2] To see the mapping, note that $\sigma_t^2 - \bar{\sigma}^2 \approx \rho_h\left(\sigma_{t-1}^2 - \bar{\sigma}^2\right) + 2\sigma_h\sqrt{1-\rho_h^2}\eta_{h,t}$.

It is simply an AR1

$$e_t = \rho_e e_{t-1} + \sigma_e \eta_{e,t} \tag{6}$$
$$\eta_{e,t} \sim N(0,1) \quad \text{i.i.d.}$$

and is thus identified by the price-dividend ratio. e_t captures everything drives market volatility that is *not* long run growth, long run volatility, or habit.

2.2. Model Frequency and Data

We assume the model frequency is annual, the same frequency as the data we use. This differs from the typical approach in the literature which tests monthly models against annual data moments.

We choose this approach for two reasons. The first is that monthly consumption and dividends exhibit stark seasonality which is entirely unaccounted for by models. The enormous end-of-quarter boosts to dividend growth and spikes in consumption at the end of the year suggest that risk is properly understood at an annual horizon. Indeed, if monthly risk is relevant to agents in the economy, why would we observe such stark seasonality in equilibrium?

Moreover, modeling this seasonality is not a simple task. Simple deterministic month or quarter fixed effects do a poor job, leading to the sophisticated Census Bureau's X-13ARIMA-SEATS seasonal adjustment approach. As discussed in Ferson and Harvey (1992), the Census Bureau adjustments are forward-looking: They boost the current month's observation if the future months are high. The resulting series is difficult to interpret in a model of consumption risk.

The second reason we use an annual model is that the robustness of asset pricing frameworks to changes in model frequency is an interesting question in itself. The annual frequency is particularly relevant, as annual data is far more accessible and uncontroversial. Indeed, the nondurable consumption at the monthly level is never directly observed, and instead is calculated by holding fixed shares observed every five years (Wilcox (1992)). Time-aggregating a monthly model to the annual horizon is possible but dramatically increases the complexity of model evaluation (Schorfheide, Song, and Yaron (2016)).

Thus, we estimate the model using annual consumption, dividend, and stock price data from the Bureau of Economic Analysis and the Center for Research on Security Prices. Consumption is real non-durable and services consumption.

Dividends and prices correspond to the CRSP index. The sample runs from 1929 to 2014.

2.3. Estimation Method

The model contains a significant number of unobserved state variables, so it's important to use an estimation approach which takes full advantage of the data available. To this end, we estimate the model using Bayesian MCMC methods. Such methods utilize the full likelihood of the data, while maintaining computational tractability. This approach also avoids the potentially contentious choice of moment conditions.

To evaluate the likelihood of our nonlinear model, we use a particle filter (Herbst and Schorfheide (2014)). We also take advantage of the conditionally Gaussian nature of the model to adapt the filter, using the approach of Schorfheide, Song, and Yaron (2016). To estimate the model parameters, we embed the filter in a standard random-walk Metropolis-Hastings algorithm. Details of the particle filter and Metropolis-Hastings algorithms can be found in the Appendix.

We fix some parameters outside of the estimation that are uninteresting or difficult to identify. The (uninteresting, for our purposes) means of all observables μ_{pd}, μ_c, μ_d are fixed to be their sample means.

$\bar{S}$ and A_s are difficult to identify separately as they both control the volatility of the habit contribution to the price-dividend ratio. Thus we chose $\bar{S} = 0.06$, close to the Campbell and Cochrane (1999) value. In Section 4 we estimate this parameter and find that it is poorly identified but does not affect the main results. Experiments with assuming alternative values of $\bar{S}$ also did not have a significant impact on the main results.

Similarly, σ_e and A_e both control the volatility of the residual contribution. Thus, we set $\sigma_e = 1$.

2.4. Prior Parameters

Priors are chosen to be as diffuse as possible, while maintaining the economic interpretation of the model. Overall, our consumption and dividend priors are similar to those in Schorfheide, Song, and Yaron (2016). However, the main results are not at all sensitive to the choice of prior, as we show in Section 4.

Prior distributions are independent and uniform for simplicity. Uniform priors are also useful because they imply that the posterior is simply a plot of the likelihood function.[3]

The left half of Table 1 shows the range spanned by the priors. The priors are diffuse. For example, our prior persistence parameters are uniform between 0 and 1, and consumption volatility is between 0.1% and 4.0% annually. The upper bound on the relative volatility of long run growth is 0.20, an order of magnitude larger than Bansal, Kiku, and Yaron (2012a)'s choice of 0.038. Nevertheless, we will see that the data suggest that these priors need to be revised significantly, and that there is a large predictable component.

We chose priors on the price dividend ratio coefficients that allow for the possibility that each individual state variable can account for 100% of the variance of the price dividend ratio at the standard parameters in the literature. Explicitly, for the long run growth coefficient, we choose the upper bound on A_x to solve

$$\text{Var}(\Delta pd_t) \approx A_x \varphi_x \bar{\sigma} \tag{7}$$

where $\varphi_x = 0.038$ and $\bar{\sigma} = 0.0072 \times \sqrt{12}$ as in Bansal, Kiku, and Yaron (2012b), and $\text{Var}(\Delta pd_t) = 0.23$ in our data sample. We use the analogous expressions to equation (7) for the other state variables. The signs of the price dividend coefficients are also restricted to be intuitive. That is, we restrict the coefficients on long run growth and surplus consumption to be positive, and we restrict the coefficients on long run volatility and to be negative.

2.5. Posterior Parameter Estimates

The right half of Table 1 shows the posterior estimates. The posteriors on simple consumption and dividend parameters are standard. The steady state consumption volatility $\bar{\sigma}$ is about 1% per year, and dividends are roughly 6 times as volatile as consumption.

The estimator finds evidence of significant long run risks in real economic

[3] This is just the result of Bayes Rule and the constancy of uniform priors

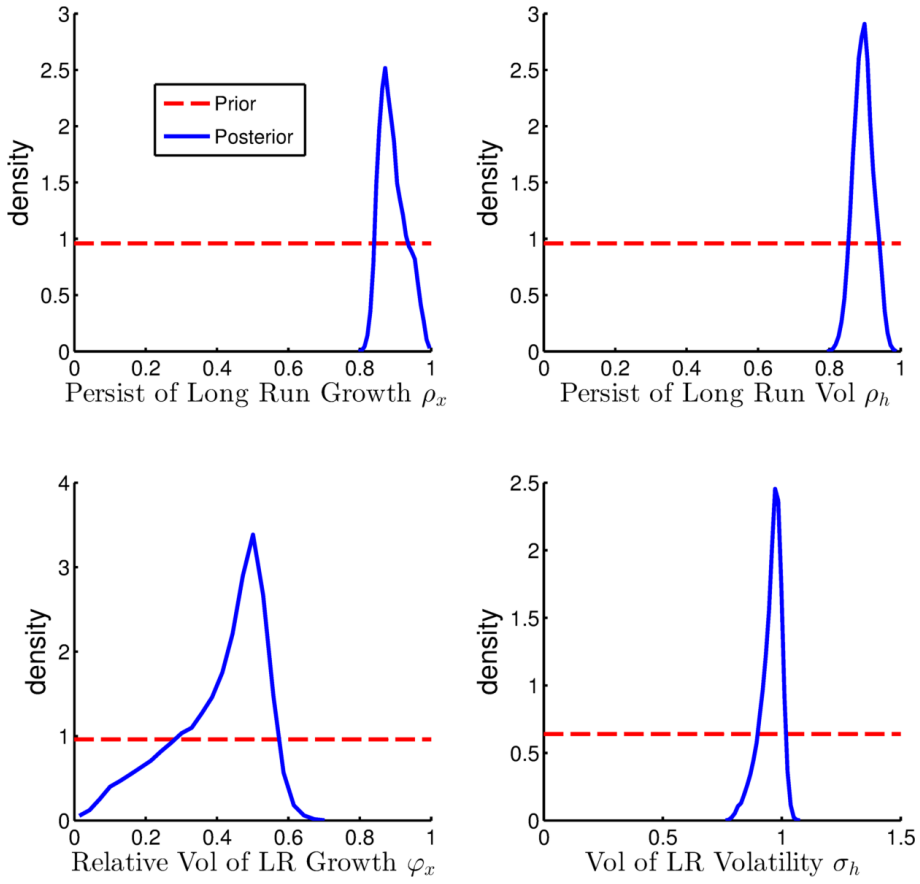
$$\begin{aligned} p(\text{parameters}|\text{data}) &= [\text{Constants}]\, p(\text{data}|\text{parameters})\, p(\text{parameters}) \\ &= [\text{Constants}]\, p(\text{data}|\text{parameters}). \end{aligned}$$

Table 1: Parameter Estimates

Model (equations (1) - (6)) and parameters are annual. All priors distributions are uniform. Posteriors are computed using annual consumption, dividend, and stock prices from 1929-2014, particle filter, and Metropolis Hastings. $\sigma_e = 1$ and $\bar{S} = 0.06$ are chosen outside of the estimation. Plots of the distributions can be found in the Appendix.

Parameter		Prior		Posterior			
		0%	100%	Mean	5%	50%	95%
Simple Consumption and Dividends							
Consumption Vol	$\bar{\sigma}$	0.001	0.04	0.0098	0.0047	0.0094	0.0162
Div Loading on Cons Shock	$\phi_{\eta c}$	0	10	1.15	0.396	1.2	1.75
Relative Vol of Dividends	φ_d	0	10	6.13	5.37	6.19	6.71
Long Run Risks							
Persistence of LR Growth	ρ_x	0	1	0.892	0.842	0.886	0.96
Relative Vol of LR Growth	φ_x	0	1	0.183	0.14	0.443	0.562
Div Loading on LR Growth	ϕ_x	0	10	2.37	1.85	2.38	2.84
Persistence of LR Vol	ρ_h	0	1	0.895	0.85	0.895	0.943
Volatility of LR Vol	σ_h	0	1.5	0.954	0.864	0.964	1.01
Habit and Residual							
Persistence of Habit	ρ_s	0	1	0.933	0.804	0.954	0.999
Persistence of Residual	ρ_e	0	1	0.939	0.891	0.94	0.983
Price Dividend Coefficients							
LR Growth Coefficient	A_x	0	243	39.1	22.4	37.9	59.2
LR Vol Coefficient	A_V	-2.4e4	0	-87.7	-149	-84.2	-39.4
Habit Coefficient	A_s	0	0.93	0.282	0.0818	0.294	0.45
Residual Coefficient	A_e	0	0.23	0.147	0.11	0.147	0.186

Figure 1: Long Run Risk Parameter Estimate Details. Plots show posterior distributions of long run risks parameters from Table 1. The estimator finds significant evidence of persistent changes in expected growth and volatility.



growth— that is, expected consumption growth and consumption volatility both contain highly persistent components, with autocorrelations of about 0.90 annually. Since the identification of long run risk is an important issue in the literature, Figure 1 takes a closer look and plots the parameter distributions.

The figure shows that the high persistence of long run growth and volatility are estimated rather precisely. The entire the distribution of these parameters is above 0.80. These long run risks vary over time, that is, the relative volatility of long run growth φ_x and the volatility of long run volatility σ_h are statistically and economically significant. In terms of magnitudes, these parameters are similar to Schorfheide, Song, and Yaron (2016)'s estimates which omit asset price data.

Moving down Table 1, habit is estimated to be highly persistent with an autocorrelation of about 0.95, similar to Campbell and Cochrane (1999)'s calibration.

The persistence of the residual is also about 0.95, and moreover it is precisely estimated, with a lower bound of about 0.90. This high persistence illustrates a critical issue with long run risks and habit: The portion of asset prices that they can't explain is very long lived.

Now we come to the main parameters of interest: the price-dividend ratio coefficients. These parameters determine the contribution of each state to market volatility.

Critically, the residual coefficient is estimated to be quite high at about 15% per quarter. These parameters imply that the residual has an unconditional volatility of above 50%, more than large enough to account for the entire unconditional volatility of the log price dividend ratio (roughly 40%).

The large role of the residual is seen in the small posteriors of the other price-dividend coefficients. Since the priors were chosen so that each individual state could account for 100% of movements in the price dividend ratio, most coefficients are shrunk dramatically toward zero.

2.6. Main Result: Price Dividend Ratio Decompositions

With parameter estimates in hand, we can now address the main question of the paper: Which source of risk is the most important?

Figure 2 shows the estimated contributions of each state variable to the historical price dividend ratio. To create this plot, we find expected states using a particle filter and mean posterior parameters from Table 1. We then multiply the expected states by their respective posterior coefficients (the contribution of long run growth is $A_x \mathbb{E}(x_t)$).

The figure shows that the residual (yellow bars) played a dominant role in market volatility between 1929 and 2014. The residual is responsible for the relatively low asset prices in the 1940s and 50s, the bear market of the mid-1970s, and the big boom in the late 1990's. Indeed the residual closely tracks the price dividend ratio (blue line) for the vast majority of the sample.

Long run risks and habit play a non-trivial role. In particular, they weigh heavily on asset prices during the Great Depression and Great Recession. Long run growth and habit also boost prices somewhat in the 1960s.

Compared to the residual, however, long run risks and habit are relatively

Figure 2: Decomposition of the Historical Log Price-Dividend Ratio. We apply a particle filter to data on consumption, dividends, and stock prices using mean posterior parameter values (Table 1). A state's contribution to price/dividends is computed by multiplying the state's estimated price dividend ratio coefficient by the filtered mean state (see equation (1)). The residual contribution is dominant and closely tracks the price-dividend ratio.

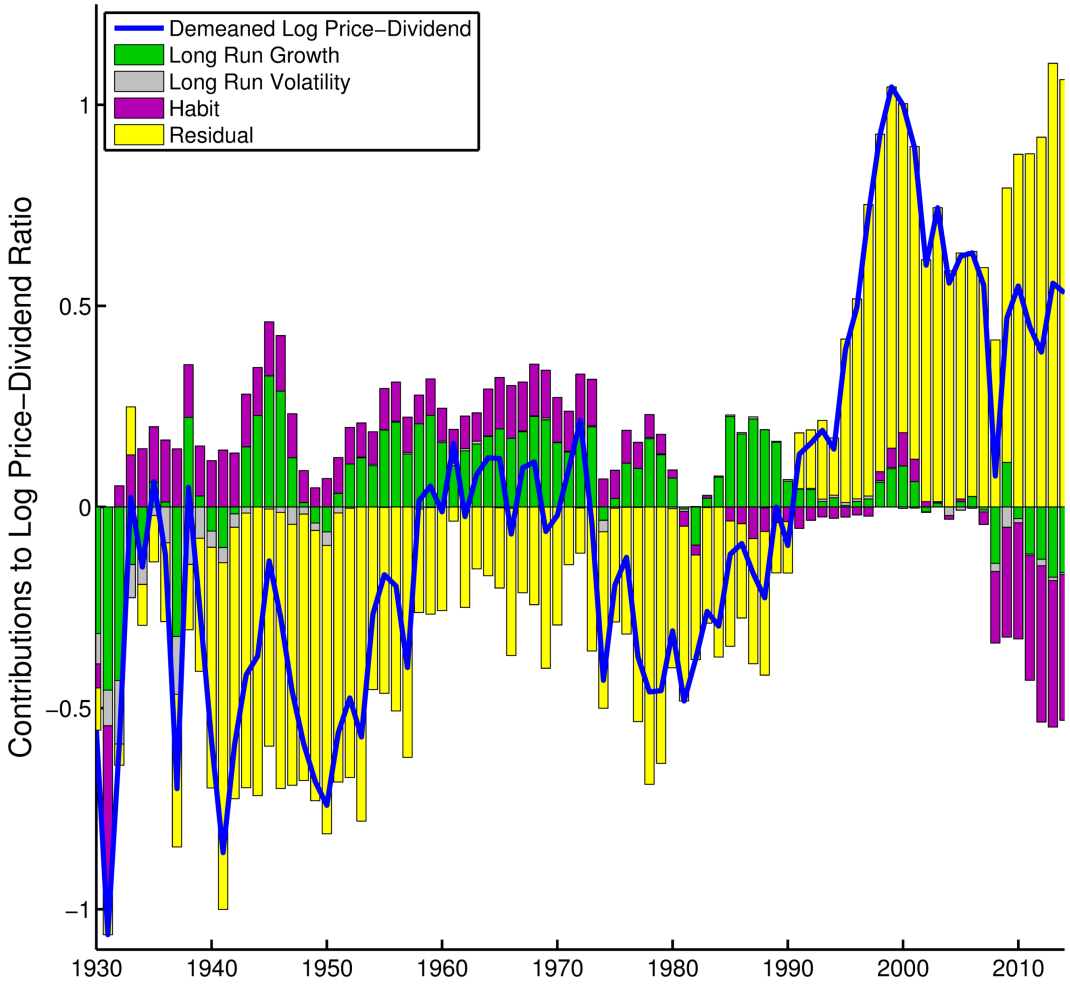
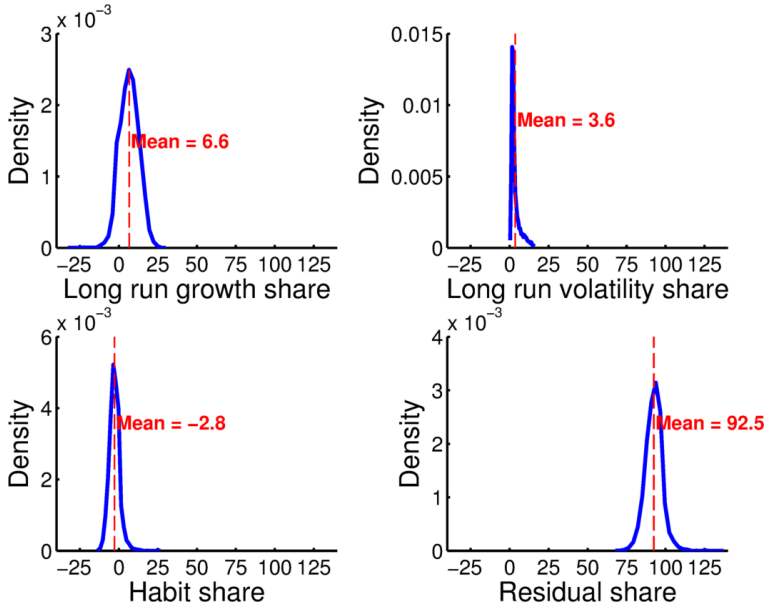


Figure 3: Price/Dividend Variance Decomposition. Shares are the percent of the variance of price/dividends accounted for by the given state variable. We draw parameters from the posterior, use the draw to filter expected states, and calculate variance contributions according to equation (8). The residual's dominant role is robust to estimation uncertainty.



unimportant. Outside of crises, long run volatility does almost nothing. And while long run growth and habit have more consistent effects, their effects are often the opposite of the overall pattern in asset prices. Indeed, while economic growth declined slowly between 1940 and 2014, the price-dividend ratio has trended up.

Figure 2 is created using mean posterior parameters, which do not account for estimation uncertainty. Is the dominant role of the residual robust to the uncertainty seen in the posterior parameters (Table 1)?

Figure 3 shows that the answer is yes. The figure plots variance decomposition of the price dividend ratio using the entire distribution of posterior parameters. The variance decomposition is calculated with covariances

$$\begin{aligned}\mathrm{Var}(pd_t) = \mathrm{Cov}(A_x x_t, pd_t) + \mathrm{Cov}\left(A_V \sigma_t^2, pd_t\right) \\ + \mathrm{Cov}(A_s s_t, pd_t) + \mathrm{Cov}(A_e e_t, pd_t)\end{aligned} \tag{8}$$

which can lead to negative shares if a state variable has a negative in-sample

correlation with pd_t.

Figure 3 shows that the residual's share of price dividend variance is almost entirely above 75%. Indeed, its posterior mean share is 93%, showing that it accounts for essentially all of the market volatility we have seen in the past 100 years.

Some of the distribution of the residual's share lies above 100%. This excess share is accounted for by the negative shares of long run growth and habit. Long run growth has declined over the sample while asset prices have grown, which sometimes leads to a negative share. Habit faces a similar correlation problem. We discuss both of these issue in depth in Section 3.2.

Long run growth, long run volatility, and habit have very small shares overall. On average, the three risks account for just 10% of the variance of price-dividends. There is some uncertainty in these estimates, however. The distributions of the shares for long run growth and volatility cover up to 25% and 15% respectively.

This section illustrates the main message of the paper: while long run risks played a non-trivial role in asset prices, something else is behind the vast majority of market volatility between 1929 and 2014. This result is robust to estimation uncertainty, and indeed, Section 4 shows that it is also robust to several prior and model specifications.

3. Supporting Results

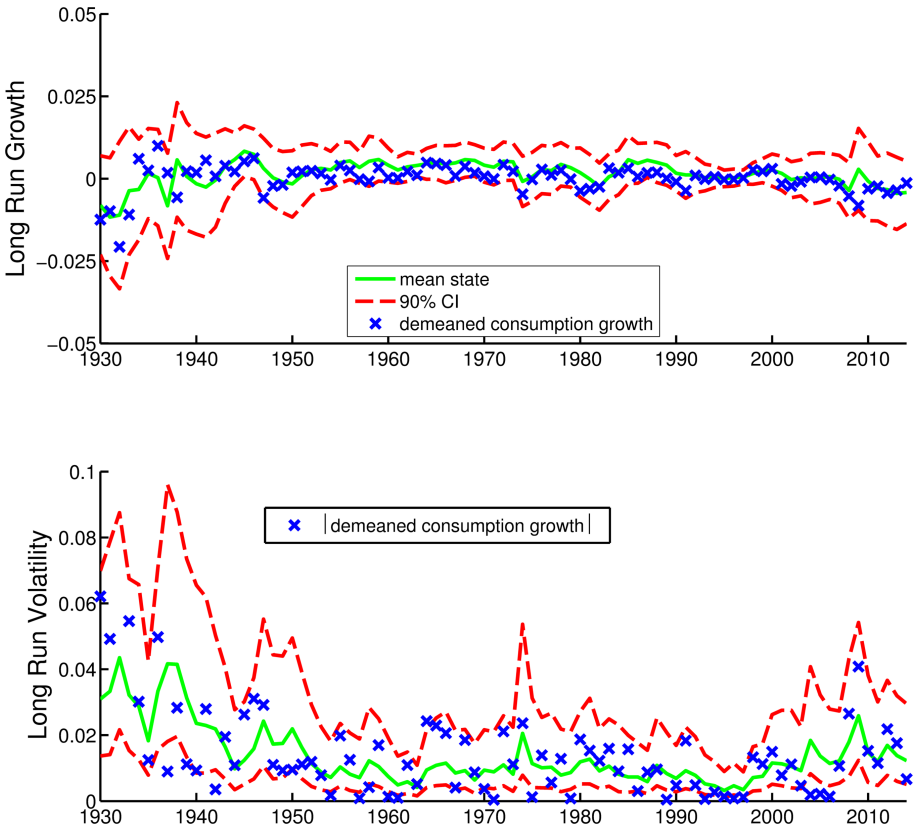
We now present evidence in support of our main results. We show that the estimated states are intuitive—that is, they match the related observables and narrative descriptions of economic history. We also show a simple OLS version of our price-dividend decomposition that generates similar results.

3.1. Estimated States Match Observables

Likelihood-based estimations generate historical estimates of latent states. These estimates provide an intuitive check on the price-dividend decompositions. With them we can ask: does the estimator do a good job describing economic history?

Figures 4 and 5 show the answer is yes. These plots show estimated historical

Figure 4: Baseline: Filtered States and Observables Part 1 of 2. We apply a particle filter to data on consumption, dividends, and stock prices, using mean posterior parameter values (Table 1). Scattered x's plot observables for comparison. The estimated states capture historical shifts in growth and volatility.

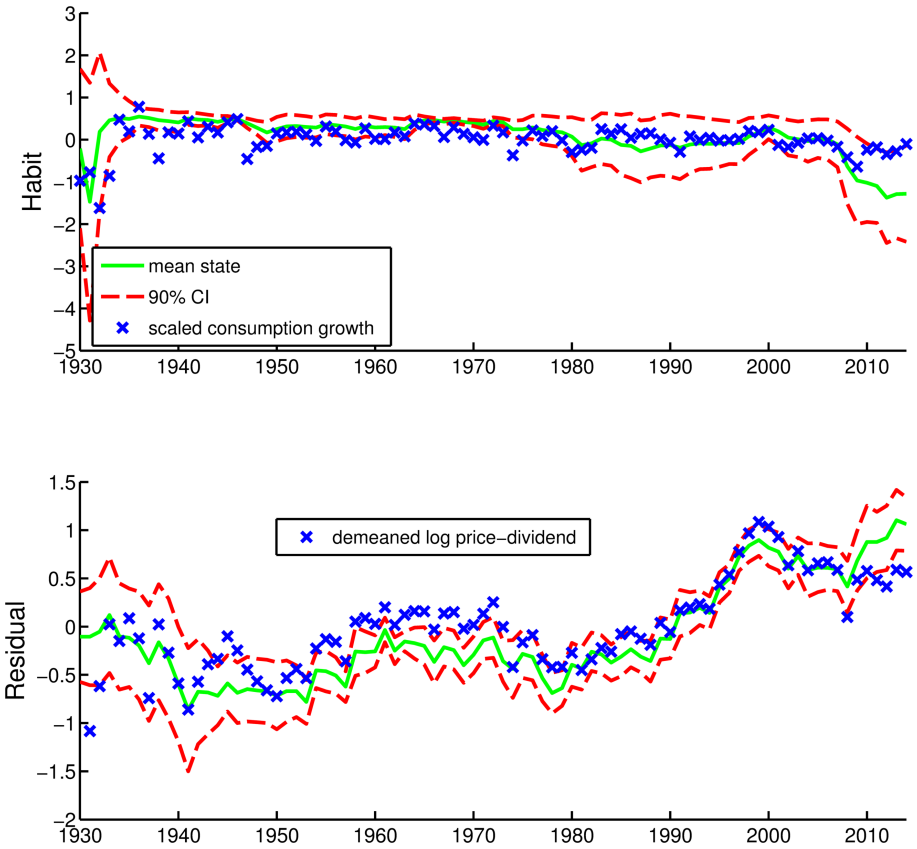


paths for long run growth, long run volatility, habit, and the residual. These paths are computed by using mean posterior parameter values (Table 1) and a particle filter.

The top panel of Figure 4 shows the estimated history of long run growth, along with demeaned consumption growth. Estimated long run growth path does a good job of capturing historical shifts in growth. The state identifies the Great Depression, the booming 60s, as well as the productivity slowdown of the 1970s. Interestingly, the estimator finds that growth has slowed since the 2008 Financial Crisis.

The bottom panel of Figure 4 shows the estimated history of long run volatility, along with the absolute value of demeaned consumption growth. The estimator does a good job of picking up key historical patterns: the decline in volatility

Figure 5: Baseline: Filtered States and Observables Part 2 of 2. We apply a particle filter to data on consumption, dividends and stock prices using mean posterior parameter values (Table 1). Habit is surplus consumption (equation (5)). x's plot observables for comparison. The top panel shows consumption growth scaled by the steady state $\lambda(s_t)$. The bottom panel shows the demeaned log price dividend ratio divided. Surplus consumption responds intuitively to consumption, and the residual closely tracks price-dividends.



after the war, the return of volatility in the 1970s, the Great Moderation, as well as the recent return of volatility in 2008.

Figure 5 shows the remaining two states: habit and the residual. The top panel shows habit, that is, surplus consumption, along with scaled consumption growth. Consumption growth is scaled by subtracting out its mean, and then multiplying by the steady state $\lambda(s_t)$, to imitate the "shock" term in the habit process (5).

Habit responds intuitively to consumption shocks. Surplus consumption is persistent, but responds to large changes in consumption growth. In particular, surplus consumption plummets in the Great Recession, as noted in Cochrane

(2011).

The bottom panel of Figure 5 shows the estimated residual, along the log price-dividend ratio. The panel shows that the two series closely follow each other. Notably, the residual tracks the broad historical patterns of the stock market: the rise in the 1960s, drop in the late 70s, and the 1990's boom. This close relationship does not necessarily follow from the fact that the residual is not observable in consumption or dividends. The residual follows the observable only if the explanatory power of the other variables is small.

3.2. OLS Price-Dividend Decomposition

Using the likelihood makes the estimator comprehensive: It accounts for all moments of the observables. But it also makes it non-trivial to dissect. What moment is driving the results?

This section provides a simple analysis of the driver. The price-dividend ratio has a relatively low correlation with consumption growth, consumption volatility, and past consumption growth. This low correlation leads to a large role for the residual.

To demonstrate this explanation, we perform a simplified version of estimation. Specifically, we construct state histories using only data on consumption and dividends. We then use OLS to regress the price-dividend ratio on the states. The resulting variance decomposition uses only correlation information, and the correlations lead OLS to conclude that a residual explains most of market volatility.

Long run risk states are constructed by running our Bayesian estimator using only consumption and dividend data. The long run risk parameter estimates are shown in Figure 6. Even without asset prices, the estimator finds evidence of significant long run risks in consumption and dividends. Both states are highly persistent, and quite volatile.

Figure 7 show the resulting state histories. The top and middle panels show that the estimated state paths are largely unaffected by the use of asset price data. The paths pick up the same features as in the baseline estimation (see Figure 4).

The bottom panel shows a simple construction of the habit process. Surplus consumption is constructed by initializing at the Campbell and Cochrane (1999)

Figure 6: Long Run Risk Parameter Estimates: Consumption and Dividends Only. Even without asset prices, the estimator finds evidence of significant long run risks in consumption and dividends.

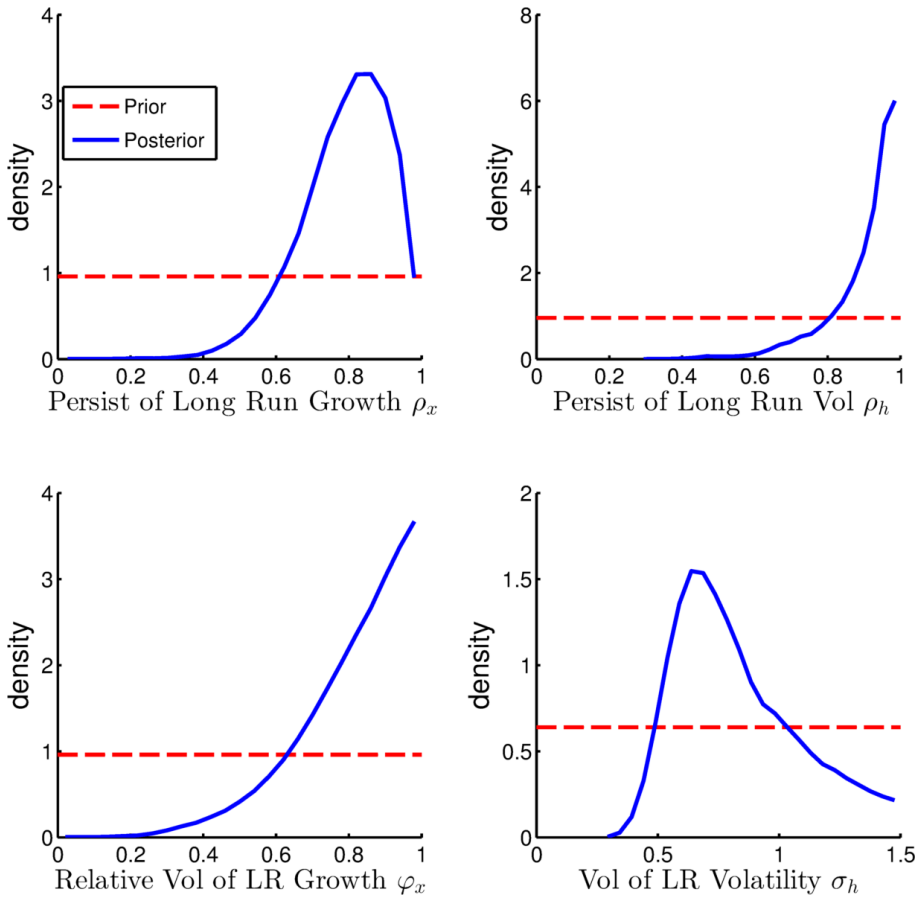
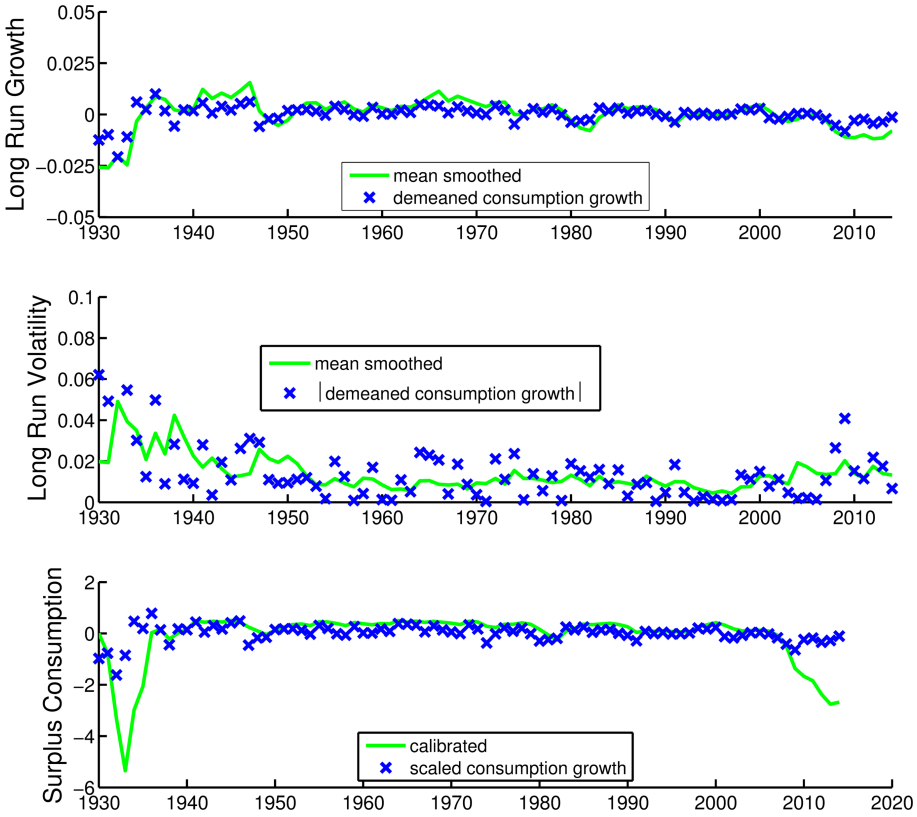


Figure 7: Simple State Histories. Long run growth and long run volatility are computed from an estimation using just consumption and dividend growth. Surplus consumption is constructed with (5) assuming that consumption growth is constant and applying parameter values from Campbell and Cochrane (1999). The bottom panel shows consumption growth scaled by the steady state $\lambda(s_t)$. The state paths are similar to Figures 4 and 5 and relatively unaffected by omitting asset prices.



steady state and then applying the surplus consumption process (5) under the assumption that consumption is i.i.d. The habit process parameters use Campbell and Cochrane (1999)'s parameter values. The resulting surplus consumption path looks much like the baseline estimation too (see Figure 5).

Comparing all three panels, Figure 7 shows that long run growth, long run volatility, and habit capture only the crisis periods of stock market history. In the panels we can see market crashes of the Great Depression and 2008 Financial Crisis. But nowhere can we see the extended stock market boom of the 1960s, the subsequent decade-long decline, or the dramatic bull market of the 1990s.

Table 2 provides a quantitative statement of this story. The table shows OLS regressions of the price-dividend ratio on the states in Figure 7— that is, an OLS estimate of our main equation of interest (1).

Table 2: OLS Price-Dividend Equation Estimates

We regress log price-dividends on the long run growth, long run volatility, and habit (Figure 7). All variables are scaled to have zero mean and a standard deviation of 1. The R^2 shows that most of the variance of log price-dividends is explained by a residual.

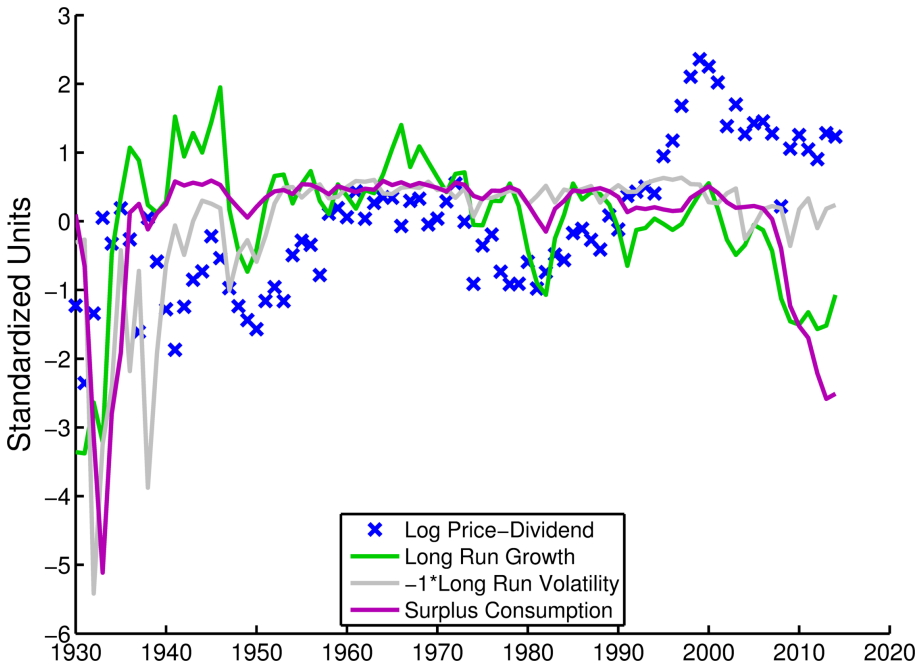
	Long Run Growth	Long Run Volatility	Surplus Consumption	
Coefficient	0.15	-0.48	-0.48	
s.e.	(0.14)	(0.12)	(0.16)	
R^2				0.17

The R^2 of this regression is 0.17, indicating that the residual accounts for the vast majority of the variance of price-dividends. Indeed, this R^2 is overstated, since the surplus consumption coefficient has the wrong sign. The baseline Bayesian estimation avoids these wrong sign by imposing priors, leading to an even larger role for the residual (Figure 3).

Figure 8 illustrates where the coefficients come from. The figure plots the state paths along with price-dividends. Long run growth generally trends downward over the sample period, while price-dividends grows. Thus, even though there are periods over which the two variables move together, the overall correlation is small, and the OLS estimates an insignificant coefficient.

Surplus consumption lags the price dividend ratio. It crashes after the stock

Figure 8: Simple States vs the Price-Dividend Ratio. All variables are scaled to have zero mean and a standard deviation of 1. Long run growth is estimated without asset prices, and surplus consumption is constructed using parameters from Campbell and Cochrane (1999). All states move in crises but otherwise are uncorrelated with price-dividends.



market crashes in both in the Great Depression and Great Recession. As the market recovers after these crashes, this leads to a negative correlation and the negative OLS coefficient.

Overall, Figure 8 shows that tricky mental maneuvers are required to line up asset prices with long run growth, long run volatility, and habit. There is some semblance of a relationships between the state variables and asset prices, but overall the correlations are low.

4. Robustness

As with any model-based econometrics, our method could potentially be sensitive to the model specification. The Bayesian approach raises the additional concern that the results could be sensitive to the choice of priors.

This section shows that our main result is quite robust. As long as the specification allows for the *possibility* of a large residual, the estimator concludes that the residual is dominant and closely follows the historical path of price/dividends. This result holds in (1) our baseline specification, (2) if we specify that the prior conditional volatility and persistence of long run risks are independent, (3) if we specify that habit responds to consumption growth rather than innovations, (4) if we model just long run risks and the residual, (5) if we model just habit and the residual, (6) if we rescale price dividend coefficients for the variance of the states. Indeed, this result holds for every specification that we have examined in the course of writing this paper (that allows for a residual).

Table 3 summarizes the robustness results. The table shows the shares of variance accounted for by long run growth, long run volatility, habit, and the residual across 6 model and prior specifications (see Equation 8). Under all 6 specifications, the residual accounts for the vast majority of market volatility, with a minimum share of 83%.

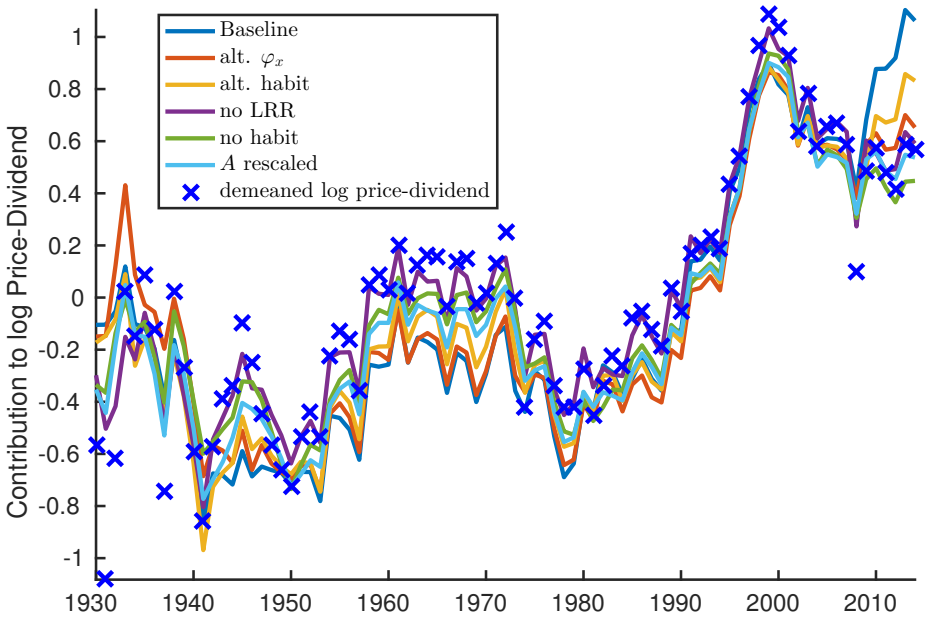
Figure 9 plots the residual under these 6 model specifications. Regardless of the specification, the residual is very highly correlated with the price dividend ratio and marks most key events in stock market history. Indeed, the role of the residual is very consistent: it tracks the stock market outside of the Great Depression and Great Recession.

Table 3: P/D variance shares in alternative model specifications.

Figures show percent contributions to the variance of the log price-dividend ratio following equation (8) under alternative model specifications. The table shows the mean and standard deviation (in parentheses) of the posterior distribution of the shares. Shares are computed using the filtered states evaluated at a sample of 5,000 draws from the posterior distribution.

	(1)	(2)	(3)	(4)	(5)	(6)
Variance share	Baseline	alt. φ_x	alt. habit	no LRR	no habit	A rescaled
Long-run Growth	9.3	4.9	4.1	-	10.7	9.9
	(6.6)	(9.2)	(3.9)		(2.8)	(2.9)
Long-Run Volatility	3.7	11.8	4.0	-	4.3	2.1
	(3.0)	(6.8)	(2.3)		(4.3)	(3.0)
Habit	-2.7	0.2	-0.8	8.8	-	1.0
	(3.3)	(1.9)	(3.6)	(5.5)		(1.3)
Residual	89.6	83.1	92.7	91.2	85.0	87.0
	(5.5)	(10.4)	(5.0)	(5.5)	(4.8)	(3.6)

Figure 9: Historical residual contributions to pd in alternative model specifications.



Plots show $A_e e_t$ under alternative specifications where A_e is the mean posterior residual coefficient on pd and e_t is the filtered expected residual computed at the mean posterior parameters.

The alternative specifications have intuitive motivations. The remainder of this section discusses these motivations and some details of each result.

Our baseline assumes that the prior parameters are independent, but of course this implies that other properties of the model are correlated. The alt. φ_x (Column (2) of Table 3) tests the importance of our independence structure. Specifically, we replace Equation (3) with

$$x_t = \rho_x x_{t-1} + \varphi_x \sigma_{t-1} \eta_{x,t},$$

thus removing the adjustment term $\sqrt{1-\rho_x^2}$ for the autocorrelation in the long-run growth process. We then place a uniform prior on the relative volatility of long-run growth: $\varphi_x \sim \mathscr{U}([0, 0.2])$. Under this specification, the variance share of long-run growth is somewhat lower than in the baseline, and the share of long-run volatility is higher. The dominant share of the residual remains.

Another concern readers may have about our specification is that we assume a specific relationship between long run risks and surplus consumption. We assume that habit responds to consumption innovations, which leads to the two state variables interacting in a specific way. The alt. Habit specification (Table 3 Column (3)) shows that an alternative and intuitive specification leads to similar results. Specifically, we replace Equation (5) with

$$\tilde{s}_t = \rho_s \tilde{s}_{t-1} + \lambda(\tilde{s}_{t-1})(\Delta c_t - \mu_c).$$

Under this alternative habit process, surplus consumption changes in response to all changes in consumption growth, not only unexpected changes. This means that long-run growth x_{t-1} also enters the current habit state. This alternative specification introduces a strong theoretical correlation between the long-run growth state and the habit state, but is closer to the original formulation of Campbell and Cochrane (1999). Under this alternative, the residual takes up even more of the variation in the price-dividend ratio than in the baseline.

A common theme in our results is that long run risks and habit both capture crises. A natural concern is that this correlation pollutes our results regarding the residual's share of market volatility. The no LRR (Table 3, Column (4)) and no habit (Column (5)) specifications examine this concern. We remove in turn the habit and long-run risk factors from the price-dividend ratio equation (1). When long-run risks are removed, the estimation assigns a variance share of about 9%

to habit, while when habit is removed, long-run risks account for a variance share of about 15%. These results make it clear that the high share of the residual is not due to competition between the different macro-asset pricing factors, but a robust feature coming from the data.

Finally, some readers may be concerned about our price-dividend coefficient priors, as these are not a standard type of variable to place priors over. Indeed, baseline prior for these variables was chosen for simplicity rather than a careful statistical or economic argument (Section 2.4).

The A rescaled specification (Table 3, Column (6)) places a different prior structure on the four coefficients of the price-dividend equation (1). Specifically, we choose the priors such that the theoretical variance of the factors in the price-dividend equation are identically and log-normally distributed. In doing so, we avoid as much as possible that the prior favors of any state variable in the variance decomposition, which is our prime object of interest in this paper. We construct four independent random variables T_x, T_V, T_s, T_e that are log-normally distributed with $T_i \sim \log\mathcal{N}\left(\mu_T, \sigma_T^2\right)$, $i = x, V, s, e$. We then construct the A_x coefficients conditional on the values of the remaining model parameters θ as follows:

$$
\begin{aligned}
A_x &= \sqrt{\frac{T_x}{\mathbb{V}\left[x_t \mid \theta\right]}}, & A_V &= -\sqrt{\frac{T_V}{\mathbb{V}\left[\tilde{\sigma}_t \mid \theta\right]}}, \\
A_s &= \sqrt{\frac{T_s}{\mathbb{V}\left[\tilde{s}_t \mid \theta\right]}}, & A_e &= \sqrt{\frac{T_e}{\mathbb{V}\left[e_t \mid \theta\right]}}.
\end{aligned} \tag{9}
$$

Here, $\mathbb{V}\left[x_t \mid \theta\right]$ etc. are the theoretical variances of the state variables conditional on the other model parameters. Note that we restrict the signs of the coefficients to conform to economic intuition. That is, we restrict the coefficients on long run growth and surplus consumption to be positive, and that on long run volatility to be negative. The result of this prior choice is that the prior distribution of the variances of the factors conditional on any θ are given simply by the T_i's, and in particularly iid among each other. We set $\sigma_T^2 = 2$ and μ_T such that the unconditional prior variance of the price-dividend ratio equals the observed variance in the data. Other, similarly diffuse distributions of the T_i's produce very similar results. Column (6) makes it clear that the prior structure on the factor loadings in Equation (1) do not matter much for the historical variance decomposition: The results are very similar to the baseline.

5. Interpretation of the Residual

We've shown that the residual is responsible for the bulk of market volatility. But what does this residual represent?

Broadly speaking, the fluctuations in the residual are a kind of excess stock market volatility. The residual moves closely with the price-dividend ratio, is unrelated to average economic growth (past or future), and is also unrelated to real volatility.

This description matches several theories in the literature. The theories fit into two broad categories: tractable representative agent models with hard-to-observe shocks to risk (such as variable disaster risk) and more complex models that link expected returns to observables other than consumption and dividends (such as incomplete market models). We cannot distinguish among these theories in this paper, but this section explains how these theories are consistent with our evidence, and suggests avenues for future resarch.

5.1. The Residual as a Hard-to-Observe, Time-Varying Risk

As the residual represents excess volatility, it naturally maps to hard-to-observe variations in risk. This kind of modeling has the virtue of being highly tractable, and thus leads to explicit predictions about a variety of asset market phenomena (Tsai and Wachter (2015)).

To see how the residual can be modeled as hard-to-observe variations in risk, suppose consumption growth experiences rare disaster shocks J_t

$$\Delta c_t = \mu_c + \sigma \eta_{c,t} + J_t \tag{10}$$

$$\Delta d_t = \mu_d + \phi_{\eta c} \sigma \eta_{c,t} + \varphi_d \sigma_t \eta_{d,t} + \phi_J J_t$$

$$J_t = \begin{cases} \bar{J}, & \text{with prob} \quad e_t \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

and that the probability of a disaster e_t is an AR(1) process

$$e_t = \bar{e} + \rho_e e_{t-1} + \sigma_e \eta_{e,t}. \tag{12}$$

Close the model with a representative Epstein-Zin household, and standard log-

linear approaches show that the price-dividend ratio is approximately

$$pd_t \approx \mu_{pd} + A_e(e_t - \bar{e}) \tag{13}$$

$$A_e = -\frac{\exp\left[(\phi_J - \gamma)\bar{J}\right] - 1 + \frac{\gamma - \frac{1}{\psi}}{1-\gamma}\left(\exp\left[(1-\gamma)\bar{J}\right] - 1\right)}{1 - \kappa_1 \rho_e} \frac{\sigma_e}{\sqrt{1-\rho_e^2}}. \tag{14}$$

where ψ and γ are the intertemporal substitution and the risk aversion parameters of the representative household.

Equations (10)-(13) show that the price-dividend ratio moves around in response to a variable e_t that is almost entirely unconnected to consumption and dividend growth. e_t shows up in equation (10) as the probability that $J_t > 0$, but the rare nature of these disasters means that (10) is empirically equivalent to one in which $J_t = 0$ all the time. More formally, simulating this model and applying our Bayesian estimation to the simulated data would result in A_e coefficients that are similar to what we found in U.S. data.

Thus, the probability of disaster functions just like a residual in the price-dividend equation. But other kinds of hard-to-observe risks act similarly, for example, the changes in the magnitude of ambiguity (Sbuelz and Trojani (2008)) or white noise shocks to habit (Bekaert, Engstrom, and Xing (2009)). Indeed, one could add hard-to-observe shocks to other models of asset prices and likely achieve a similar results.

The simplicity of this modeling approach means that it has the potential to be extended to generate additional quantitative predictions. In production economies, increases in hard-to-observe risks lead to clearly visible declines in output and investment (Gourio (2012), Ilut and Schneider (2014)). Similar real effects are seen in response to changes in habit (Chen (2016)) or beliefs (Winkler (2016)). Whether production economies can help distinguish between these theories is an interesting question for future research.

5.2. More Complex Models of the Residual

Directly linking the residual to observables other than aggregate consumption is possible, but requires more complicated models. Broadly speaking, there are two kinds of complications that achieve this result: (1) incomplete markets, and (2) imperfectly rational agents.

Incomplete market assume consumption risk is not shared efficiently, so aggregate consumption is no longer relevant for asset prices. This notion has a long history going back to Mankiw (1986). The most parsimonious way to model incomplete markets is by introducing idiosyncratic income risk (for example, Constantinides and Duffie (1996)). Schmidt (2015) finds that this channel can be made quantitatively significant with idiosyncratic disaster risk and Epstein-Zin preferences. The time series of idiosyncratic disaster risk is not readily observable, but Schmidt (2015) argues that initial claims for unemployment is a reasonable proxy, and finds that this measure is highly correlated with the price-dividend ratio.

Incomplete markets can also be modeled by focusing on institutional features, namely the fact that financial intermediaries appear to play a critical role in asset prices (Muir (2015)). In such models, only a subset of agents in the economy trade stocks, and these agents are capital constrained (He and Krishnamurthy (2013), Brunnermeier and Sannikov (2014)). As a result of these constraints, financial sector leverage becomes closely tied to the price-dividend ratio. As all sector valuations tend to move together, this proxy most certainly has a high correlation with the aggregate price-dividend ratio.

Models with imperfectly rational agents goes back to De Long et al. (1990). Most of this literature assumes irrational expectations motivated by psychology (for example, Hirshleifer, Li, and Yu (2015)). Barberis et al. (2015) apply this approach in a heterogeneous agent model that is qualitatively consistent with the data on survey expectations of returns. This qualitative relationship is difficult to match in completely rational models (Amromin and Sharpe (2013), Greenwood and Shleifer (2014), Koijen, Schmeling, and Vrugt (2015)).

A more recent literature assumes agents are rational, but form beliefs from a misspecified law of motion for stock prices (Adam and Marcet (2011), Adam, Marcet, and Nicolini (2016)). Since agents rationally update beliefs about stock prices based on observables, this approach naturally leads to relationships between the price-dividend ratio and non-consumption data. Adam, Marcet, and Beutel (2015) find that this approach leads to predictions about the price-dividend ratio and past returns which are quantitatively consistent with the data. Their model is also able to match the evidence on valuations and surveys expectations of returns.

6. Conclusion

We develop a model of asset prices that involves multiple sources of risk: long run growth, long run volatility, habit, and a persistent residual. The model is estimated using Bayesian methods which account for the entire likelihood of the data. We find that the residual is the most important source of risk, accounting for at least 80% of the variance of the price dividend ratio, as well as most recognizable historical features of the price-dividend series. Long run risks and habit play a role, but primarily in crisis periods.

This analysis raises the bar for asset pricing models. Many macro finance models which are quite successful at matching moments struggle when confronted with the entire likelihood of the data. Simply put, the conditional correlations between asset prices and real variables is too small for the estimator to put a lot of stock in real factors.

Models with hard-to-observe changes in risk (such as variable disaster risk) pass these tests, but only do so because they hide the mechanism from empirical scrutiny. Indeed it is difficult to falsify a model in which asset prices are driven by fluctuations in the conditional density of rare events. More complex models can link risk changes to observables, but typically can only be evaluated based on their qualitative predictions.

Nevertheless, the results of this paper illustrate the importance of unobservable drivers of asset price data. Policy makers, market participants, and academic economists which desire to understand why valuations are currently elevated, or why valuations have recently plummeted should be careful when attributing these changes to movements in long run growth, long run volatility, or habit-based risk aversion.

7. Appendix

7.1. State Space Formulation

To estimate the model, we write it in a state space formulation following Schorfheide, Song, and Yaron (2016). We make some small altertations to their approach to accommodate habit and to simplify the large matricies.

In the end, we have transition equations

$$\begin{aligned} h_t &= \rho_h h_{t-1} + \sigma_h \sqrt{1-\rho_h^2} w_t \\ s_t &= \Phi(s_{t-1}) s_t + \Sigma_s(s_{t-1}) \eta_t. \end{aligned} \tag{15}$$

and observation equations

$$y_t = \mu_y + Zs + A_V \bar{\sigma}^2 (\exp(2h_t) - \exp(2\sigma_h^2)) \tag{16}$$

where s_t, y_t are vectors of augmented states and observables, w_t, η_t, ϵ_t are vectors of standard normal independent noise, and $\Phi(s_{t-1}), \Sigma_s(s_{t-1}), \mu_y, Z$ are vectors and matricies that describe the evolution of the system. Note that these vectors and matricies can depend on the previous state or on the current time period.

We'll now derive these vectors and matricies.

Observables and States Denote the innovations on consumption and dividends by:

$$\tilde{\eta}_{c,t} = \bar{\sigma} \exp(h_t) \eta_{c,t} \tag{17}$$

$$\tilde{\eta}_{d,t} = \bar{\sigma} \exp(h_t) \eta_{c,t} \left(\phi_{\eta c} \varphi_d \eta_{d,t} \right). \tag{18}$$

Further, denote the log surplus consumption ratio by u_t. Stack the non-volatility standard state variables $z_t \equiv [x_t, u_t, e_t]'$. Then we can map observables

to an augmented state s_t (as in (16)) with:

$$\underbrace{\begin{bmatrix}\Delta c_t\\ \Delta d_t\\ pd_t\end{bmatrix}}_{y_t} = \underbrace{\begin{bmatrix}\mu_c\\ \mu_d\\ \mu_{pd}\end{bmatrix}}_{\mu_y} + \underbrace{\begin{bmatrix}0 & 1 & 1 & 0 & 0 & 0\\ 0 & \phi_x & 0 & 1 & 0 & 0\\ [A_x, A_s, \sigma_e] & 0 & 0 & 0 & 0 & 0\\ 0 & A_x & 0 & 0 & A_s & \sigma_e\end{bmatrix}}_{Z} \underbrace{\begin{bmatrix}z_t\\ x_{t-1}\\ \tilde{\eta}_{c,t}\\ \tilde{\eta}_{d,t}\\ u_{t-1}\\ e_{t-1}\end{bmatrix}}_{s_t}$$

$$+ A_V \bar{\sigma}^2 (\exp(2h_t) - \exp(2\sigma_h^2)).$$

State Transition Finally, we can relate the augmented state s_t to its lag as in (15):

$$\underbrace{\begin{bmatrix}z_t\\ x_{t-1}\\ \tilde{\eta}_{c,t}\\ \tilde{\eta}_{d,t}\\ u_{t-1}\\ e_{t-1}\end{bmatrix}}_{s_t} = \underbrace{\begin{bmatrix}\begin{bmatrix}\rho_x & 0 & 0\\ \lambda(u_{t-1}) & \rho_u & 0\\ 0 & 0 & \rho_e\end{bmatrix} & 0 & 0 & 0 & 0 & 0\\ \begin{bmatrix}1 & 0 & 0\end{bmatrix} & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0\\ \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix} & 0 & 0 & 0 & 0 & 0\end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix}z_{t-1}\\ x_{t-2}\\ \tilde{\eta}_{c,t-1}\\ \tilde{\eta}_{d,t-1}\\ u_{t-2}\\ e_{t-2}\end{bmatrix}}_{s_{t-1}}$$

$$+ \underbrace{\begin{bmatrix}\begin{bmatrix}\varphi_x \bar{\sigma} \exp(h_{t-1})\\ 0\\ 0\end{bmatrix} & \begin{bmatrix}0\\ \lambda(u_{t-1})\bar{\sigma}\exp(h_{t-1})\\ 0\end{bmatrix} & \begin{bmatrix}0\\ 0\\ \sigma_e\end{bmatrix} & 0\\ 0 & \bar{\sigma}\exp(h_{t-1}) & 0 & 0\\ 0 & 0 & 0 & \varphi_d \bar{\sigma}\exp(h_{t-1})\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix}}_{\Sigma_s(s_{t-1})} \underbrace{\begin{bmatrix}\eta_{x,t}\\ \eta_{c,t}\\ \eta_{e,t}\\ \eta_{d,t}\end{bmatrix}}_{\eta_t}.$$

7.2. Particle Filter Details

We first describe the big picture of the algorithm. We then go on to give the details of how each distribution is defined. Index particles by a superscript $i = 1, ..., M$.

1. Begin with a set of particles $[s^i_{t-1}, h^i_{t-1}]$ and weights π^i_{t-1}.
2. Draw $h^i_t \sim q(h^i_t | pd^i_t, h^i_{t-1}, s^i_{t-1})$ for each i (q will be defined later)
3. Draw $s^i_t \sim p(s_t | y^o_t, h^i_t, h^i_{t-1}, s^i_{t-1})$ for each i (p will be derived later)
4. Update particle weights using

$$\pi^i_t = \pi^i_{t-1} [\text{update factor}]^i \tag{19}$$

$$[\text{update factor}]^i = p(y^o_t | h^i_t, h^i_{t-1}, s^i_{t-1}) \left[\frac{p(h^i_t | h^i_{t-1})}{q(h^i_t | pd_t, h^i_{t-1}, s^i_{t-1})} \right] \tag{20}$$

   We'll explain how to derive this update factor later.

5. Estimate log-likelihood contribution

$$\log \hat{p}(y^o_t) = \log \left(\sum_i \pi^i_{t-1} [\text{update factor}]^i \right) \tag{21}$$

6. Resample: if $\frac{1}{M^2 \sum_i (\pi^i_t)^2} < 0.5$ redraw $\{\pi^i_t\}$ using a multinomial distribution with probabilities $\{\pi^i_t\}$.

Since the remainder of this section discusses operations applied to every particle i, we drop the superscript for ease of reading.

7.2.1. Proposal Distribution for $h_t \sim q(h_t | pd_t, h_{t-1}, s_{t-1})$

We draw h_t based off of pd_t. This differs from Schorfheide, Song, and Yaron (2016) and helps a bit in making the filter more accurate.

The basic idea is that we want to draw h_t as close to the true probability $p(h_t | pd_t, h_{t-1}, s_{t-1})$ as possible. Unfortunately, the relationship between pd_t and h_t is nonlinear, so we have to work with an approximation for

$q(h_t|pd_t, h_{t-1}, s_{t-1}) \approx p(h_t|pd_t, h_{t-1}, s_{t-1})$ based off of the linear system

$$\begin{aligned} pd_t &= pd_0 + \tilde{A}_{V,t} h_t + \tilde{\sigma}_{pd} \eta_{pd,t} \\ h_t &= \rho_h h_{t-1} + \sigma_h \sqrt{1-\rho_h^2} w_t \end{aligned} \tag{22}$$

where $\eta_{pd,t} \sim N(0,1) i.i.d.$ and

$$\begin{aligned} pd_0 &= \Omega_{0pd} + \Omega_{pd}(Z_{t0} + Z_t s_{t|t-1} + A_V \bar{\sigma}^2 \left[\exp(2\rho_h h_{t-1})(1-2\rho_h h_{t-1}) - \exp(2\sigma_h^2)\right] \\ \tilde{A}_V &= A_V \bar{\sigma}^2 2 \exp(2\rho_h h_{t-1}) \\ \tilde{\sigma}_{pd} &= \sqrt{(\Omega_{pd} Z \Sigma_s)^2 + \sigma_{pd}^2} \end{aligned}$$

That is, (22) is based off Taylor expanding (1) around $h_t = \rho_h h_{t-1}$. We find this approximation is good for relevant parameter values.

We apply a Kalman filter to system (22) to obtain $h_{t|t}, Vh_{t,t}$. Finally, we draw using $q(h_t|pd_t, h_{t-1}, s_{t-1}) \sim N(h_{t|t}, Vh_{t,t})$.

7.2.2. The proposal distribution for $s_t \sim p(s_t|y_t^o, h_t, h_{t-1}, s_{t-1})$

to be completed

7.2.3. Simplifying the Update Factor

We simplify the particle filter update step by taking advantage of the conditional Gaussian properties of the model and using Bayes' theorem.

The standard generic particle filter update follows

$$\pi_t = \pi_{t-1} \text{update weight} \tag{23}$$

where the update weight is

$$\text{update weight} \equiv p(y_t|s_t, h_t) \frac{p(s_t, h_t|s_{t-1}, h_{t-1})}{q(s_t, h_t|s_{t-1}, h_{t-1}, y_t)}$$

and q is the proposal distribution in the propogation step (see Herbst and Schorfheide (2014)). We can simplify the update weight by using the the proper-

ties of the proposal distribution as well as Bayes' rule.

$$\text{update weight} = p(y_t|s_t,h_t)\frac{p(s_t|s_{t-1},h_{t-1})}{p(s_t|y_t,h_t,s_{t-1},h_{t-1})}\frac{p(h_t|h_{t-1})}{q(h_t|y_t,h_{t-1},s_{t-1})} \tag{24}$$

$$= p(y_t|h_t,h_{t-1},s_{t-1})\left[\frac{p(h_t|h_{t-1})}{q(h_t|y_t,h_{t-1},s_{t-1})}\right] \tag{25}$$

7.3. Bayesian MCMC Method

We wrap the filter in a standard Random Walk Metropolis-Hastings algorithm in order to derive parameter estimates (Herbst and Schorfheide (2014)). We run standard initial tuning runs of the algorithm in order to choose a good proposal distribution. That is, we begin by finding the highest likelihood on a short random search over the prior distribution (500 vectors). We then run a 500 vector chain and use the variance of the posterior as a step direction. Last we test various step sizes using 500 vector chains in order to find a step size which produces an acceptance rate of about 0.3. The final MCMC chain is 500,000 parameter vectors long and we burn the first 100,000 vectors to focus on the ergodic distribution.

7.4. Detailed Baseline Posteriors

This section shows the prior and posteriors for all of the parameters. XXX

Figure 10: Baseline: Posterior Details 1 of 2.

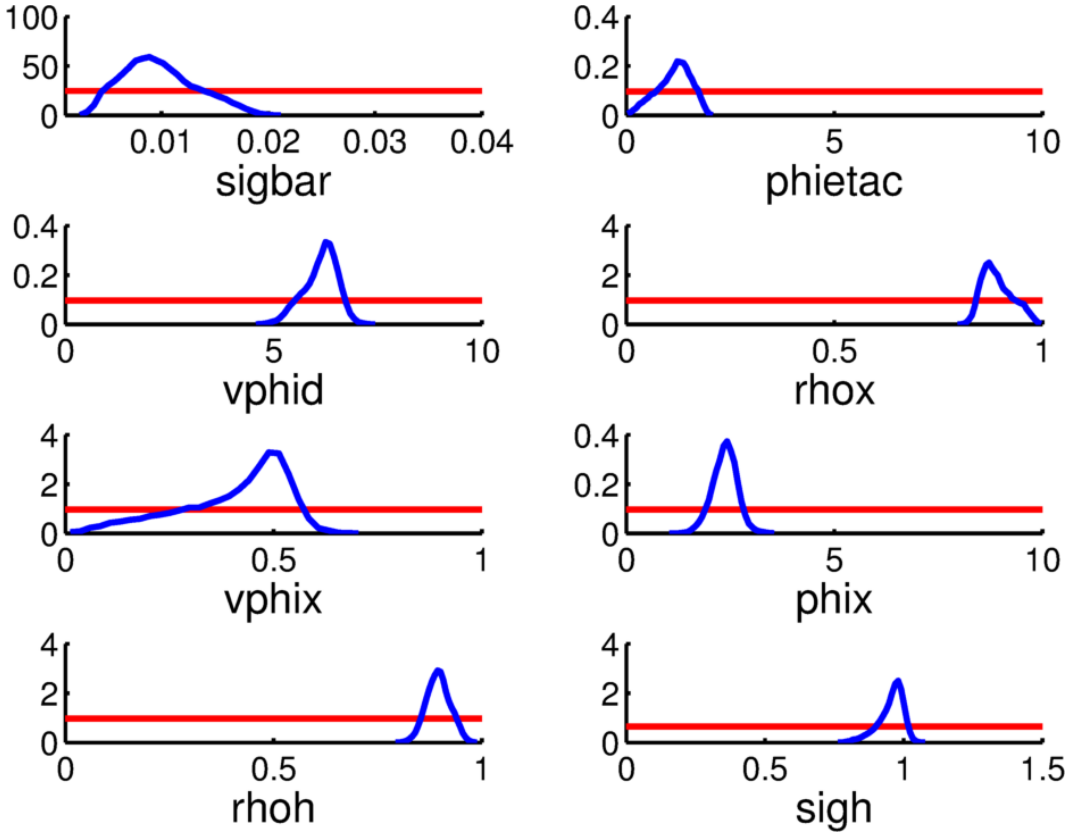
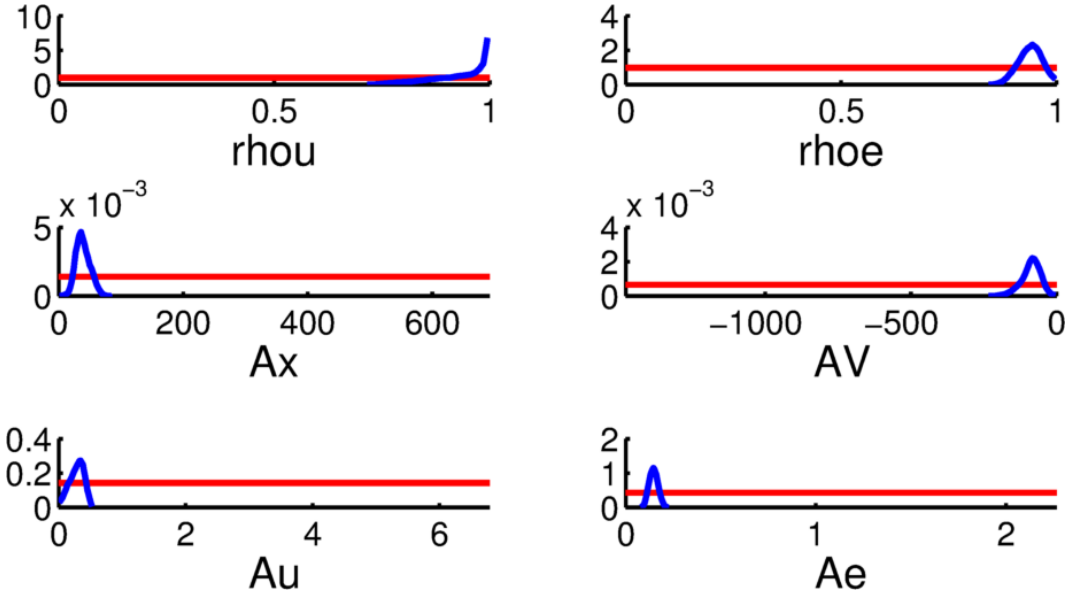


Figure 11: Baseline: Posterior Details 2 of 2.



References

Adam, Klaus and Albert Marcet. "Internal rationality, imperfect market knowledge and asset prices". *Journal of Economic Theory* 146.3 (2011), pp. 1224–1252.

Adam, Klaus, Albert Marcet, and Johannes Beutel. "Stock price booms and expected capital gains". *Available at SSRN 2634053* (2015).

Adam, Klaus, Albert Marcet, and Juan Pablo Nicolini. "Stock market volatility and learning". *The Journal of Finance* 71.1 (2016), pp. 33–82.

Aldrich, Eric M and A Ronald Gallant. "Habit, long-run risks, prospect? a statistical inquiry". *Journal of Financial Econometrics* 9.4 (2011), pp. 589–618.

Amromin, Gene and Steven A Sharpe. "From the horse's mouth: economic conditions and investor expectations of risk and return". *Management Science* 60.4 (2013), pp. 845–866.

Bansal, Ravi, A Ronald Gallant, and George Tauchen. "Rational pessimism, rational exuberance, and asset pricing models". *The Review of Economic Studies* 74.4 (2007), pp. 1005–1033.

Bansal, Ravi, Dana Kiku, and Amir Yaron. "An Empirical Evaluation of the Long-Run Risks Model for Asset Prices". *Critical Finance Review* 1.1 (2012), pp. 183–221.

— *Risks for the long run: Estimation with time aggregation*. Tech. rep. National Bureau of Economic Research, 2012.

Bansal, Ravi and Amir Yaron. "Risks for the long run: A potential resolution of asset pricing puzzles". *The Journal of Finance* 59.4 (2004), pp. 1481–1509.

Barberis, Nicholas et al. "X-CAPM: An extrapolative capital asset pricing model". *Journal of Financial Economics* 115.1 (2015), pp. 1–24.

Barro, Robert J. "Rare disasters and asset markets in the twentieth century". *The Quarterly Journal of Economics* 121.3 (2006), pp. 823–866.

Barro, Robert J. and Tao Jin. *Rare Events and Long-Run Risks*. Working Paper 21871. National Bureau of Economic Research, Jan. 2016. DOI: 10.3386/w21871. URL: http://www.nber.org/papers/w21871.

Beeler, Jason and John Y. Campbell. "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment". *Critical Finance Review* 1.1 (2012), pp. 141–182. ISSN: 2164-5744. DOI: 10.1561/104.00000004. URL: http://dx.doi.org/10.1561/104.00000004.

Bekaert, G., E. Engstrom, and Y. Xing. “Risk, uncertainty, and asset prices”. *Journal of Financial Economics* 91.1 (2009), pp. 59–82.

Brunnermeier, Markus K and Yuliy Sannikov. “A macroeconomic model with a financial sector”. *The American Economic Review* 104.2 (2014), pp. 379–421.

Campbell, John Y. and John H. Cochrane. “By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior”. *Journal of Political Economy* 107.2 (1999), pp. 205–251.

Campbell, John Y and John H Cochrane. “Explaining the Poor Performance of Consumption-based Asset Pricing Models”. *The Journal of Finance* 55.6 (2000), pp. 2863–2878.

Carvalho, Carlos et al. “Particle learning and smoothing”. *Statistical Science* 25.1 (2010), pp. 88–106.

Chang, Andrew C and Phillip Li. “Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say’Usually Not’” (2015).

Chen, Andrew Y. “External Habit in a Production Economy: A Model of Asset Prices and Consumption Volatility Risk”. *Available at SSRN 2479372* (2016).

Cochrane, J.H. “Presidential address: Discount rates”. *The Journal of Finance* 66.4 (2011), pp. 1047–1108.

Cochrane, John H. *Macro-finance*. Tech. rep. National Bureau of Economic Research, 2016.

Constantinides, George M. “Habit formation: A resolution of the equity premium puzzle”. *Journal of political Economy* (1990), pp. 519–543.

Constantinides, George M and Darrell Duffie. “Asset pricing with heterogeneous consumers”. *Journal of Political economy* (1996), pp. 219–240.

De Long, J Bradford et al. “Noise trader risk in financial markets”. *Journal of political Economy* (1990), pp. 703–738.

Dew-Becker, Ian et al. *The price of variance risk*. Tech. rep. National Bureau of Economic Research, 2015.

Ferson, Wayne E and Campbell R Harvey. “Seasonality and Consumption-Based Asset Pricing”. *The Journal of Finance* 47.2 (1992), pp. 511–552.

Gabaix, Xavier. “Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance”. *The Quarterly Journal of Economics* 127.2 (2012), pp. 645–700.

Gourio, François. “Disaster Risk and Business Cycles”. *The American Economic Review* 102.6 (2012), pp. 2734–66.

Greenwood, Robin and Andrei Shleifer. "Expectations of returns and expected returns". *Review of Financial Studies* (2014), hht082.

Guvenen, Fatih. "A parsimonious macroeconomic model for asset pricing". *Econometrica* 77.6 (2009), pp. 1711–1750. ISSN: 1468-0262.

He, Zhiguo and Arvind Krishnamurthy. "Intermediary asset pricing". *The American Economic Review* 103.2 (2013), pp. 732–770.

Herbst, Edward and Frank Schorfheide. "Bayesian inference for DSGE models". *Manuscript. Board of Governors of the Federal Reserve System, Washington, DC* (2014).

Hirshleifer, David, Jun Li, and Jianfeng Yu. "Asset pricing in production economies with extrapolative expectations". *Journal of Monetary Economics* 76 (2015), pp. 87–106.

Ilut, Cosmin L and Martin Schneider. "Ambiguous business cycles". *The American Economic Review* 104.8 (2014), pp. 2368–2399.

Koijen, Ralph SJ, Maik Schmeling, and Evert B Vrugt. "Survey expectations of returns and asset pricing puzzles". *Available at SSRN 2513701* (2015).

Mankiw, N Gregory. "The equity premium and the concentration of aggregate shocks". *Journal of Financial Economics* 17.1 (1986), pp. 211–219.

Mankiw, N Gregory and Stephen P Zeldes. "The consumption of stockholders and nonstockholders". *Journal of financial Economics* 29.1 (1991), pp. 97–112.

Mehra, Rajnish and Edward C Prescott. "The equity premium: A puzzle". *Journal of monetary Economics* 15.2 (1985), pp. 145–161.

Muir, Tyler. "Financial crises and risk premia". *Available at SSRN 2379608* (2015).

Rietz, Thomas A. "The equity risk premium a solution". *Journal of monetary Economics* 22.1 (1988), pp. 117–131.

Sbuelz, Alessandro and Fabio Trojani. "Asset prices with locally constrained-entropy recursive multiple-priors utility". *Journal of Economic Dynamics and Control* 32.11 (2008), pp. 3695–3717.

Schmidt, Lawrence DW. "Climbing and falling off the ladder: Asset pricing implications of labor market event risk". *Available at SSRN 2471342* (2015).

Schorfheide, Frank, Dongho Song, and Amir Yaron. *Identifying long-run risks: A bayesian mixed-frequency approach*. Tech. rep. National Bureau of Economic Research, 2014.

— *Identifying long-run risks: A bayesian mixed-frequency approach*. Tech. rep. Mimeo, 2016.

Shiller, RJ. "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71.3 (1981), pp. 421–436.

Tsai, Jerry and Jessica A Wachter. *Disaster risk and its implications for asset pricing*. Tech. rep. National Bureau of Economic Research, 2015.

Van Binsbergen, Jules, Michael Brandt, and Ralph SJ Koijen. "On the timing and pricing of dividends". *The American Economic Review* 102.4 (2012), pp. 1596–1618.

Wachter, Jessica A. "Can Time-Varying Risk of Rare Disasters Explain Aggregate Stock Market Volatility?" *The Journal of Finance* 68.3 (2013), pp. 987–1035.

Wilcox, David W. "The construction of US consumption data: some facts and their implications for empirical work". *The American economic review* (1992), pp. 922–941.

Winkler, Fabian. "The role of learning for asset prices and business cycles" (2016).

Yang, Wei. "Intertemporal Substitution and Equity Premium". *Review of Finance* (2015). DOI: 10.1093/rof/rfv004. eprint: http://rof.oxfordjournals.org/content/early/2015/03/21/rof.rfv004.full.pdf+html. URL: http://rof.oxfordjournals.org/content/early/2015/03/21/rof.rfv004.abstract.