Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

A Likelihood-Based Comparison of Macro Asset Pricing Models

Andrew Y. Chen, Rebecca Wasyk, and Fabian Winkler

2017-024

Please cite this paper as:
Chen, Andrew Y., Rebecca Wasyk, and Fabian Winkler (2017). "A Likelihood-Based Comparison of Macro Asset Pricing Models," Finance and Economics Discussion Series 2017-024. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.024.

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

March 2017*

Abstract

We estimate asset pricing models with multiple risks: long-run growth, long-run volatility, habit, and a residual. The Bayesian estimation accounts for the entire likelihood of consumption, dividends, and the price-dividend ratio. We find that the residual represents at least 80% of the variance of the price-dividend ratio. Moreover, the residual tracks most recognizable features of stock market history such as the 1990's boom and bust. Long run risks and habit contribute primarily in crises. The dominance of the residual comes from the low correlation between asset prices and consumption growth moments. We discuss theories which are consistent with our results.

*We thank Francesco Bianchi, Ed Herbst, Francisco Palomino, and Missaka Warusawitharana for advice and comments, and Carter Bryson and Chris Campano for excellent research assistance. 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. The evaluation uses a model which divides the price-dividend ratio into identifiable macro risks (habit, long run growth, and long run volatility), and hard-to-observe fluctuations in risk (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 estimation leads to a decomposition 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 model specifications. Moreover, the smoothed residual tracks most of the recognizable features of the U.S. stock market's history, such as the booms of the 1950s and 1990s, and the busts of the 1970s and early 2000s. 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

[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 target moments, as our Bayesian estimation accounts for the entire likelihood of consumption, dividends, and the price-dividend ratio. This methodology cuts through the problem of weighing disparate pieces of evidence from the moment matching literature. How important is the excessively strong dividend predictability in long run risks models? 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 provides a succinct answer to these questions.

Fluctuations in the residual represent "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 or beliefs (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 that compel us to deviate from the standard approach of looking for equilibrium among optimizing agents.

The semi-structural model lets the data 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 because 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 stripped-down version of our estimator. Specifically, we extract state paths using only data on consumption and dividends. We then use OLS to regress the price-dividend ratio on the states. The variance decomposition from these procedure uses only covariances between price/dividends and the state paths, and this information leads OLS 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 the 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 estimation suggests that the data frequency is not critical. Nevertheless, some studies suggest that a monthly model is important (Bansal, Kiku, and Yaron (2012a), 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) clarify the picture by using a Bayesian framework to compare habit, long run risks, and prospect theory. Our results echo theirs: long run risks are critical for addressing the volatile 1930s, but are less important for other time periods. We differ from Aldrich and Gallant, however, by allowing a residual 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, Giglio, Le, and Rodriguez (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, if one accounts for the entire likelihood of the data.

Our paper also owes a large intellectual debt to Schorfheide, Song, and Yaron (2016), who pioneer the use a particle filter and Bayesian Markov chain Monte Carlo (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 equation. 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 describes the model, explains the estimation method, and ends with the main results: decompositions of the price-dividend ratio (Section 2.6). Along the way, we discuss 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 is long run growth, $\tilde{\sigma}_t$ is long run volatility, $\tilde{s}_t$ is surplus consumption (habit), and e_t is a residual. The tildes over $\tilde{\sigma}_t$ and $\tilde{s}_t$ indicate that these variables are demeaned ($\tilde{\sigma}_t^2 = \sigma_t^2 - \mathbb{E}(\sigma_t^2)$ and $\tilde{s}_t = s_t - \bar{s}$). These transformations imply that μ_{pd} is the mean log price-dividend ratio.

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 through several economic models (see Section 5).

Our goal is to estimate the coefficients A_x, A_V, A_s, A_e and find the smoothed paths of the states $x_t, \tilde{\sigma}_t^2, \tilde{s}_t$ and e_t. The coefficients and smoothed 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 (2016)'s Epstein-Zin habit model to include time-varying disaster probability.

All of the states are latent, but they can be identified by their linkages with observables. 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

AR(1)

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

and long run volatility σ_t evolves according to

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

The volatility specification borrows a technical fix from Schorfheide, Song, and Yaron (2016), but otherwise the above consumption and dividends are equivalent to Bansal, Kiku, and Yaron (2012a)'s processes.[2] Schorfheide et al's fix ensures that volatility is always positive.

Importantly, our specification does not include the multiple volatility processes of Schorfheide et al. Using a single volatility process is consistent with the bulk of the long run risk literature (for example, Bansal, Kiku, and Yaron (2012a)) and makes the estimation simpler. However, this assumption is restrictive in that it assumes that the impact of volatility on the price-dividend ratio can be identified with "short-run" realized consumption volatility (see equation (2)).

Surplus consumption $\tilde{s}_t$ is also identified by consumption growth, but is more "backward looking." $\tilde{s}_t$ is an AR(1)-like process

$$\tilde{s}_t = \rho_s \tilde{s}_{t-1} + \lambda(\tilde{s}_{t-1})(\Delta c_t - \mathbb{E}_{t-1}\Delta c_t) \tag{5}$$
$$\lambda(\tilde{s}_{t-1}) = \begin{cases} \exp(-\overline{s})\sqrt{1-2\tilde{s}_{t-1}} - 1, & \tilde{s}_t \le \frac{1}{2}[1-\exp(2\overline{s})] \\ 0, & \text{otherwise} \end{cases}$$

which is equivalent to Campbell and Cochrane (1999)'s habit process. This process means that surplus consumption is the average of past consumption growth and that habit is the average of past consumption levels (Campbell (2003), Chen (2016)). In the robustness section (Section 4), we examine a pro-

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

cess in which $\tilde{s}_t$ responds to consumption growth itself rather than innovations. This alternative assumption does not change the main results.

Unlike the other state variables, the residual is not identified by either consumption or dividends. It is simply an AR(1)

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

and is thus identified primarily by the price-dividend ratio. e_t captures everything in 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 Census Bureau's sophisticated X-13ARIMA-SEATS seasonal adjustment. 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. This forward-looking and smoothed 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, monthly nondurable consumption 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 (BEA) and the Center for Research on Security Prices (CRSP).[3] 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 number of latent state variables, so it's important to use an estimation approach that takes full advantage of the data. 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, following 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 jointly determine the volatility of the habit contribution to the price-dividend ratio. Thus we chose $\bar{s} = \log 0.06$, 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. Assuming alternative values of $\bar{S}$ also did not have a significant impact on the main results.

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

[3] CRSP data is from Wharton Research Data Services (WRDS). wrds-web.wharton.upenn.edu/wrds/about/databaselist.cfm.

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.[4]

The left half of Table 1 shows the wide range spanned by the priors. Our prior persistence parameters are all very diffuse, uniform between 0 and 1. Similarly, the prior relative volatility of long run growth shocks, φ_x, is uniform between 0 and 1. Other priors are diffuse too. For example, the prior on the mean volatility of consumption shocks is between 0.1% and 4.0% annually.

Priors on price-dividend coefficients are empirically motivated. The bounds allow for the possibility that each state variable can account for all of market volatility (assuming standard consumption growth parameters in the literature). For example, the upper bound on the long run growth coefficient A_x solves

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

where $\varphi_x = 0.038$ and $\overline{\sigma} = 0.0072 \times \sqrt{12}$ as in Bansal, Kiku, and Yaron (2012a), 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 exact magnitude of the bounds is not important under uniform priors, however. As long as the bound exceeds the mass of the likelihood, the posterior is largely independent of the bound.

The signs of the price-dividend coefficients are also restricted to be consistent with theory. That is, we restrict the coefficients on long run growth and surplus consumption to be positive, and we restrict the coefficients on long run

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

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

volatility and to be negative.

2.5. Posterior Parameter Estimates

The right half of Table 1 shows the posterior estimates. Beginning at the top of the table, the posteriors on simple consumption and dividend parameters are standard. The mean volatility of consumption innovations $\overline{\sigma}$ is about 1% per year, and dividend innovations are roughly 6 times as volatile as the consumption innovations.

In the long run risks section, we see that the estimator finds evidence of significant long run risks in real economic 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 risks is an important issue in the literature, Figure 1 takes a closer look and plots the posterior parameter distributions.

The figure shows that the high persistence of long run growth and volatility are estimated rather precisely. The entire distribution of these parameters is above 0.80. Long run risks vary over time, that is, the relative volatility of long run growth shocks φ_x and the volatility of log 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 problem with long run risks and habit: The portion of asset prices that they do not 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 component of price/dividends has an unconditional volatility of above 50%, more than

Table 1: Parameter Estimates

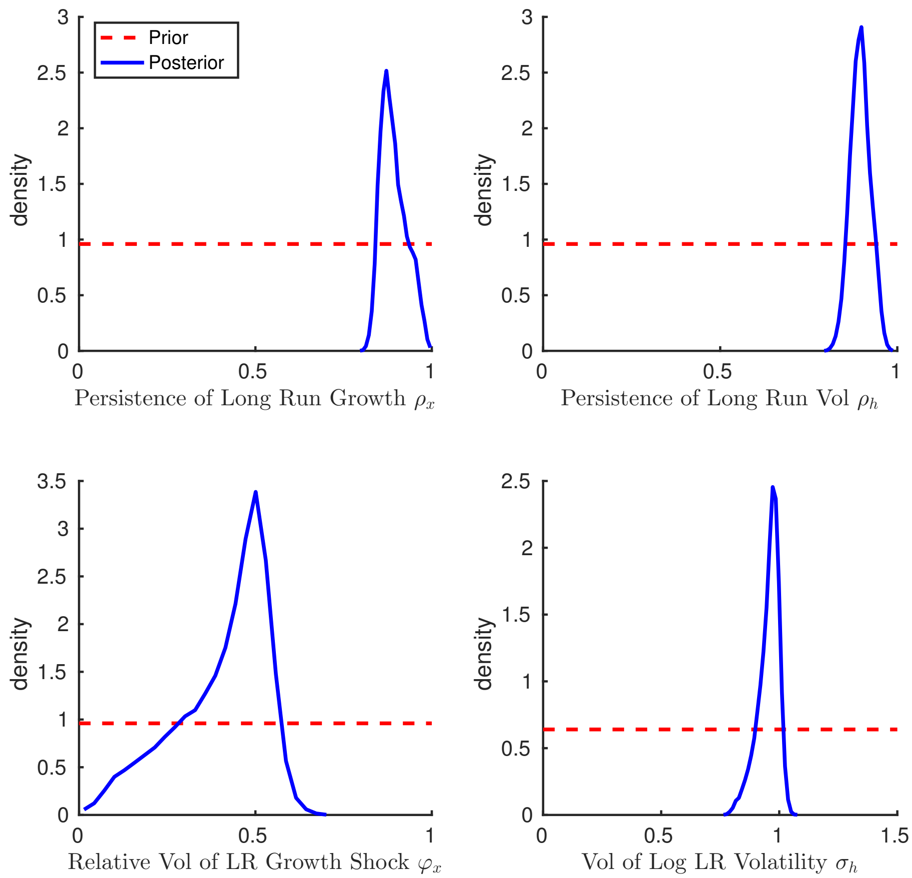
The table shows prior and posterior parameter estimates for the model

$$\begin{aligned}
pd_t &= \mu_{pd} + A_x x_t + A_V \tilde{\sigma}_t^2 + A_s \tilde{s}_t + A_e e_t \\
\Delta c_t &= \mu_c + x_{t-1} + \sigma_{t-1}\eta_{c,t} \\
\Delta d_t &= \mu_d + \phi_x x_{t-1} + \phi_{\eta c}\sigma_{t-1}\eta_{c,t} + \varphi_d \sigma_{t-1}\eta_{d,t} \\
x_t &= \rho_x x_{t-1} + \varphi_x \sqrt{1-\rho_x^2}\sigma_{t-1}\eta_{x,t} \\
h_t &= \rho_h h_{t-1} + \sigma_h \sqrt{1-\rho_h^2}\eta_{h,t}, \quad \sigma_t = \overline{\sigma}\exp(h_t) \\
\tilde{s}_t &= \rho_s \tilde{s}_{t-1} + \lambda(\tilde{s}_{t-1})(\Delta c_t - \mathbb{E}_{t-1}\Delta c_t) \\
e_t &= \rho_e e_{t-1} + \sigma_e \eta_{e,t}
\end{aligned}$$

where pd_t is the log price-dividend ratio, Δc_t is consumption growth, Δd_t is dividend growth, and η_t's are standard normal independent noise. Prior distributions are independent and uniform. Posteriors are computed using annual consumption, dividend, and stock prices from 1929-2014, particle filter, and Metropolis Hastings. $\sigma_e = 1$, $\exp \bar{s} = 0.06$ are chosen outside of the estimation, as are μ_{pd}, μ_c, and μ_d, which are chosen to be their sample means.

Parameter		Prior		Posterior			
		0%	100%	Mean	5%	50%	95%
Simple Consumption and Dividends							
Mean Vol of Cons Shocks	$\overline{\sigma}$	0.001	0.040	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
Rel Vol of Dividend Shocks	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
Rel Vol of LR Growth Shocks	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 Log LR Vol	σ_h	0	1.5	0.954	0.864	0.964	1.01
Habit and Residual							
Persistence of Surplus Cons.	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
Surplus Cons. 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. Other posterior distributions can be found in the Appendix.



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 shrunken posteriors of the other price-dividend coefficients. Since priors were chosen so that each state could account for all market volatility, the posterior coefficients on long run growth, long run volatility, and surplus consumption shrink 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 market volatility is the most important?

Figure 2 provides an answer to this question. It plots the historical path of the price-dividend ratio and decomposes the path into contributions from each state variable (see equation (1)). The states are extracted with a particle smoother (Godsill, Doucet, and West (2004)) using mean posterior parameters from Table 1. The contribution of a state is the mean smoothed state multiplied by its respective mean posterior coefficient (the contribution of long run growth is $A_x x_t$).

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, the boom of the 1950s, the bear market of the 1970s, the epic rise of valuations between 1980 and 2000, and the sharp crash in the early 2000s. 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 throughout the 1950s and 1960s.

Compared to the residual, however, long run risks and habit are relatively unimportant. Indeed, outside of the two major crises, long run volatility has almost no effect. And while the effects of long run growth and habit are more visible, their contributions are economically small and often have the opposite pattern of the movements seen in asset prices. This last result is, perhaps, intuitive: economic growth declined slowly between 1960 and 2014, while the price-dividend ratio has trended upwards overall.

Figure 2: Decomposition of the Historical Log Price-Dividend Ratio. The figure plots the decomposition implied by equation (1)

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

where pd_t is log price/dividends, x_t is long run growth, $\tilde{\sigma}_t$ is long run volatility, $\tilde{s}_t$ is surplus consumption, and e_t is the residual. The states are the mean state found by particle smoother using mean posterior parameters (Table 1). Coefficients are the mean posterior coefficients in Table 1. The residual contribution (yellow) is dominant and closely tracks the price-dividend ratio.

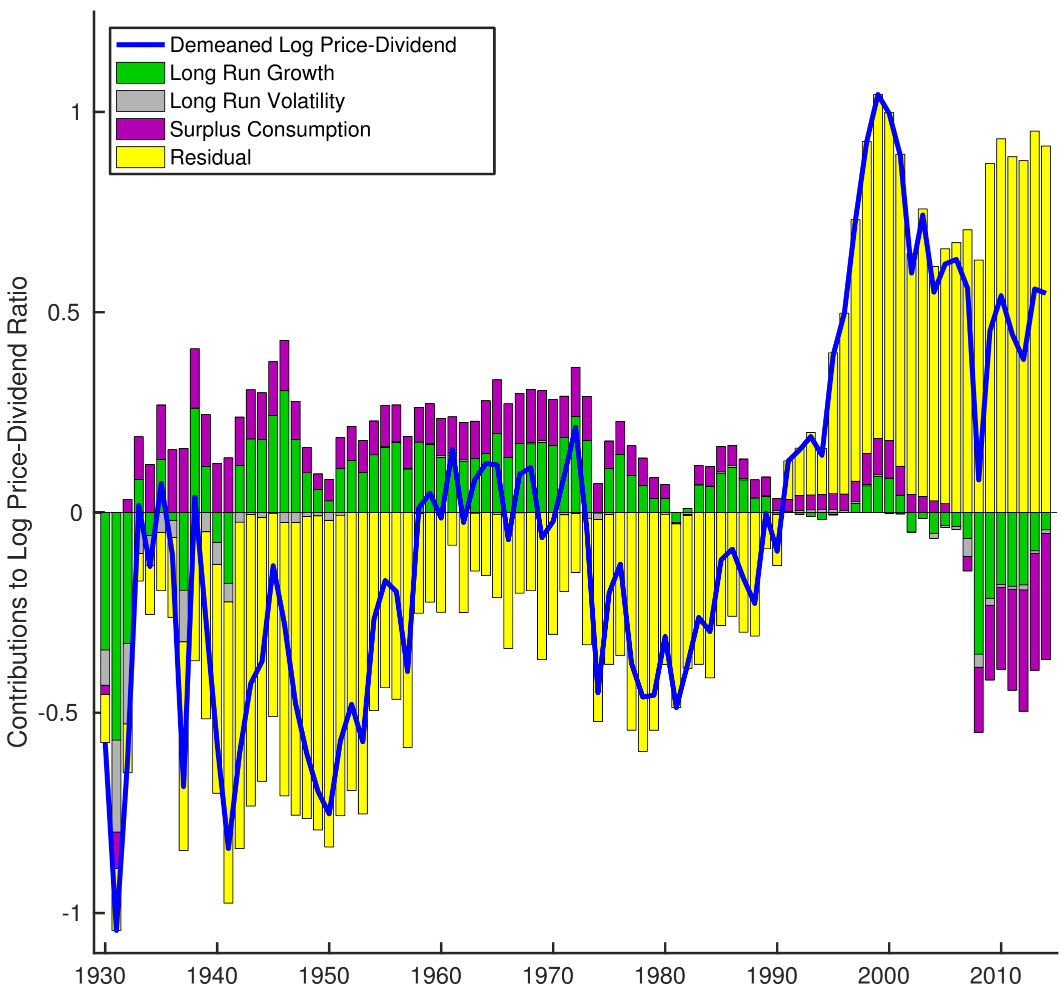


Figure 2 is created using mean posterior parameters, which do not account for estimation uncertainty. Is the dominant role of the residual robust to the dispersion seen in the posterior parameters (Table 1)?

Figure 3 shows that the dominant role of the residual is far too large to be due to estimation uncertainty. The figure plots a variance decomposition of the price-dividend ratio using the entire distribution of posterior parameters. The variance decomposition is calculated by taking the covariance of equation (1) with pd_t

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

which can lead to negative shares if a state variable has a negative sample correlation with pd_t.

The figure shows that the residual's share of price-dividend variance is almost entirely above 75%. Indeed, the residual's mean share is about 95%, showing that something other than long run risks and habit accounts for nearly all of the past 85 years of market volatility.

Long run growth, long run volatility, and habit have very small shares, accounting for the remaining 5% of market variance altogether. These small shares are consistent with the scarcity of crises in our 85 year sample. There is some uncertainty in these estimates. The distributions of the shares for long run growth and volatility cover up to 25% and 15% respectively. Nevertheless these shares seem to be negatively correlated, as the residual share very rarely falls below 75%.

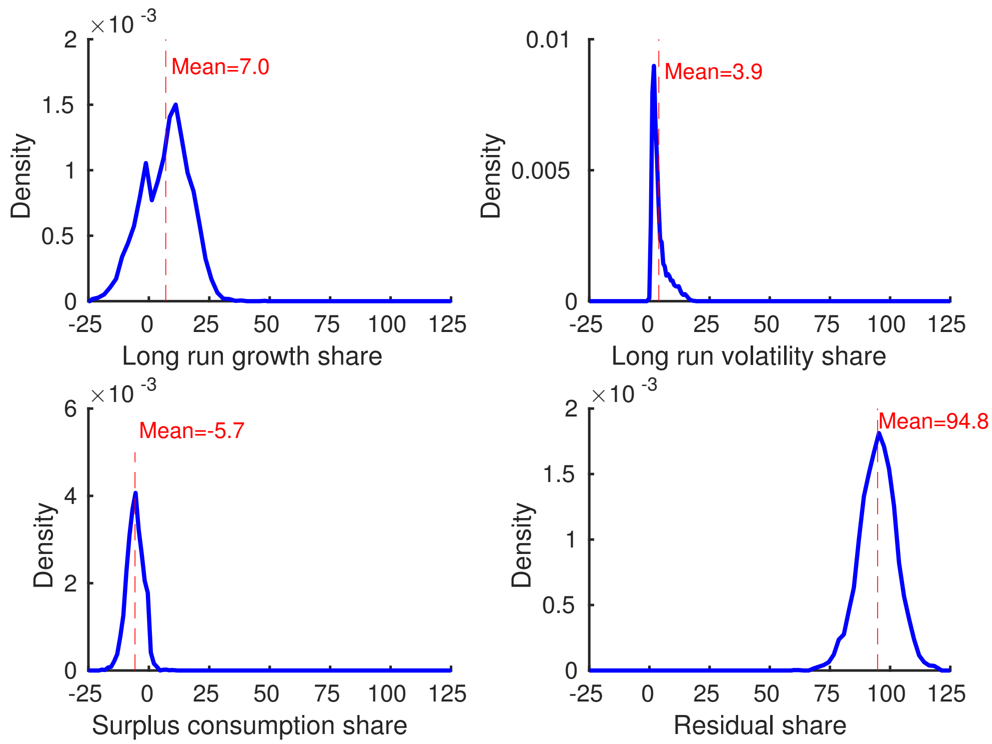
This accounting for market volatility is somewhat complicated by the fact that shares are sometimes negative or above 100%. These negative shares are due to the fact that the variance decomposition is computed using sample covariances (equation (8)), which can be negative. Indeed, economic growth declined over the sample while asset prices grew, which intuitively leads to a negative share. Habit faces a similar in-sample correlation problem. We discuss both of these issues in more detail in Section 3.2.

This section illustrates the main message of the paper: while long run risks and habit play a non-trivial role in asset prices, something else is behind the vast majority of market volatility. This result is robust to estimation uncertainty,

Figure 3: Price/Dividend Variance Decomposition and Estimation Uncertainty. Shares are in percent and implied by the decomposition

$$\begin{aligned}\mathrm{Var}(pd_t) = &\ \mathrm{Cov}(A_x x_t, pd_t) + \mathrm{Cov}\left(A_V \tilde{\sigma}_t^2, pd_t\right) \\ &+ \mathrm{Cov}(A_s \tilde{s}_t, pd_t) + \mathrm{Cov}(A_e e_t, pd_t)\end{aligned}$$

where pd_t is log price/dividends, x_t is long run growth, $\tilde{\sigma}_t$ is long run volatility, $\tilde{s}_t$ is surplus consumption, e_t is the residual, and A's are the coefficients from equation (1). The densities are computed by drawing parameters from the posterior (Table 1), using the draw to find smoothed mean states, and calculating variance contributions according to the above equation. Coefficients are mean posteriors. The plots show that the residual's dominant role 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 present a simple OLS version of our price-dividend decomposition that generates similar results.

The OLS decomposition also helps explain why the residual is so dominant. The large role for the residual comes from the low correlation between asset prices and real economic growth or volatility (past and future).

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 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 smoother.

The top panel of Figure 4 shows the historical path of long run growth along with demeaned consumption growth. The long run growth path does a good job of capturing historical shifts in growth. The path identifies the Great Depression, the relatively booming 60s, as well as the productivity slowdown of the 1970s. Movements in expected growth are small but persistent, consistent with the long run risks story.

The bottom panel of Figure 4 shows the historical path 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 after World War 2, the return of volatility in the 1970s, the Great Moderation, as well as the recent return of volatility in 2008.

Figure 4: Smoothed States and Observables Part 1 of 2. We apply a particle smoother to data on consumption, dividends, and stock prices, using mean posterior parameters (Table 1). Scattered x's plot observables for comparison. The smoothed paths are intuitive and capture historical shifts in growth and volatility.

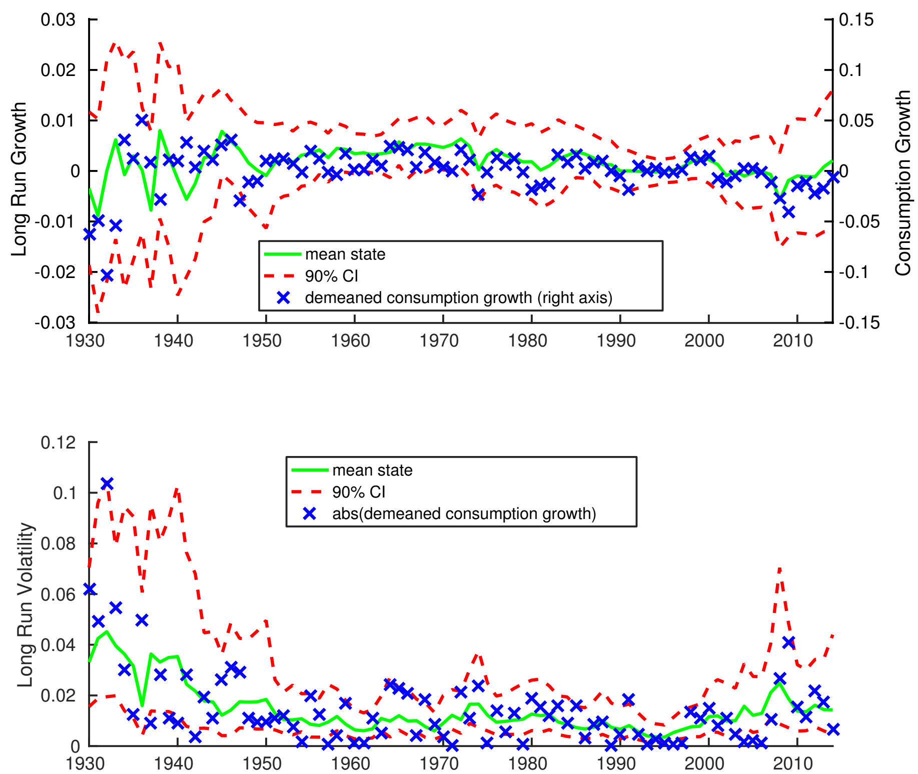


Figure 5: Smoothed States and Observables Part 2 of 2. We apply a particle smoother to data on consumption, dividends and stock prices using mean posterior parameters (Table 1). x's plot observables for comparison. Scaled consumption growth is demeaned consumption growth multiplied by the steady state $\lambda(s_t)$ (see equation (5)). Surplus consumption responds slowly to consumption growth, and the residual closely tracks price-dividends.

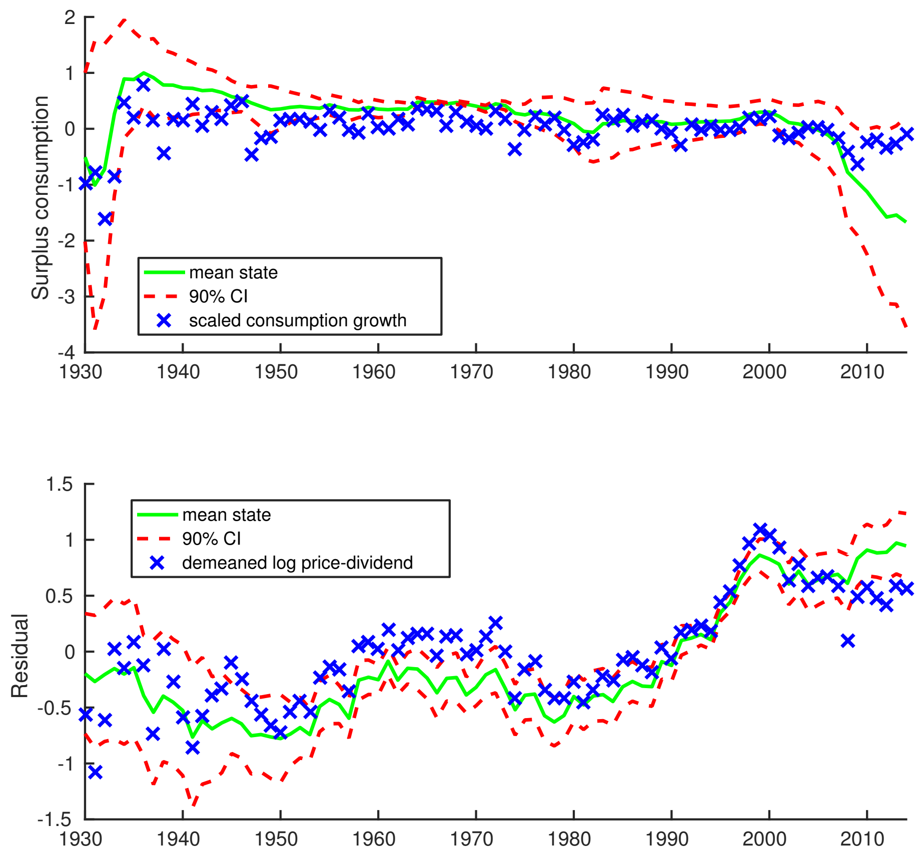


Figure 5 shows the remaining two states: surplus consumption and the residual. The top panel shows the historical path of surplus consumption along with scaled consumption growth. Consumption growth is scaled by subtracting out its mean, and then multiplying by the steady state $\lambda(s_t)$, to imitate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

The historical path of surplus consumption is intuitive. Surplus consumption is very persistent, and responds to slowly to 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 historical path of the residual, along with the log price-dividend ratio. As noted in Section 2.6, these two series track each other closely throughout most of the sample.

3.2. OLS Price-Dividend Decomposition

Using the likelihood makes the estimator comprehensive: It accounts for all moments of the observables. But a likelihood-based estimation is also nontrivial to dissect. Which moments are driving the results?

This section helps address this question by performing a simplified, OLS version of our estimation. We extract state histories using only data on consumption and dividends (excluding asset prices). We then use OLS to regress the price-dividend ratio on the states. The OLS estimates focus on the covariance between the price-dividend ratio and expected consumption growth, consumption volatility, and past consumption growth. These covariances lead OLS to conclude that a residual explains most of market volatility.

When extracting state histories, we skip estimation of parameter values and instead use values from Bansal, Kiku, and Yaron (2012a) and Campbell and Cochrane (1999). This approach makes clear that the estimation of cash flow parameters is not behind our decomposition results. Table 2 shows the parameter values we use. Parameters are converted from the monthly values used in the original papers by simple transformations ([annual persistence] = [monthly persistence]^{12}). The parameters shown also account for our modified volatility process and the functional form of the conditional volatilities (equations (3) and (4)).

Historical paths for long run growth and long run volatility are found by applying a particle smoother to the data on consumption and dividends using parameters from Table 2. The paths are plotted in the top two panels of Figure 6. Long run growth and long run volatility follow paths that are similar to those from the baseline, and generally pick up the same historical features (see Figure 4). These two states deviate from their historical means primarily in the Great Depression and Great Recession. Long run growth is less volatile in the 1930s compared to the baseline, as the smoother no longer ties growth to the violent asset price movements in that era.

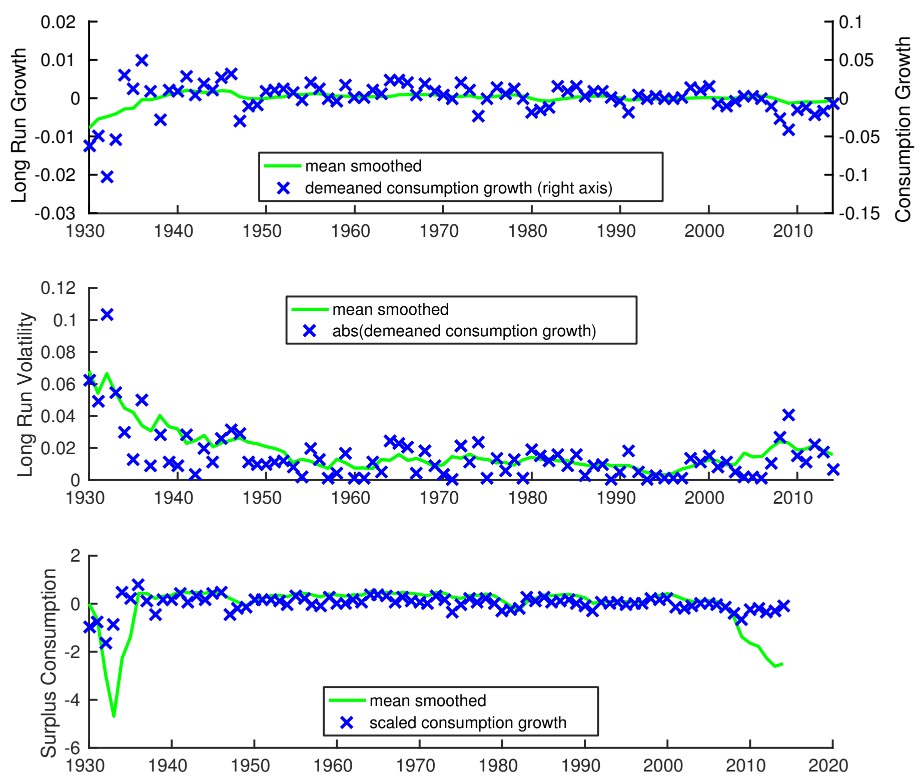
Surplus consumption is constructed by initializing at the Campbell and

Table 2: OLS Decomposition: Cash Flow Parameter Choices

For the OLS decomposition, we apply a particle smoother to consumption and dividend data using parameter values from Bansal, Kiku, and Yaron (2012a) (BKY) and Campbell and Cochrane (1999). The model is the same as the baseline (equations (1)-(6)). The data, model, and parameter values are annual. Parameters are converted from the monthly values used the original papers by applying simple transformations: [annual persistence] = [monthly persistence]^{12}, [annual volatility] = [monthly volatility] $\times\sqrt{12}$. Parameter are also transformed to account for the volatility specification (see footnote on page 7).

		Value	Source
Simple Consumption and Dividends			
Consumption Vol	$\overline{\sigma}$	0.0249	BKY
Div Loading on Cons Shock	$\phi_{\eta c}$	2.60	BKY
Relative Vol of Dividends	φ_d	5.96	BKY
Long Run Risks			
Persistence of LR Growth	ρ_x	0.74	BKY
Relative Vol of LR Growth	φ_x	0.17	BKY
Div Loading on LR Growth	ϕ_x	2.50	BKY
Persistence of LR Vol	ρ_h	0.99	BKY
Volatility of LR Vol	σ_h	2.09	BKY
Habit			
Persistence of Surplus Cons.	ρ_s	0.87	Campbell-Cochrane
Steady State Surplus Cons.	$\bar{s}$	log 0.06	Campbell-Cochrane

Figure 6: Simple State Histories. Long run growth and long run volatility are found by applying a particle smoother to consumption and dividend growth using parameter values from Bansal, Kiku, and Yaron (2012a) (Table 2). Surplus consumption is constructed using consumption growth data, the surplus consumption process (5), and 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 those from the baseline estimation (Figures 4 and 5).



Cochrane (1999) steady state and then applying the surplus consumption process (5), using our long run growth estimate as $\mathbb{E}_{t-1}\Delta c_t$. We assume $\rho_s = 0.87$, as in Campbell and Cochrane (1999). The resulting path (bottom panel of Figure 6) looks similar to the baseline (see Figure 5). Like the long run risk states, surplus consumption experiences large movements in the Great Depression and Great Recession, but otherwise is relatively constant.

The final step of our simplified estimator uses OLS to regress the price-dividend ratio on the extracted state histories. Table 3 shows the resulting coefficients, R^2's, and price-dividend variance decomposition. The table shows two specifications: one which includes all state histories and one which includes only variables that have the "right sign."

Table 3: OLS Estimates and Variance Decomposition

We find smoothed paths for long run growth, volatility, and surplus consumption without using asset prices (Figure 6). We then regress log price-dividends on the state paths. Var. share is the share of price-dividend variance implied by

$$\begin{aligned} \text{Var}(pd_t) = &\text{Cov}(A_x x_t, pd_t) + \text{Cov}\left(A_V \tilde{\sigma}_t^2, pd_t\right) \\ &+ \text{Cov}(A_s \tilde{s}_t, pd_t) + \text{Cov}(A_e e_t, pd_t) \end{aligned}$$

where pd_t is log price/dividends, x_t is long run growth, $\tilde{\sigma}_t$ is long run volatility, $\tilde{s}_t$ is surplus consumption, e_t is the residual, and A's are the coefficients from equation (1). The "Long Run Volatility Only" specification omits variables whose coefficients have the opposite sign of that implied by theory. In both regressions, the R^2 shows that most of the variance of log price-dividends is explained by a residual.

Specification		Long Run Growth	Long Run Volatility	Surplus Consumption	
Including All Variables	Coefficient	-128.9	-460.8	-0.15	
	s.e.	(53.0)	(87.3)	(0.06)	
	R^2 (%)				30.0
	Var. Share (%)	-3.2	29.1	4.0	
Long Run Volatility Only	Coefficient		-190.2		
	s.e.		(54.5)		
	R^2 (%)				12.0
	Var. Share (%)		88.0		

The "Including All Variables" specification shows that long run growth and surplus consumption have the "wrong sign," in that their signs are the opposite of that implied by theory. Unlike the Bayesian estimator which can constrain these coefficients to be positive, OLS ignores any priors and finds that a negative sign is the best fit for the data.

The "Long Run Volatility Only" specification keeps only variables that have the right sign. These results should thus be fairly close to the baseline, which imposes priors consistent with theory. Indeed, the coefficient on long run volatility of -190 is within two standard errors of the -88 value from the full estimation (Table 1).

The "Long Run Volatility Only" specification produces an R^2 of 12%, indicating that the residual accounts for the vast majority of the variance of price-dividends. By construction, OLS minimizes the role of the residual, leading to a slightly smaller residual share of price-dividend variance (88%) compared to our baseline estimate of 95%. The higher R^2 of 30% found in the "Including All Variables" specification still means that the residual accounts for 70% of market volatility.

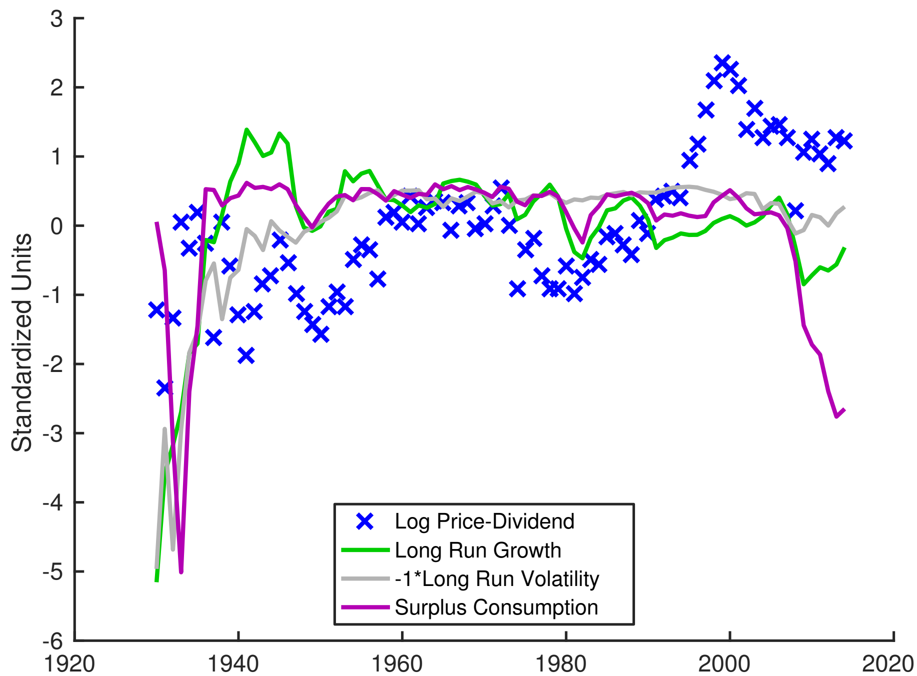
Figure 7 illustrates why the residual plays such a large role in the OLS estimates. The figure plots state paths along with the price-dividend ratio, with all variables standardized. Long run growth, long run volatility, and surplus consumption comove with the price-dividend ratio primarily in crisis periods (Great Depression and 2008 Financial Crisis). But they fail to capture most of the other patterns in asset price history. Indeed, a crucial feature of the price-dividend history is its long upward trend over the past 85 years, something that is absent from all three state variables.

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 remove long run risks from the model, (3) if we remove habit from the model, (4)

Figure 7: Simple States Histories vs the Price-Dividend Ratio. All variables are scaled to have zero mean and a standard deviation of 1. Long run growth and long run volatility are found by applying a particle smoother to consumption and dividend growth using parameter values from Bansal, Kiku, and Yaron (2012a) (Table 2). Surplus consumption is constructed using consumption growth data, the surplus consumption process (5), and parameter values from Campbell and Cochrane (1999). All states move in crises but otherwise have little correlation with price-dividends.



if we alter the prior correlation structure, (5) if we specify that habit responds to consumption growth rather than innovations, (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 4 summarizes the robustness results. The table shows the shares of variance (equation 8) accounted for by long run growth, long run volatility, habit, and the residual across the six model and prior specifications listed above. Under all six specifications, the residual accounts for the vast majority of market volatility, with a minimum share of 80%.

Table 4: P/D Variance Shares in Alternative Model Specifications

Figures show percent contributions to the variance of the log price-dividend ratio implied by

$$\begin{aligned}\text{Var}(pd_t) = &\text{Cov}(A_x x_t, pd_t) + \text{Cov}\left(A_V \tilde{\sigma}_t^2, pd_t\right)\\ &+ \text{Cov}(A_s \tilde{s}_t, pd_t) + \text{Cov}(A_e e_t, pd_t)\end{aligned}$$

where pd_t is log price/dividends, x_t is long run growth, $\tilde{\sigma}_t$ is long run volatility, $\tilde{s}_t$ is surplus consumption, e_t is the residual, and A's are the coefficients from equation (1). The table shows the mean and standard deviation (in parentheses) of the posterior distribution of the shares. Shares are computed using the smoothed states evaluated at 5,000 draws from the posterior parameter distribution. Regardless of the specification, the residual is the dominant source of market volatility.

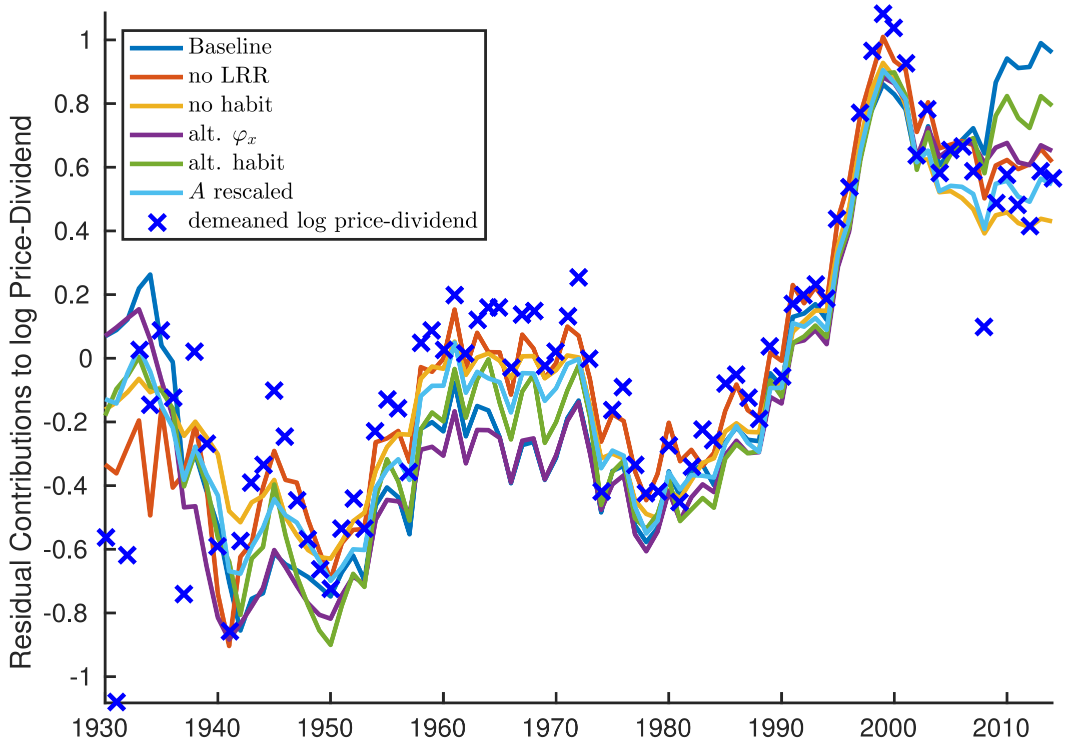
Variance share	(1) Baseline	(2) no LRR	(3) no habit	(4) alt. φ_x	(5) alt. habit	(6) A rescaled
Long-run Growth	6.99		15.27	-1.59	-2.61	14.66
	(9.83)		(3.62)	(8.21)	(7.61)	(3.26)
Long-run Volatility	3.90		4.49	14.09	4.13	1.95
	(3.28)		(4.04)	(6.35)	(2.84)	(2.62)
Surplus Consumption	-5.72	6.81		0.24	-2.30	-0.86
	(3.49)	(4.26)		(2.28)	(3.52)	(1.62)
Residual	94.83	93.19	80.24	87.25	100.77	84.25
	(7.71)	(4.26)	(4.38)	(9.21)	(7.09)	(2.85)

Figure 8 plots the residual under these six model specifications. Regardless of the specification, the residual is very highly correlated with the price-dividend ratio and marks most of the key events in stock market history. Indeed, the role of the residual is very consistent: regardless of the specification,

Figure 8: Residual Contributions to the Price-Dividend Ratio Under Alternative Specifications. Lines show $A_e e_t$ from

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

where pd_t is log price/dividends, x_t is long run growth, $\tilde{\sigma}_t$ is long run volatility, $\tilde{s}_t$ is surplus consumption, and e_t is the residual. Each line shows a different model specification (Table 4). e_t is the smoothed mean residual computed at the mean posterior parameters, and A_e is the mean posterior. x's show the demeaned log price-dividend ratio for comparison. Regardless of the specification, the residual tracks recognizable features of the price-dividend ratio outside of the Great Depression and 2008 Financial Crisis.



the residual drives the stock market outside of the Great Depression and Great Recession.

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

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 4, Column (2)) and no habit (Column (3)) specifications examine this concern. Column (2) removes habit and Column (3) removes the long-run risk factors from the price-dividend ratio equation (1).

Table 4 shows that, when long-run risks are removed, the estimation assigns a variance share of about 9% to habit. Similarly, when habit is removed, long-run risks accounts for a variance share of about 15%. Under both specification, the residual plays a dominant role, making it clear that competition between the different macro-asset pricing factors is not driving our results.

Our baseline assumes that the prior parameters are independent, but this implies that other properties of the model are correlated. The alt. φ_x specification (Column (3) of Table 4) 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: $\varphi_x \sim \mathscr{U}([0, 0.2])$. Table 4 shows that 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 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 4 Column (5)) 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). Table 4 shows that, under this alternative, the residual takes up even more of the variation in the price-dividend ratio than in the baseline.

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 4, 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:

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

Here, $\mathbb{V}[x_t \mid \theta]$ 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 i.i.d. 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.

Table 4's 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

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

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

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

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 a process 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. There are two kinds of complications which are consistent with our results and lead to additional observables: (1) incomplete markets, and (2) imperfectly rational agents. The complexity of these models, however, makes direct model evaluation difficult.

Under incomplete markets, 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 simplest way to model incomplete markets is by introducing idiosyncratic income risk (for example, Constantinides and Duffie (1996)). Constantinides and Ghosh (2017) use GMM to estimate a model with idiosyncratic risk, but they do not use the correlation between asset prices and idiosyncratic risk in their estimation. Schmidt (2015) argues that initial claims for unemployment is a reasonable proxy for idiosyncratic risk, 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 brings in survey data. Their model replicates the positive correlation between survey expectations of returns and the price-dividend ratio. 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 only in crisis periods.

This analysis raises the bar for asset pricing models. Many macro finance models that 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 are 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 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 who 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 consumption-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). In the end, we have transition equations

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

and observation equations

$$y_t = \mu_y + Z m_t + Z_v(\exp(2h_t) - \exp(2\sigma_h^2)) \tag{16}$$

where m_t is a vector of mean "states," y_t is a vector of observables, w_t, η_t are vectors of standard normal independent noise, $\Phi(m_{t-1}), \Sigma_s(m_{t-1})$ are matricies that describe the evolution of the mean "states," and μ_y, Z, Z_v are vectors and matricies that map states to observables. We put quotes around "states" because the elements of m_t include terms which are not state variables in the traditional economic sense. These additional "states" help simplify notation.

Equations (15) and (16) are convenient forms for the asset pricing models with time varying volatility and normal shocks. As this class of models is conditionally normal, it helps to express the model as close to a state space form as possible. Moreover, this formulation allows the model to be extended to account for mixed frequency data.

These equations are mapped to the model in Section 2 by a careful definition of vectors and matricies. We'll now define these vectors and matricies.

Observables and States Equations (1) and (2) can be mapped into the observation equation (16) as follows:

$$\underbrace{\begin{bmatrix}\Delta c_t\\ \Delta d_t\\ pd_t\end{bmatrix}}_{y_t}=\underbrace{\begin{bmatrix}\mu_c\\ \mu_d\\ \mu_{pd}\end{bmatrix}}_{\mu_y}+\underbrace{\begin{bmatrix}0 & 1 & 1 & 0\\ 0 & \phi_x & \phi_{\eta c} & 1\\ [A_x, A_s, A_e] & 0 & 0 & 0\end{bmatrix}}_{Z}\underbrace{\begin{bmatrix}z_t\\ x_{t-1}\\ \tilde{\eta}_{c,t}\\ \tilde{\eta}_{d,t}\end{bmatrix}}_{m_t}$$
$$+\underbrace{\begin{bmatrix}0\\ 0\\ A_V\overline{\sigma}^2\end{bmatrix}}_{Z_v}(\exp(2h_t)-\exp(2\sigma_h^2)).$$

where

$$\begin{aligned}z_t &\equiv [x_t, \tilde{s}_t, e_t]'\\ \tilde{\eta}_{c,t} &= \overline{\sigma}\exp(h_{t-1})\eta_{c,t}\\ \tilde{\eta}_{d,t} &= \varphi_d\overline{\sigma}\exp(h_{t-1})\eta_{d,t}\end{aligned}$$

and $'$ indicates a transpose.

State Transition Then state transitions (3)-(6) can be expressed in terms of the augmented state space transitions (15) as follows:

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

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

where ψ_x is an indicator variable which depends on the habit specification ($\psi_x = 0 \Rightarrow$ habit responds to innovations in consumption growth).

7.2. Particle Filter Details

With the state space formulation in hand, we can now write down the particle filter algorithm in a compact form. We first describe the big picture of the algorithm. We then go on to give the details of how each distribution is defined.

For each $t = 1, ..., T$, do the following for particles $i = 1, ..., M$.

1. Begin with a set of particles $[m_{t-1}^i, h_{t-1}^i]$ and weights π_{t-1}^i.
2. Draw $h_t^i \sim q(h_t^i | pd_t^i, h_{t-1}^i, m_{t-1}^i)$ for each i, where $q(h_t^i | pd_t^i, h_{t-1}^i, m_{t-1}^i)$ is a proposal distribution which we'll describe shortly.
3. Draw $m_t^i \sim p(m_t | y_t, h_t^i, h_{t-1}^i, m_{t-1}^i)$ for each i, where $p(m_t | y_t, h_t^i, h_{t-1}^i, m_{t-1}^i)$ is the conditional density of m_t. We'll describe how this density is computed shortly. Throughout this Appendix $p(x|y)$ means the conditional density of x given y.

4. Update particle weights using

$$\pi_t^i = \pi_{t-1}^i [\text{update factor}]^i \tag{17}$$

$$[\text{update factor}]^i = p(y_t | h_t^i, h_{t-1}^i, m_{t-1}^i) \left[\frac{p(h_t^i | h_{t-1}^i)}{q(h_t^i | pd_t, h_{t-1}^i, m_{t-1}^i)} \right] \tag{18}$$

In a simple bootstrap particle filter, the update factor is just the likelihood of y_t given m_t^i and h_t^i. We'll explain how to derive the above update factor shortly.

5. Estimate log-likelihood contribution

$$\log \hat{p}(y_t) = \log \left(\sum_i \pi_{t-1}^i [\text{update factor}]^i \right) \tag{19}$$

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

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 $h_t \sim q(h_t | pd_t, h_{t-1}, m_{t-1})$

We draw h_t based off of pd_t and the previous state (h_{t-1}, m_{t-1}). The basic idea is that we want to draw h_t as close to the true probability $p(h_t | pd_t, h_{t-1}, m_{t-1})$ as possible in order to minimize Monte Carlo noise in the particle filter. Unfortunately, the relationship between pd_t and h_t is nonlinear (equations (1) and (4)). We can, however, use the following Taylor expansion

$$\exp(2h_t) \approx \exp(2\rho_h h_{t-1})(1 - 2\rho_h h_{t-1}) + 2\exp(2\rho_h h_{t-1}) h_t \tag{20}$$

which leads to an approximation (16) that is linear in h_t

$$\begin{aligned} y_t &= \mu_y + Z m_t + Z_v(\exp(2h_t) - \exp(2\sigma_h^2)) \\ &\approx \mu_y + Z m_t + Z_v \exp(2\rho_h h_{t-1})(1 - 2\rho_h h_{t-1}) + Z_v 2\exp(2\rho_h h_{t-1}) h_t. \end{aligned} \tag{21}$$

This approximation, combined with the equation (15) and the definition of y_t lets us write a mini state space system

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

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

$$\begin{aligned} pd_0 &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} [\mu_y + Z\Phi(m_{t-1})m_{t-1} + Z_v \exp(2\rho_h h_{t-1})(1-2\rho_h h_{t-1})] \\ \tilde{A}_V &= A_V \bar{\sigma}^2 2 \exp(2\rho_h h_{t-1}) \\ \tilde{\sigma}_{pd} &= \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} Z\Sigma_s(m_{t-1}). \end{aligned}$$

In this approximation, $h_t | pd_t, h_{t-1}, m_{t-1}$ is normally distributed, and a one-step Kalman filter gives the mean and variance. We use this distribution as $q(h_t | pd_t, h_{t-1}, m_{t-1})$.

In principle, this approximation could be used to generate a proposal distribution for all states $[h_t, s_t]$. But drawing h_t separately lets us nest Schorfheide, Song, and Yaron (2016)'s specification and helps error checking.

7.2.2. Proposal Distribution for $m_t \sim p(m_t | y_t, h_t, h_{t-1}, m_{t-1})$

We draw m_t in a similar way to h_t. The only difference is we draw m_t given h_t, and thus y_t is linear in the unobserved m_t and so no approximations are needed.

Explicitly, given h_t, h_{t-1}, the state space system

$$y_t = \mu_y + Z m_t + Z_v(\exp(2h_t) - \exp(2\sigma_h^2)) \tag{23}$$
$$m_t = \Phi(m_{t-1}) m_{t-1} + \Sigma_s(m_{t-1}) \eta_t$$

shows that $p(m_t | y_t, h_t, h_{t-1}, m_{t-1})$ is normally distributed, and a one-step Kalman filter gives the mean and variance of this distribution. We use this distribution to draw m_t in the particle filter.

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

$$\pi_t = \pi_{t-1} \text{update weight} \tag{24}$$

where the update weight is

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

and q is the proposal distribution in the propogation step (see Herbst and Schorfheide (2014)). In our case, $p(m_t, h_t | m_{t-1}, h_{t-1})$ and $q(m_t, h_t | m_{t-1}, h_{t-1}, y_t)$ can be broken up into mean and volatility components

$$\text{update weight} = p(y_t | m_t, h_t) \frac{p(m_t | m_{t-1}, h_{t-1})}{p(m_t | y_t, h_t, m_{t-1}, h_{t-1})} \frac{p(h_t | h_{t-1})}{q(h_t | y_t, m_{t-1}, h_{t-1})} \tag{25}$$

The above expression can be further simplified. First note that (m_t, h_t) are sufficient to determine the density of y_t. Similarly, h_t adds no information regarding m_t given (m_{t-1}, h_{t-1}). Thus,

$$\text{update weight} = \left[p(y_t | m_t, h_t, m_{t-1}, h_{t-1}) \frac{p(m_t | h_t, m_{t-1}, h_{t-1})}{p(m_t | y_t, h_t, m_{t-1}, h_{t-1})} \right] \frac{p(h_t | h_{t-1})}{q(h_t | y_t, m_{t-1}, h_{t-1})} \tag{26}$$

The term in the brackets can be simplified using Bayes' rule twice

$$\begin{aligned}
p(y_t | m_t, h_t, m_{t-1}, h_{t-1}) & \frac{p(m_t | h_t, m_{t-1}, h_{t-1})}{p(m_t | y_t, h_t, m_{t-1}, h_{t-1})} \\
& = p(m_t | y_t, h_t, m_{t-1}, h_{t-1}) \frac{p(y_t, h_t, m_{t-1}, h_{t-1})}{p(m_t, h_t, m_{t-1}, h_{t-1})} \frac{p(m_t | h_t, m_{t-1}, h_{t-1})}{p(m_t | y_t, h_t, m_{t-1}, h_{t-1})} \\
& = \frac{p(y_t, h_t, m_{t-1}, h_{t-1})}{p(m_t, h_t, m_{t-1}, h_{t-1})} p(m_t | h_t, m_{t-1}, h_{t-1}) \\
& = \frac{p(y_t, h_t, m_{t-1}, h_{t-1})}{p(m_t, h_t, m_{t-1}, h_{t-1})} \frac{p(m_t, h_t, m_{t-1}, h_{t-1})}{p(h_t, m_{t-1}, h_{t-1})} \\
& = p(y_t | h_t, m_{t-1}, h_{t-1}).
\end{aligned} \tag{27}$$

Finally, combining equations (26) and (27) gives the update factor expression in the particle filter algorithm (18).

7.3. Particle smoother details

We use a variant of the backward-simulation particle smoother of Godsill, Doucet, and West (2004). Our procedure explicitly takes care of the possibility that the transitional likelihood is degenerate whenever the habit innovation $\lambda(u_{t-1}) = 0$ is zero.

We start with a set of filtered particles $\left(m_t^i, h_t^i\right)$ with weights π_t^i, $t = 1, \ldots, T$, $i = 1, \ldots, M$. We then compute a set of smoothed particles $\left(\tilde{m}_t^i, \tilde{h}_t^i\right)$ as follows.

- Draw k_{iT} from the distribution π_t. Set $\left(\tilde{m}_T^i, \tilde{h}_T^i\right) = \left(m_T^{k_{iT}}, h_T^{k_{iT}}\right)$.
- For $t = T-1 \ldots 1$:
  1. Check whether $\lambda\left(u_t^{k_{it+1}}\right) = 0$:
     - If it is, then the filtered particle drawn in $t+1$ came from a degenerate distribution, and so

       $$\tilde{\pi}_t^k = p\left(m_t^k, h_t^k \mid \tilde{m}_{t+1}^i, \tilde{h}_{t+1}^i, y_{1:T}^o\right) = I(i = k)$$

       where $I()$ is an indicator function. Set $k_{it} = k_{it+1}$.
     - If not, then

       $$\tilde{\pi}_t^k = p\left(m_t^k, h_t^k \mid \tilde{m}_{t+1}^i, \tilde{h}_{t+1}^i, y_{1:T}^o\right) \sim \pi_t^j p\left(\tilde{m}_{t+1}^i, \tilde{h}_{t+1}^i \mid m_t^k, h_t^k\right)$$

       and these densities are finite. Draw k_{it} from the distribution $\tilde{\pi}_t$.
  2. Set $\left(\tilde{m}_t^i, \tilde{h}_t^i\right) = \left(m_t^{k_{it}}, h_t^{k_{it}}\right)$.

In particular, whenever $\lambda\left(u_t^{k_{it+1}}\right) > 0$ we have that $p\left(\tilde{m}_{t+1}^i, \tilde{h}_{t+1}^i \mid m_t^k, h_t^k\right) = 0$ for all k with $\lambda\left(u_t^k\right) = 0$.

7.4. 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 a local maximum of the likelihood function using numerical optimization. We then run a chain of length 5,000 with a symmetric step direction and use the variance of the posterior as the step direction in the next steps. We also choose the step size such that the acceptance rate is a little larger than 0.3. The final MCMC chain has length 500,000.

7.5. Baseline Posterior Destils

These figures show the distributions of the all posterior parameters in baseline estimation (Table 1).

Figure 9: Baseline: Posterior Details 1 of 2.

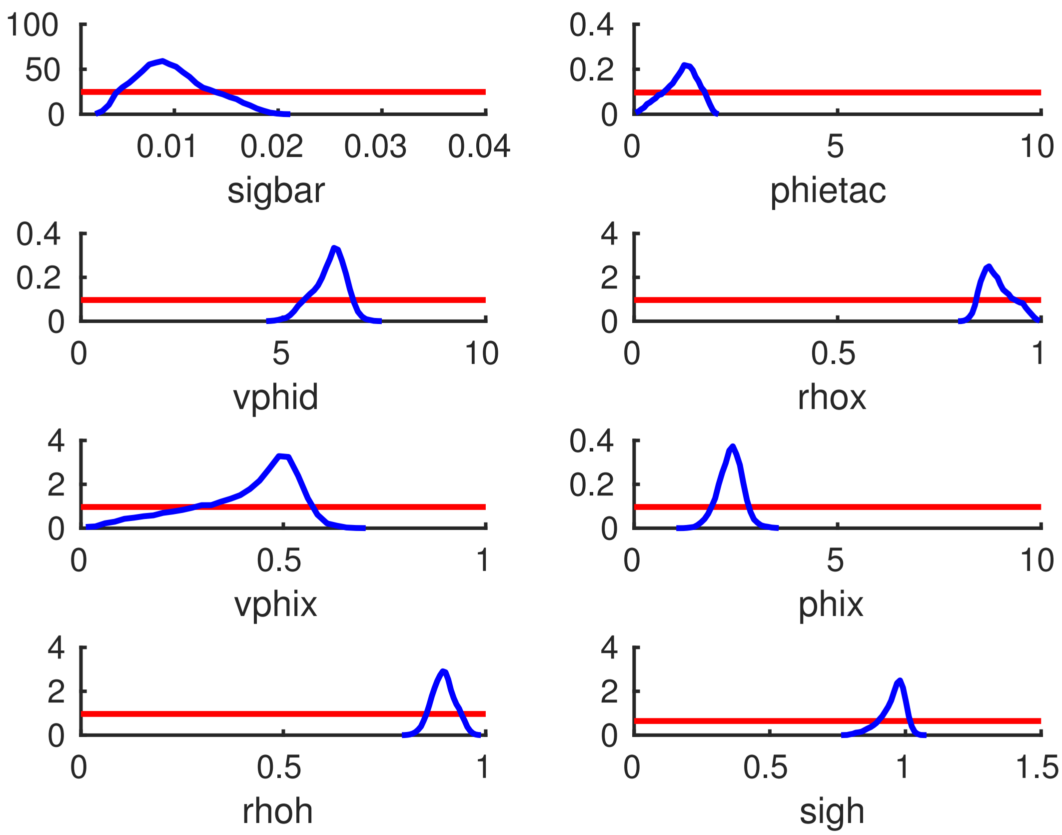
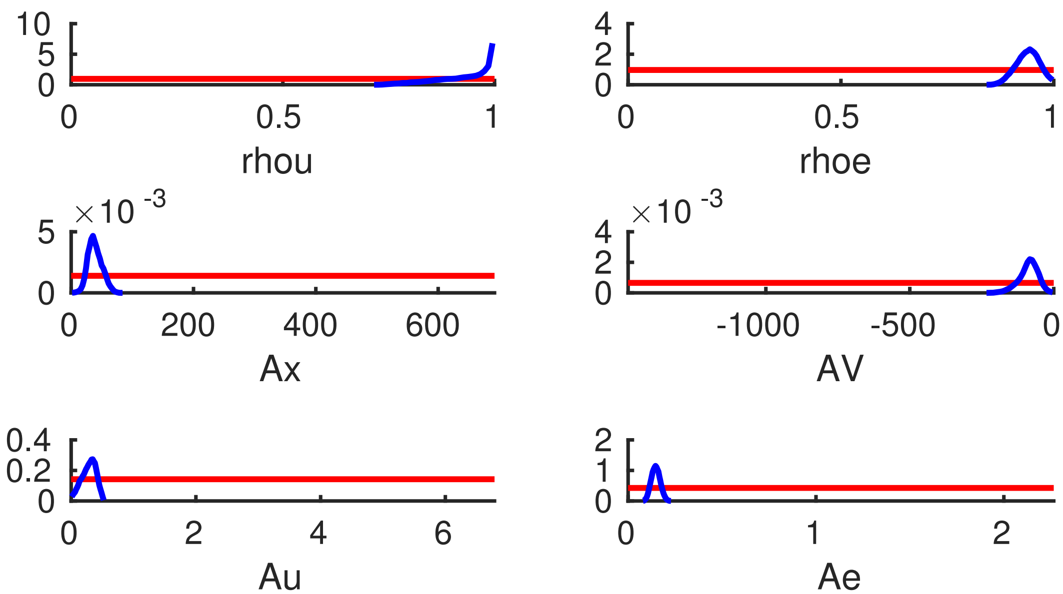


Figure 10: Baseline: Posterior Details 2 of 2.



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. 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.

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. "Consumption-based asset pricing". *Handbook of the Economics of Finance* 1 (2003), pp. 803–887.

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.

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.

Constantinides, George M and Anisha Ghosh. "Asset Pricing with Countercyclical Household Consumption Risk". *The Journal of Finance* (2017).

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.

Godsill, Simon J, Arnaud Doucet, and Mike West. "Monte Carlo smoothing for nonlinear time series". *Journal of the american statistical association* 99.465 (2004), pp. 156–168.

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. 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* 20.1 (2016), pp. 403–445.