*Asset Return Dynamics under Habits and Bad-environment-Good Environment Fundamentals**

Geert Bekaert

Columbia University and NBER

Eric Engstrom

Federal Reserve Board of Governors†

March 15, 2013

*We appreciate comments from seminar participants at the Massachusets Institute of Technology, the Wharton School, Northwestern University, New York University, and over 20 others universities and conferences. We have also benefitted tremendously from the comments of three referees and the editor, Monika Piazzesi. All errors are the sole responsibility of the authors.

†The views expressed in this document do not necessarily reflect those of the Federal Reserve System, its Board of Governors, or staff.

Abstract

We introduce a "bad environment-good environment" (BEGE) technology for consumption growth in a consumption-based asset pricing model with external habit formation. The model generates realistic non-Gaussian features of fundamentals, and fits standard salient features of asset prices including the means and volatilities of equity returns and risk free rates. BEGE dynamics are essential for the model to generate realistic features of the "risk-neutral" conditional density of equity returns, including the variance premium.

I Introduction

A number of consumption-based asset pricing models have emerged that can claim some empirical success in matching salient features of asset return data. For example, Campbell and Cochrane (1999, CC henceforth) develop an external habit framework in which counter-cyclical risk aversion is the essential driver of asset return dynamics. CC keep the exogenous process for consumption growth deliberately simple and Gaussian, yet the model fits the equity premium, the low risk free rate, the variability of equity returns and dividend yields, and long horizon return predictability. In the CC model, habit is external but moves slowly and nonlinearly in response to consumption shocks, yielding a theoretically appealing and empirically successful link between macro-economic events and asset prices.

In this article, we propose a stochastic process for consumption growth that is non-Gaussian, following what we call a "Bad Environment – Good Environment" framework, "BEGE" for short. The consumption growth process receives two types of shocks, both drawn from potentially fat-tailed, skewed distributions. While one shock has positive skewness, the other shock has negative skewness. Because the relative importance of these shocks may vary through time, there are "good times" where the good distribution dominates, and "bad times" where the bad distribution dominates and recession risk is high. An implication of the framework is that even during bad times, large good shocks can occur, and vice versa. We demonstrate that post-war U.S. consumption growth data indeed exhibits the kind of non-Gaussian behavior that is accommodated by the BEGE framework. One particularly striking feature of the post-war US data is that the distribution of consumption growth one quarter after a consumption slowdown has a significantly more negative left tail than after normal periods. It is such non-Gaussian recession dynamics that a BEGE model can fit quite naturally.

Economically, the BEGE model creates a riskier consumption growth environment, which, in equilibrium, leads to a large equity premium and substantial precautionary savings demands, keeping risk free rates low. The non-Gaussian dynamics in consumption growth

are critical for the model to confront option price dynamics. A particularly powerful empirical feature of the data is the so-called variance premium, which is the difference between the "risk neutral" expected conditional variance of the stock market index and the actual expected variance under the physical probability measure. The CBOE's VIX index essentially provides direct readings on the risk-neutral variance as detailed by Bollerslev, Gibson and Zhou (2011) and Carr and Wu (2008). The variance premium is reliably positive and changes considerably through time. Other statistically "strong" stylized facts about the risk neutral conditional distribution of returns include time-varying (but generally negative) skewness, and time-varying, but reliably fatter-than-Gaussian tails (see, for instance, Figlewski (2008)). To generate these features of the risk-neutral distribution, structural models must endogenously generate time-varying non-Gaussianity in returns, which many existing models fail to do. We show that the classic CC model can match important properties of volatility and option price dynamics when appended with a BEGE shock structure for consumption growth.

The remainder of the article is organized as follows. Section II introduces the model and discusses, inter alia, solutions for the risk free rate and the variance premium. Section III confronts the BEGE model with actual consumption data, showing that it matches time-varying non-linearities in the consumption growth process. Section IV discusses our parameter calibration and the fit of the model with respect to asset prices. The final section offers some concluding remarks.

II The Bad Environment-Good Environment (BEGE) Model

In this section, we formally introduce the representative agent model. We begin with a discussion of the assumed data generating process for fundamentals, and then describe preferences.

II.A The BEGE distribution for fundamentals

Our model for consumption growth is given by the following equation:

$$\Delta c_{t+1} = \overline{g} + \sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1} \tag{1}$$

where $\Delta c_t = \ln(C_t) - \ln(C_{t-1})$ is the logarithmic change in real, per-capita consumption, and $\overline{g}$ is the mean rate of consumption growth, which we assume is constant. The final two terms reflect non-Gaussian innovations. The parameters σ_{cp} and σ_{cn} are both positive. The shocks, $\omega_{p,t+1}$ and $\omega_{n,t+1}$, are zero-mean, independent innovations following the centered gamma distribution,[1]

$$\begin{aligned} \omega_{p,t+1} &\sim \widetilde{\Gamma}(p,1) \\ \omega_{n,t+1} &\sim \widetilde{\Gamma}(n,1) \end{aligned} \tag{2}$$

The shape parameters, p and n may be static or vary through time according to a stochastic process, a feature that we shall introduce below. These parameters govern the higher-order moments of Δc_t. Specifically, p governs the width of the positive tail, and n governs the width of the negative tail.

To illustrate what this model implies for the conditional moments of Δc_{t+1}, we calculate the conditional moment generating function (MGF) of Δc_{t+1}. For a scalar, v,

$$\begin{aligned} MGF_v(\Delta c_{t+1}) &\equiv E_t[\exp(v\Delta c_{t+1})] \\ &= \exp(v(\overline{g}) - pf(v\sigma_{cp}) - nf(-v\sigma_{cn})) \end{aligned} \tag{3}$$

where, for notational simplicity, we have defined the function,

$$f(x) = x + \ln(1-x).$$

Note that $f(x)$ is always negative as long as $x < 1$, a technical condition that is always met

[1] The centered gamma distribution with shape parameter k and scale parameter θ, which we denote $\widetilde{\Gamma}(k,\theta)$, has probability density function,

$$\phi(x) = \frac{1}{\Gamma(k)\theta^k}(x+k\theta)^{k-1}\exp\left(-\frac{1}{\theta}(x+k\theta)\right)$$

for $x > -k\theta$, and $\Gamma(\cdot)$ representing the gamma function. Unlike the standard gamma distribution, this distribution has mean zero.

in our applications. This follows directly from the MGF of the gamma distribution and the fact that $\omega_{p,t+1}$ and $\omega_{n,t+1}$ are independent.[2] Using sequential derivatives of the MGF evaluated at $\nu = 0$ yields the first few conditional uncentered moments of Δc_{t+1}. This leads to the centered moments:

$$\begin{aligned} E_t\left[(\Delta c_{t+1} - \overline{g})^2\right] &= \sigma_{cp}^2 p + \sigma_{cn}^2 n \equiv vc \\ E_t\left[(\Delta c_{t+1} - \overline{g})^3\right] &= 2\sigma_{cp}^3 p - 2\sigma_{cn}^3 n \equiv sc \\ E_t\left[(\Delta c_{t+1} - \overline{g})^4\right] - 3vc^2 &= 6\sigma_{cp}^4 p + 6\sigma_{cn}^4 n \equiv kc \end{aligned} \tag{4}$$

The top line of Equation (4) shows that both p and n contribute positively to the conditional variance of consumption, defined as vc. They differ, however, in their implications for the conditional skewness of consumption growth. As can be seen in the expression for the centered third moment, sc, skewness is positive when p is relatively large, and negative when n is large. This is the essence of the BEGE model: the bad environment refers to an environment in which the $\omega_{n,t}$ shocks dominate leading to negative skewness; in the good environment the $\omega_{p,t}$ shocks dominate. Of course, in both environments shocks are zero on average, but there is a higher probability of large positive shocks in a "good environment" and vice versa. Whether good or bad shocks dominate depends on the relative values of p and n, and the sensitivity of consumption growth to both shocks. Finally, the third line of the equation is the excess centered fourth moment, kc, which is increasing in both p and n. Note that there is a linear dependence among higher order moments of Δc_t, all of which are linear in p and n.

Figure 1 plots four examples of BEGE densities under various combinations for p, n σ_{cp}, and σ_{cn}. For ease of comparison of the higher order moments, the mean and variance of all the distributions are the same and $\sigma_{cp} = \sigma_{cn}$. The black line plots the density under large, equal values for p and n. This distribution very closely approximates the Gaussian distribution. The red line plots a BEGE density with smaller, but still equal values for p

[2] To see this, note that for $x \sim \Gamma(k, 1)$, $E[\exp(mx)] = \exp(-k\ln(1-m))$, and for independent random variables, x_1 and x_2, $E[\exp(m(x_1 - x_2))] = E[\exp(mx_1)]\,E[\exp(-mx_2)]$.

and n. This density is more peaked and has fatter tails than the Gaussian distribution. The blue line plots a BEGE density with large p but small n and is duly right-skewed. Finally, the green line plots a density with large n and small p, and is left-skewed. This demonstrates the flexibility of the BEGE distribution and makes tangible the role of p as the "good environment" variable and n as "the bad environment" variable.

For dividend growth, we follow CC's specification in which dividend growth follows a process symmetric to that for consumption growth:

$$\Delta d_t = \overline{g} + \sigma_{dp}\omega_{p,t} - \sigma_{dn}\omega_{n,t} \tag{5}$$

where Δd_t is real logarithmic dividend growth.

II.B Preferences

We use the preference structure of the representative agent of CC:

$$E_0\left[\sum_{t=0}^{\infty}\beta^t\frac{(C_t - H_t)^{1-\gamma} - 1}{1-\gamma}\right], \tag{6}$$

where C_t is aggregate consumption and H_t is an exogenous "external habit stock" with $C_t > H_t$.

In CC, H_t is an exogenously modelled subsistence or habit level. Hence, the local coefficient of relative risk aversion equals $\gamma\frac{C_t}{C_t - H_t}$, where $\left(\frac{C_t - H_t}{C_t}\right)$ is defined as the surplus ratio and denoted S_t. As the surplus ratio goes to zero, the consumer's local risk aversion goes to infinity. Of course, this is not actual risk aversion defined over wealth, which depends on the value function. The appendix to Campbell and Cochrane (1995) examines the relation between "local" curvature and actual risk aversion, which depends on the sensitivity of consumption to wealth. CC derive an expression for risk aversion (RA_t) as a function of local curvature and the relative valuation of two other assets:

$$RA_t = \gamma\frac{C_t}{C_t - H_t} \cdot \frac{PC_t}{PCS_t} = \gamma S_t^{-1} \cdot \frac{PC_t}{PCS_t}$$

where PC_t is the price-dividend ratio for a claim to the consumption stream, and PCS_t is the price-dividend ratio of a claim to the stream of payouts, $\{C_{t+\tau} \cdot S_{t+\tau}\}_{\tau=1}^{\infty}$.

The intertemporal marginal rate of substitution in this model determines the real pricing kernel, which we denote by M_t. Taking the ratio of marginal utilities at time $t+1$ and t, we obtain:

$$m_{t+1} = \ln \beta - \gamma \Delta c_{t+1} - \gamma \left(s_{t+1} - s_t \right) \tag{7}$$

where $s_t = \ln(S_t)$ and $m_{t+1} = \ln\left(M_{t+1}\right)$.

We specify the process for s_t exactly as in CC:

$$s_{t+1} = \left(1 - \phi\right) \overline{s} + \phi s_t + \lambda_t \left(\Delta c_{t+1} - \overline{g} \right) \tag{8}$$

where $\overline{s}$, and ϕ are scalar parameters. As in CC, λ_t is the so-called sensitivity function, and is decreasing in s_t. Negative consumption shocks decrease the surplus ratio and increase λ_t, making stocks riskier and the equity premium larger. This feature of the model generates counter-cyclical expected returns and prices of risk. Because we follow CC in choosing the sensitivity function to make the interest rate constant, we defer a full characterization of the sensitivity function in our framework to our presentation of asset prices in Section II.E.

It will be convenient to denote the exposure of the pricing kernel to the two consumption shocks, $\omega_{p,t}$ and $\omega_{n,t}$, as follows

$$\begin{aligned} a_{p,t} &\equiv \partial m_{t+1} / \partial \omega_{p,t+1} = -\gamma \left(1 + \lambda_t\right) \sigma_{cp} \\ a_{n,t} &\equiv \partial m_{t+1} / \partial \omega_{n,t+1} = +\gamma \left(1 + \lambda_t\right) \sigma_{cn} \end{aligned} \tag{9}$$

Naturally, positive $\omega_{p,t}$ shocks lower marginal utility, while positive $\omega_{n,t}$ shocks raise it.

II.C Variance premiums

The variance premium is defined as the difference between the conditional variance of equity returns under the risk-neutral and physical measures.

$$vprem_t = VAR_t^Q \left[ret_{t+1} \right] - VAR_t^P \left[ret_{t+1} \right] . \tag{10}$$

and ret_{t+1} denotes log equity returns. As noted above, the CBOE's VIX index essentially provides direct readings on the risk-neutral variance. Irrespective of the measures used for physical variance, the variance premium is reliably positive and changes considerably

through time. In this subsection, we show why the standard CC model with Gaussian shocks is unlikely to generate a variance premium. To do so, we revert to the original CC formulation for consumption growth

$$\Delta c_{t+1} = \overline{g} + \sigma_{cc}\varepsilon_{t+1} \tag{11}$$

where ε_{t+1} is distributed as a standard Gaussian random variable. To price the equity claim, we follow CC in assuming that the dividend growth process is:

$$\Delta d_{t+1} = \overline{g} + \sigma_{dc}\sigma_{cc}\varepsilon_{t+1} + \sigma_{dd}\nu_{t+1} \tag{12}$$

where ν_{t+1} is also a Gaussian shock that is independent of ε_t, which allows for imperfect correlation between consumption and dividend growth. Log equity returns are, by definition,

$$ret_{t+1} = \Delta d_{t+1} - pd_t + gpd_{t+1} \tag{13}$$

where pd_t is the log price-dividend ratio, and $gpd_{t+1} = \ln\left(1 + \exp\left(pd_{t+1}\right)\right)$. Both pd_t and gpd_t, are nonlinear functions of s_t, which is the only state variable. CC show that the price dividend ratio is a nearly linear function of the surplus ratio (See their Figure 3).

To elucidate the nature of the variance premium in this model, we examine a second order approximation to the function describing the relation between gpd_t and the state variable, s_t:

$$gpd_t = \kappa_0 + \kappa_1 s_t + \kappa_2 s_t^2 \tag{14}$$

where κ_0, κ_1, and κ_2 are coefficients that are functions of deeper model parameters. As derived in Appendix A, in this case, the variance of returns under the physical measure equals:

$$VAR_t^P\left(ret_{t+1}\right) = \sigma_{dd}^2 + \left(\sigma_{dc} + \kappa_1\lambda_t + 2\kappa_2\overline{s}_t\lambda_t\right)^2\sigma_{cc}^2 + 2\kappa_2^2\lambda_t^4\sigma_{cc}^4 \tag{15}$$

where $\overline{s}_t = E_t\left[s_{t+1}\right]$. This simplifies in the linear case ($\kappa_2 = 0$) to

$$VAR_t^P\left(ret_{t+1}\right) = \sigma_{dd}^2 + \left(\sigma_{dc} + \kappa_1\lambda_t\right)^2\sigma_{cc}^2$$

Clearly, the CC model generates time-varying volatility in returns as λ_t varies over time regardless of whether κ_2 is nonzero. We also derive in Appendix A an expression for the

risk-neutral variance:

$$\begin{aligned} VAR_t^Q\left(ret_{t+1}\right) &= \sigma_{dd}^2+\left(\sigma_{dc}^2+\kappa_1^2\lambda_t^2+2\kappa_2\overline{s}_t\lambda_t\right)\sigma_{cc}^2+2\kappa_2^2\lambda_t^4\sigma_{cc}^4 \\ &\quad -4\kappa_2\lambda_t^2\sigma_{cc}^2 a_{cc,t}\left(\sigma_{dc}+\kappa_1\lambda_t+2\kappa_2\lambda_t\overline{s}_t\right)\sigma_{cc} \end{aligned} \tag{16}$$

where $a_{cc,t}=-\gamma\left(1+\lambda_t\right)\sigma_{cc}$ is defined as the sensitivity of the pricing kernel to the consumption shock under the original CC formulation. Clearly, if $\kappa_2=0$ (linearity) or $\gamma=0$ (so that $a_{cc,t}=0$, implying risk-neutrality) then $VAR_t^P\left(ret_{t+1}\right)=VAR_t^Q\left(ret_{t+1}\right)$ and $vprem_t=0$. Hence, under the second order approximation, the traditional CC model with Gaussian shocks generates a variance premium only to the extent that gpd_t is nonlinear in s_t. These results also highlight a more general feature of conditionally Gaussian models. Unless the dependence of the price-dividend ratio on the state variable is nonlinear, Gaussian models will not generate a variance premium. This is discussed further in Appendix A. For instance, unless a model incorporating stochastic volatility into the process for fundamentals generates strong non-linearities in the dependence of equity prices on the state variable, it cannot generate a variance premium. We show below that empirically the original CC model with Gaussian shocks does not generate a meaningful variance risk premium.

We next demonstrate that, in contrast to Gaussian models, the BEGE model generates a "first-order" variance risk premium. That is, even if equity prices are linear in the state variable, the model still generates a meaningful variance premium. To make this point as stark as possible, we restrict p and n to be constant, so that s_t remains as the only state variable. Now we assume a linear approximation of the functional form of gpd_t

$$gpd_t=\kappa_0+\kappa_1\ s_t \tag{17}$$

where κ_0 and κ_1 are linearization coefficients. As shown in Appendix A, under the physical measure, the variance of returns is

$$VAR_t^P=p\left(\sigma_{dp}+\lambda_t\kappa_1\sigma_{cp}\right)^2+n\left(\sigma_{dn}+\lambda_t\kappa_1\sigma_{cn}\right)^2 \tag{18}$$

while the risk-neutral variance is

$$VAR_t^Q=p\left(\frac{\sigma_{dp}+\lambda_t\kappa_1\sigma_{cp}}{1-a_{p,t}}\right)^2+n\left(\frac{\sigma_{dn}+\lambda_t\kappa_1\sigma_{cn}}{1-a_{n,t}}\right)^2 \tag{19}$$

Because $a_{p,t} < 0$, the exposure of returns to the $\omega_{p,t+1}$ shock contributes negatively to the variance premium. On the other hand, because $a_{n,t} > 0$, and assuming that $a_{n,t} < 1$, exposure to the $\omega_{n,t+1}$ shock contributes positively to the variance premium.

These results are related to existing results in the options pricing literature. Bakshi and Madan (2006), for example, study the determination of volatility spreads (which is our variance premium expressed as a percent of the physical variance) in a dynamic economy with general preferences that however only depend on the stock market return. Using a second order Taylor series expansion of the utility function, they show that the volatility spread is related to risk aversion, and the skewness and excess kurtosis of the physical distribution of returns. If the return distribution is Gaussian, the volatility spread is zero.

II.D Variants of the BEGE model

Given that the BEGE structure can generate a first-order variance premium even with constant shape parameters, we start our analysis simply appending an unconditional BEGE structure with constant p and n to the CC model. However, it is straightforward to consider a dynamic BEGE model by having p and/or n vary through time. This dynamic version of the model is critical for matching features of the consumption and asset pricing data, as will be demonstrated in Sections III and IV.

For n, we assume a simple, autoregressive process with "square-root-like" volatility dynamics,

$$n_t = (1 - \rho_n)\,\overline{n} + \rho_n n_{t-1} + \rho_n \sigma_{nn} \omega_{n,t} \tag{20}$$

where $\overline{n}$ is the unconditional mean of the process, ρ_n is the autocorrelation coefficient, and σ_{nn} governs the conditional volatility of the process. Specifically, the conditional volatility of n_t is $\rho_n \sigma_{nn} \sqrt{n_{t-1}}$. Importantly, the minimum possible value of $\omega_{n,t}$ is $-n_t$, an artifact of the one-sided nature of the gamma distribution. A very useful implication of the finite left tail of the distribution is that, under the additional technical assumption that $\sigma_{nn} < 1$, the n_t process can never become negative, even at discrete intervals. This is a highly

desirable, however uncommon, property for a volatility process, which is not shared by Gaussian stochastic volatility processes. Note that if we introduce n_t as a state variable, we must assume that the representative agent observes n_t as well as s_t.

We could model p_t symmetrically to n_t, but choose not to do so for simplicity. With time-varying p_t and n_t, the BEGE model can be interpreted as a stochastic volatility model with "good" (governed by p_t) and "bad" (governed by n_t) volatility. However, as volatility changes the entire conditional distribution of consumption growth changes as well.

II.E Asset Prices Under the BEGE-Habits Model

The main mechanism of the model in generating countercyclical risk premiums follows directly from the habits model in CC. We therefore keep the theoretical analysis of the asset price implications limited, but show empirically the contributions of the BEGE framework in Section IV. Here we limit ourselves to first explaining how we model the risk free rate and what the BEGE set-up adds to the expression for the Sharpe ratio, a critical variable in CC's empirical approach as they select parameters to match it exactly.

II.E.1 The real short rate

To solve for the real risk free short rate, rrf_t, we use the usual first order condition for the consumption-saving choice,

$$\exp\left(rrf_t\right) = E_t\left[\exp\left(m_{t+1}\right)\right]^{-1} \tag{21}$$

Under the above BEGE dynamics and CC preference structure, this simplifies to

$$\begin{aligned} rrf_t = & -\ln\beta + \gamma\overline{g} + \gamma\left(1-\phi\right)\left(\overline{s}-s_t\right) \\ & +p\ f\left(a_{p,t}\right) + n_t\ f\left(a_{n,t}\right) \end{aligned} \tag{22}$$

where we, again, use the function, $f\left(x\right) = x + \ln\left(1-x\right)$. The first line in the solution for rrf_t has the usual consumption and utility smoothing effects: to the extent that marginal utility is expected to be lower in the future (that is, when $s_t < \overline{s}$), investors desire to borrow to smooth marginal utility, and so risk free rates must rise. The bottom line captures

higher-order precautionary savings effects. Notice that because the function $f(x)$ is always negative, the precautionary savings effects are also always negative. A third-order Taylor expansion of the log function helps with the interpretation of rrf_t:

$$rrf_t \approx \begin{pmatrix} -\ln\beta + \gamma\overline{g} + \gamma(1-\phi)(\overline{s} - s_t) \\ + \left(-\frac{1}{2}a_{p,t}^2 - \frac{1}{3}a_{p,t}^3\right)p \\ + \left(-\frac{1}{2}a_{n,t}^2 - \frac{1}{3}a_{n,t}^3\right)n_t \end{pmatrix} \tag{23}$$

The second order terms capture the usual precautionary savings effects: higher volatility generally leads to increased savings demand, depressing interest rates. Under a Gaussian distribution all higher order terms are zero. The cubic terms represent a novel feature of the BEGE model. The third order term for p is positive (recall that $a_{p,t} < 0$), so that it mitigates the precautionary savings effect. This makes perfect economic sense. When good environment shocks are dominant, the probability of large positive shocks is relatively large, and the probability of large negative shocks is small, decreasing precautionary demand. Conversely, both the second and third order (and all higher order) terms premultiplying n_t are negative (recall that $a_{n,t} > 0$) indicating that precautionary savings demands are exacerbated when n_t is large. That is, when consumption growth is likely to be impacted by large, negative shocks, risk free rates are depressed over and above the usual precautionary savings effects relative to volatility. Through this mechanism, our model may generate the kind of extremely low risk free rates witnessed in the 2007-2009 crisis period.

To abstract from term structure dynamics, we follow CC in ensuring that the short rate is constant. We do so by choosing the sensitivity function to be a function of s_t and n_t such that the short rate is constant:

$$\lambda_t = \lambda(s_t, n_t, p; \theta) : rrf_t = \overline{rrf}$$

No closed-form solution for λ_t exists, but it is easily found numerically. Appendix B explores the evaluation of λ_t in more detail. As in CC, for some combinations of $\left(\overline{rrf}, s_t, n_t; \theta\right)$, it may not be possible to find a value for λ_t that satisfies the above condition. In these unusual cases, which occur when s_t approaches zero, we set λ_t to zero, as do CC. Appendix B shows

that the BEGE sensitivity function looks similar to the one in a standard CC model, but lower sensitivity is required for extremely negative values of s_t. This is because the BEGE model generates stronger precautionary savings effects than a Gaussian model does, which helps to offset the intertemporal smoothing demands generated by deviations of s_t from its mean value.

II.E.2 The Maximum Sharpe Ratio

In the CC model, the pricing kernel is log normal and there exists a closed form solution for the maximum possible Sharpe ratio, which is the coefficient of variation of the pricing kernel:

$$\max_{\{\text{all assets}\}} \frac{E_t\left[R^e_{t+1}\right]}{\sigma_t\left[R^e_{t+1}\right]} = \frac{\sigma_t\left[M_{t+1}\right]}{E_t\left[M_{t+1}\right]} \approx \sqrt{a^2_{cc,t}} \tag{24}$$

where R^e_{t+1} is the excess return of a generic asset, and we again denote $a_{cc,t} = -\gamma\sigma_{cc}\left(1+\lambda_t\right)$, the sensitivity of the pricing kernel to the consumption shock under the original CC formulation. The Sharpe ratio is also easy to compute in our framework, given the presence of two independent gamma processes, but the resulting expression is not very intuitive. In Appendix C, we derive an approximation to the Sharpe ratio under the BEGE-habit model:

$$\frac{\sigma_t\left[M_{t+1}\right]}{E_t\left[M_{t+1}\right]} \approx \sqrt{p\left[a^2_{p,t} + 2a^3_{p,t}\right] + n_t\left[a^2_{n,t} + 2a^3_{n,t}\right]} \tag{25}$$

Absent the third order terms, this looks very much like Equation (24) except with two sources of uncertainty. Intuitively, the third order effect of p is to lower the maximum Sharpe ratio (recall that $a_{p,t} < 0$), while the third order effect of n_t is to raise it further (recall that $a_{n,t} > 0$). This further highlights the role of p_t (n_t) as governing "good" ("bad") volatility.

III Consumption Growth Dynamics

In this section, we examine the properties of U.S. post-war consumption growth data to determine whether significant non-Gaussianities are present, and how well a BEGE model

can fit them.

III.A The Unconditional Picture

The key innovation of the BEGE model is to accommodate potential non-Gaussianities in the stochastic process for consumption growth. But do we observe such non-linearities in the data? When we allow the representative agent to take into account the Great Depression, the evidence of non-Gaussianities is abundantly clear. For instance, in the U.S. National Income and Product Accounts (NIPA, published by the Bureau of Economic Analysis), the sample skewness of the annual growth rate of real personal consumption expenditures from 1930 through 2012 is $-\dot{0}.81$ and the sample *excess* kurtosis is 3.49. Having just suffered another deep recession, it is conceivable that economic agents would think that a recurrence of a Great Depression-like economic contraction is possible. However, it is well-known that the consumption growth data were collected differently before World War II.[3] We therefore focus attention on post-war data to present a conservative view of potential non-Gaussianities in the data.

Our consumption growth series is constructed using NIPA data. We first add together quarterly, nominal consumption expenditures for nondurables and services over the period 1947-2012. We deflate the resulting series using a weighted average of the deflators for nondurables and services, where the weights are determined by the respective real expenditure shares. Finally, we further deflate the resulting series using the growth rate of the U.S. population from the U.S. Census.[4] The first-order autocorrelation coefficient of the resulting series is 0.32. Such autocorrelation may well be due to temporal aggregation bias (see Working, 1960). We therefore pre-whiten the data using an AR(1) model as in Wachter (2005) and work with serially uncorrelated consumption growth data henceforth.

In Table 1, we estimate unconditional consumption growth models by maximum likeli-

[3] We thank an anonymous referee for pointing this out.

[4] Quarterly population data are available starting in 1958, but only annual data are avialable prior to that date. For the early part of the sample, we assume constant quarterly growth rates within each year.

hood. The left column reports the Gaussian model parameters, followed by log-likelihood value and a specification test building on the standard Jarque-Bera (1987, JB henceforth) test for normality. We apply the normality test to the inverse normal cumulative density function of the assumed cumulative distribution function (cdf henceforth) of the data. If the model is correctly specified, this transformation should lead to a normally distributed variable. This is true because for a correctly specified model, the cumulative distribution function applied to the data should be distributed as uniform on the [0,1] interval, and by the inverse probability integral transform, taking the inverse Gaussian distribution function of a uniform distributed random variable should yield a Gaussian random variable. This test thus reduces to the standard Jarque-Bera test. The Gaussian model is rejected at the 1% level using the JB test.

In the column on the right, we show estimates of the BEGE model with constant p and n. We estimate p to be well above 1, but n is below 1, implying strong non-Gaussianities for the "bad environment" part of the distribution. Even though the likelihood value for this model is much higher than for the Gaussian model, note that the two models are not strictly nested, so that we cannot perform a standard likelihood ratio test. However, the Jarque-Bera test now has a p-value of 0.50, and so the specification test fails to reject the model.[5]

In the top panel of Figure 2 we show the two probability density functions from the estimates in Table 1 overlaid with a histogram of the data. The better fit of the BEGE model relative to the Gaussian model is visible in the graph. In the bottom panel, we focus on the cumulative distribution function (cdf) of the two distributions. Shown are "mountain" plots: the left portion of the curves (i.e. at values less than the median) plot the cdf, and the right portion shows 1 minus the cdf. This enables us to read tail probabilities for both the left and right tails directly. For reference, the horizontal line indicates the

[5]Monte Carlo results suggests that this test procedure has roughly the proper empirical size: Using simulations in which the null hypotheis is true (that is, the data is generated by the static BEGE model under the parameters reported in the right column), the test procedure rejects in 3.9 percent of draws when the nominal size is set to 5 percent.

estimated probability of a "catastrophic" collision of a large asteroid with the Earth in a given year, as computed by NASA. We view such an event as (hopefully) rare and unlikely. Yet, according to the Gaussian distribution, having annualized consumption growth of -7 percent is equally rare. Moreover, a realization of consumption growth of -10 percent or worse at an annual rate is one million times less likely under the Gaussian distribution than such a collision. This seems economically unreasonable as recessions of that magnitude have occurred multiple times during modern history. For example, in the US real consumption growth registered -9 percent in 1932, and real GDP fell by 7 percent in Greece in 2011.[6] We conclude that the BEGE model displays more plausible tail behavior than the Gaussian model.

III.B The Conditional Distribution of Consumption Growth

While there is clear evidence of unconditional non-Gaussianities in the data, this does not necessarily imply that consumption growth is conditionally non-Gaussian. For instance, a standard conditionally Gaussian stochastic volatility model may also generate unconditional non-Gaussianities. In this subsection, we examine the evidence for conditional non-Gaussianity

In Table 2, we use the approximate likelihood procedure of Bates (2006) to estimate the BEGE model with a time-varying n_t process as in Equation (20). The estimation procedure is described in more detail in Appendix D. We also estimate a conditionally Gaussian counterpart to the dynamic BEGE model, a standard stochastic volatility model,

$$\begin{aligned} \Delta c_{t+1} &= g + \sigma_{cs}\varepsilon_{t+1}^1 - \sigma_{cv}\sqrt{v_t}\varepsilon_{t+1}^2 \\ v_t &= \overline{v} + \rho_v\left(v_{t-1} - \overline{v}\right) + \rho_v\sigma_{vv}\sqrt{v_{t-1}}\varepsilon_t^2 \end{aligned} \tag{26}$$

where both ε_t^1 and ε_t^2 are IID Gaussian innovations. Table 2 reveals that the stochastic volatility model suggests very persistent volatility (autocorrelation of 91%) but is still

[6] One could also argue that these observations are drawn from different distributions with larger variances, in which case they may not represent such extreme outliers.

strongly rejected by the Jarque-Bera test (at the 1% level). The BEGE model delivers a much higher likelihood and is again not rejected by the Jarque-Bera test. The n_t process, which generates time-variation in the higher order moments, is also very persistent with a 0.93 autocorrelation coefficient. Compared to the static BEGE model of Table 1, we continue to estimate the (constant) value of p to be around 4, but the unconditional mean of n_t is now higher, about 2. Both of these parameters have fairly large standard errors, which is not surprising since it is rather difficult to extract the dynamics of latent variables from relatively low frequency data such as consumption growth.

We therefore also consider asymmetric volatility models in the GARCH class, specifically, the model put forward by Glosten, Jagannathan and Runkle (1993), where a negative shock has potentially a different effect on next period's volatility than a positive shock. The standard Gaussian model is specified as:

$$\begin{aligned} \Delta c_{t+1} &= g + \sigma_{cs}\varepsilon^1_{t+1} - \sigma_{cv}\sqrt{v_t}\varepsilon^2_{t+1} \\ v_t &= v_0 + \rho_v v_{t-1} + \phi_{v1}\ (\Delta c_t - g)^2 + \phi_{v2}\ (\Delta c_t - g)^2 \cdot I_{\Delta c<g} \end{aligned} \tag{27}$$

where ε^1_t and ε^2_t are $N(0,1)$. The right column shows parameter estimates for the BEGE model:

$$\begin{aligned} \Delta c_t &= g + \sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1} \\ n_t &= n_0 + \rho_n n_{t-1} + \phi_{n1}\ (\Delta c_t - g)^2 + \phi_{n2}\ (\Delta c_t - g)^2 \cdot I_{\Delta c<g} \end{aligned} \tag{28}$$

While these auxiliary models do not correspond directly to the models developed in our theoretical section, they can be estimated by exact maximum likelihood estimation. Table 3 (Panel A) reports the results for both models. In both cases, the volatility is estimated to be persistent but less so than in the stochastic models reported in Table 2. While the coefficient on negative consumption growth shocks is higher than the coefficient on positive growth shocks, the evidence is statistically weak. Again, the standard volatility model is still rejected by the Jarque-Bera normality test but the BEGE model is not rejected by the Jarque-Bera test. In Panel B, we estimate the parameters of the stochastic volatility models

that best correspond to the GARCH models. We do so by matching moments of the latent process, which we compute by simulation for the GJR models. Fortunately, these are quite consistent with the estimates from the approximate likelihood procedure. For example, the mean of n_t is now 1.87 (versus 1.98 in the approximate likelihood procedure) and its persistence is 0.95 (versus 0.93 in Table 2). Note that p is again very close to 4 in this model. We therefore use the parameters of the direct estimation reported in Table 2 in further analysis.

III.C Further Assessing the Models' fits to Consumption Data

In Table 4, we consider the fit of the various models with the unconditional variance, skewness and kurtosis of both consumption and dividend growth data. The dividend series that we use is seasonally adjusted dividends per share for the S&P 500 index, deflated by the same consumption deflator and population values described above for consumption. As is well-known, dividends at the quarterly frequency are a somewhat problematic times series. We use parameters in the dividend growth model (Equation 5) to match the sample volatility of dividend growth, and set its correlation with consumption growth to the value used by CC, 0.20.

The two models on the left with an "s" indicator are the "static" models estimated in Table 1; the two models on the right are dynamic models, indicated by "d," estimated in Table 2. In the middle column, the "s2" model is a static model that matches the sample volatility, skewness and kurtosis of quarterly consumption growth exactly. We calibrate the model parameters by matching sample statistics to analytic counterparts under the model.[7] In Panel A, we investigate the fit with quarterly data and in Panel B we aggregate the quarterly data to the annual frequency. In quarterly data, the volatility is matched by construction for all models. The Gaussian models naturally cannot fit the negative skewness of -0.32, which is significant at the 5 percent level, but the stochastic volatility model also

[7] This calibration is not unique since we are matching three moments (volatility, skewness, and kurtosis), using four parameters, σ_{cp}, σ_{cn}, p, and n.

under-estimates excess kurtosis by about 2 standard errors. Unfortunately, both the static and dynamic BEGE models generate too much excess kurtosis in quarterly consumption growth data. This motivated us to also consider the "s2" model that matches skewness and kurtosis exactly. At the annual frequency, the higher order moments are not measured with sufficient precision to reject any model.

Even though we cannot do a formal likelihood ratio test, the evidence in Tables 2 and 3 strongly suggest that a dynamic BEGE model fits the consumption growth data better than a static one. In Figure 3 and Table 5, we try to give economic and statistical content to this better fit. The BEGE model ought to capture recession risk better than a Gaussian model. In Figure 3, we graph the pre-whitened consumption growth data. In the context of our model, the concept of NBER recessions is ill-defined, since NBER recessions are defined using multiple criteria and time series variables. We therefore define the alternative concept of a "consumption slow down" as any observation where the 4-quarter moving average of consumption growth is lower than 1 percent at an annual rate. These periods are shown with gray shading. As indicated by the inset figures in the graph, NBER recessions tend to coincide with consumption slowdowns. There is a 68 percent chance of a consumption slowdown when a NBER recession is observed, and only a 9 percent chance of a slowdown in other periods. Whereas NBER-defined recessions occur in 16 percent of quarters in our sample and persist for an average of 3.7 quarters, the comparable statistics for consumption slowdowns, shown at the top of Table 5, are very similar with a frequency of 18 percent and an average duration of 3.8 quarters. Both the Gaussian stochastic volatility model and the dynamic BEGE model are able to match the slowdown statistics fairly well. However, both models imply slowdown durations that are a bit too short. This likely owes to the fact that mean consumption growth falls during slowdowns, and we intentionally abstract from this effect here.[8]

[8] Apart from higher volatility and negative skewness, slowdowns are also characterized by a downward shift in mean consumption growth. We ignore this shift in the present paper but it could obviously be modeled as affected by the n_t process as well. That is, the BEGE model could use one state variable to generate three key characteristics of recessions.

The key graph of interest is in the bottom panel of Figure 3. It shows histograms for two conditional distributions, one conditional on a consumption slowdown in the previous quarter, the other conditional on no slowdown in the previous quarter. Clearly, the "slowdown' distribution is wider and more negatively skewed than the "no slowdown" distribution. If these differences are significantly different from zero, they are direct evidence of a need for conditional non-Gaussianity. We quantify these differences in Table 5. The lower tail of the distribution, represented by the 5th percentile, is about -3 percent (at an annual rate) conditional on a slowdown, but only -1 percent after normal periods. The difference is highly significantly different from zero. These statistics could, of course, be matched both by a Gaussian stochastic volatility model and a dynamic BEGE model. However, the upper tail of the distribution, represented by the 95 percentile does not exhibit a significant increase during a slowdown. A Gaussian stochastic volatility model cannot match this asymmetry in tail behavior, but the BEGE model does reasonably well: The expansion of the negative tail is matched almost perfectly by the BEGE model, but the model implies slightly too much of an increase in the positive tail during slowdowns.[9] We conclude that the dynamic BEGE model presents the more realistic representation of consumption growth data in the US.

IV Asset Pricing Model Solution and Empirical Performance

To evaluate the BEGE model's performance with respect to asset prices, we first discuss the choice of parameters and model solution technique, then summarize some univariate statistics, and end with a discussion of important cross-moments.

[9] Allowing for time variation in p_t, the "good environment variance," would likely improve performance along this dimension.

IV.A Parameterization

We closely follow the approach in CC to select parameters. For consumption growth, we use the parameters estimated for the various models and discussed in Section III. The only parameters left to select are the discount factor, β, the curvature parameter of the utility function, γ, and the parameters governing the surplus ratio (Equation 8). As in CC, we choose a value of γ to match the unconditional Sharpe ratio on equity returns in the data. The equity return data are quarterly excess returns for the S&P 500 index over the 1947-2011 period, where the nominal 90-day Treasury bill return is used as the risk free rate. This procedure implies that γ differs across models, ranging from 1.7 to 8 as reported in Table 6. The selected persistence parameter for the surplus ratio, ϕ, is taken from CC, who used it to match the persistence of the dividend-price ratio. Model BEGE(d)' is an alternative parameterization where ϕ is chosen to match the unconditional (physical) variance of equity returns exactly in addition to the Sharpe ratio. The other parameters are chosen to deliver a value of 1 percent (annual rate) for the real risk free rate and keep the real interest rate (approximately) constant.

To solve the model, we note that the price dividend ratio is the present value of future discounted dividends:

$$\frac{P_t}{D_t} = E_t \left[\sum_{i=1}^{\infty} \exp \left(\sum_{j=1}^{i} (m_{t+j} + \Delta d_{t+j}) \right) \right] \tag{29}$$

Because the surplus ratio formulation contains a significant non-linearity, no closed-form solutions are available for this expectation. Alternative formulations of time-varying risk aversion (see e.g. Bekaert, Engstrom and Grenadier, 2010) combined with BEGE fundamentals would allow closed-form solutions because of the tractable moment generating function of the gamma distribution. We explored such a model in a previous version of this article. To implement a numerical solution of Equation (29), we follow the approach proposed by Wachter (2005). We recursively solve for the individual terms in the summation in Equation

(29). That is,

$$\frac{P_t}{D_t} = \sum_{i=1}^{\infty} F_i\left(s_t, n_t\right)$$

where

$$F_i\left(s_t, n_t\right) = E_t\left[\exp\left(\sum_{j=1}^{i}\left(m_{t+j} + \Delta d_{t+j}\right)\right)\right]$$

and

$$F_i\left(s_t, n_t\right) = E_t\left[\exp\left(m_{t+1} + \Delta d_{t+1}\right) F_{i-1}\left(s_{t+1}, n_{t+1}\right)\right]$$

While $F_1\left(s_t, n_t\right)$ can be calculated analytically, subsequent terms are solved for numerically on a grid as suggested by Wachter. Wachter shows that this "series method" is more accurate and converges faster than the Euler equation method employed by CC. A full description of this procedure is presented in Appendix E. A similar method can be used to price a consumption claim, which is useful for computing the distribution of risk aversion in this economy. In panel B of Table 6, we report features of the distribution of local curvature and actual risk aversion. As is well-known, the CC model generates very high risk aversion and its median is in fact about 258. Note that the BEGE model requires less curvature and risk aversion to match the Sharpe ratio in the data. This is quite intuitive: High risk aversion enables models to match the high equity risk premium by amplifying marginal utility in extremely bad states of the world. But the BEGE distribution suggests that the probability of landing in a bad state is higher than a Gaussian model would suggest. As a result the BEGE model requires less risk aversion. In this respect, the BEGE model is similar to "rare disaster" models such as Rietz (1990) and Barro (2006).

Table 7 illustrates some key mechanisms of our preferred dynamic model, BEGE(d)'. In Panel A, we show how some key endogenous quantities vary with the two state variables, s_t and n_t. For each of the endogenous variables, in the first column, we report their values when s_t and n_t are at their sample median values. In the second column, we report values when risk aversion is high (that is s_t takes on a more extreme negative value, the 10th percentile of its unconditional distribution), but n_t remains at its median. Finally, in the

third column, we report values for the situation in which s_t is again at its median value, but n_t is now elevated (at the 90th percentile of its distribution). Results for the sensitivity function, λ_t, are shown first. As in CC, the sensitivity function is higher at more extreme negative values of s_t This is required for precautionary savings and utility smoothing effects to exactly offset and keep the short rate constant. In contrast, λ_t is decreasing in n_t. To see why, note that higher n_t increases the volatility of the consumption shock. All else equal, this would depress the risk free rate due to elevated precautionary saving demand. To keep the short rate constant, the sensitivity function must be lower. Because they are linear in λ_t, both $a_{p,t}$ and $a_{n,t}$, the sensitivities of the pricing kernel to the two consumption shocks, follow a similar pattern. The variance of the pricing kernel, under both the physical and risk neutral measures, is therefore strongly affected by changes in s_t, but little affected by changes in n_t. The increase in the variance of the kernel that is induced by a higher level of n_t is offset by the effect that n_t has in reducing the kernel sensitivities, $a_{p,t}$ and $a_{n,t}$. Finally the maximum available Sharpe ratio in the economy is affected primarily by s_t.

Despite being deliberately prevented from having strong effects on the conditional moments of the pricing kernel, n_t still has important asset pricing implications. To provide a hint of how this may be the case, the bottom two rows of the Panel A report the conditional volatility of dividend growth, which is imperfectly correlated with consumption growth, though driven by the same two shocks (Equation 5). Obviously, s_t has no effect on the conditional volatility of dividend growth, but the volatility of dividends increases in n_t under both the physical and risk-neutral measures, as does the dividend growth volatility premium.[10]

IV.B Univariate statistics

The quarterly data we use for interest rates and equity prices and returns is standard and spans the period, 1947-2011. We measure the (nominal) short rate, rf_t, as the 90-

[10]The variance premium for dividend growth is defined as the difference between its variance under the risk-neutral and physical measures.

day nominal T-bill yield provided by the Federal Reserve. For equity prices, we use the logarithmic dividend yield, dp_t, for the S&P 500, calculated as (one-fourth of) trailing four-quarter dividends divided by the quarter-end price. The nominal return to equity, ret_t, is the logarithmic change in the quarter-end level of the S&P 500 plus the quarterly dividend yield defined above. Excess returns are measured as $ret_t - rf_{t-1}$. We calculate the physical probability measure of the equity return conditional variance, $pvar_t$, in two steps. We begin with the monthly realized variance, $rvar_t$, calculated as squared daily returns over the quarter. Then we project $rvar_t$ onto lagged values of $rvar_t$, as well as realized variance measured over just the most recent month and the most recent week of data. The fitted values from this regression are used to measure $pvar_t$. This procedure is similar in spirit to Corsi's (2009) heterogeneous autoregressive model.

To calculate standard errors for our sample moments, we use a block-bootstrapping routine. In each of 10000 bootstraps, blocks of the data spanning 40 quarters are drawn, with replacement, until the full sample length is achieved. For each bootstrapped sample, the desired statistics are calculated. The standard deviation of these estimates across bootstraps are reported as standard errors for the sample statistics.

We calculate the risk-neutral conditional second, third and fourth moments of equity returns, $qvar_t$, $qskw_t$, and $qkur_t$ respectively, using the method of Bakshi, Kapadia and Madan (2003). This involves calculating the prices of portfolios of options designed to have payoffs that are determined by particular higher order moments of returns. We obtained a panel of option prices across the moneyness spectrum for the S&P 500 index from 1996-2009 from OptionMetrics, and from 1990-1995 from DeltaNeutral. We used the option contracts that have maturity closest to one quarter, and filtered out illiquid options according to the rules described in Figlewski (2000). This unfortunately results in a shorter sample for risk-neutral moments than for the rest of our data. While we cannot observe data for the 1947-1989 period, we use a procedure to sample from an approximate distribution for these observations. To do so, we use the 1990-2012 sample to regress the missing variables onto

variables that are available for the full period: consumption growth, dividend growth, the real short rate, equity returns, and the physical conditional volatility of returns. We then "fill in" the missing values for the risk-neutral moments for the early part of the sample using the projected values and random innovations drawn from the distribution of the projection errors, which is assumed to be Gaussian. This entire procedure is embedded within the bootstrapping routine, described above, that we use for calculating standard errors for all of our sample moments.

In Table 8, we report how the various models match univariate equity return statistics. The upper part of the table investigates the well-known moments of equity returns, but the lower part of the table focuses on physical and risk neutral volatility statistics and higher order risk neutral moments of equity returns.

We first focus on the standard moments of equity returns. The Gaussian habits model underestimates the equity premium of 5.5 percent observed in the data (although it is still within a two standard error band of the sample moment), underestimates the equity return volatility and the volatility of the price dividend ratio. The astute reader may be surprised as the published CC model matched these moments almost exactly. The main reason for this is that the solution reported in the original CC model used a relatively inaccurate numerical solution method, as explored in detail by Wachter (2005). Wachter (2006) shows how a slight modification of the CC model to allow for time-varying risk free rates brings back its power to match these salient features of equity return data. We demonstrate that accommodating realistic non-linearities in the consumption growth process, as the BEGE models do, has the same effect. The performance of the BEGE models is very good, but none of the models we consider is within a two standard error band of every single moment. The main culprit is that the kurtosis of excess returns is milder in the data than it is in the model. Recall that the various BEGE models over-estimate kurtosis in the consumption growth data and that is of course one reason for the overshooting of return kurtosis. However, even the BEGE model which matches the higher order moments of consumption growth (model s2)

generates return kurtosis slightly above a 3 standard error band of the data moment. The model that performs the best is the alternative dynamic BEGE model, BEGE(d)', in which ϕ is calibrated to match the physical variance of equity returns. The only moment missed by model BEGE(d)' by more than two standard errors is the mean of the price-dividend ratio.

We now examine the performance of the models in matching features of the conditional moments of returns, both physical and risk-neutral. Given our analytic results for the variance premium above, it is no surprise that the Gaussian habits model fails to generate a variance premium, or meaningful risk neutral skewness and kurtosis. The various variants of the BEGE model generate realistic variability of the physical volatility, but the benchmark dynamic model over-estimates it. All BEGE models generate an average variance premium that is too low (between 2.6 percent and 3.8 percent versus the 5.2 percent data moment), but is still within a two standard error band. The same is true for the variability of the variance premium, although a simple static model fits the moment almost exactly. The alternative dynamic model does not generate sufficient variability in the variance premium. All BEGE models generate realistic average risk neutral skewness but under-estimate its variability over time. The dynamic models generate about 5 times as much variability as the static models, illustrating the importance of the time variation in n_t. The mean of risk neutral kurtosis is best fit by the alternative static and the benchmark dynamic models, whereas the other two models over-estimate it. The variability of risk neutral kurtosis is matched best by the basic static model but overestimated by the dynamic models. In all, the BEGE habits model provides a reasonable description of option price dynamics in a very parsimonious framework, even though it is difficult to match all data moments simultaneously.

To examine the underlying mechanisms of the dynamic BEGE model that delivers these results, we return to Table 7. In Panel B we report how equity prices and their conditional moments vary with the state vector. Both s_t and n_t strongly affect the equity price-dividend ratio. For instance, a sharp increase in risk aversion (in the table, when s_t falls from its

median to its 10th percentile value), decreases equity valuations by 30 percent. Similarly, when downside risk rises substantially (n_t rises from its median to its 90th percentile value), equity valuations fall by about 10 percent. The adverse shocks to s_t and n_t cause a jump in the equity risk premium from 5.1 percent to 7.9 and 7.5 percent, respectively. From Panel A, we know that the effect of s_t is a "price of risk" effect whereas an increase in n_t increases the amount of dividend risk. Similarly, both adverse shocks increase the physical and risk-neutral volatility of equity returns, but s_t has a stronger effect on the variance premium than n_t does. An increase in n_t affects both the physical and risk neutral variance (see also Equations 18 and 19), increasing the amount of "bad volatility." However, the difference between changes in risk neutral variance versus physical variance is driven primarily by changes to the kernel sensitivities, $a_{p,t}$ and $a_{n,t}$, the absolute magnitudes of which are increasing (decreasing) in s_t (n_t). That n_t reduces the kernel sensitivities cancels out the effect that it would otherwise have on precautionary savings demand and interest rates, but also mitigates its effect on the variance premium.

IV.C Fundamentals and Asset Prices

In Table 9, we focus on the cross-correlation between the fundamentals (consumption and dividend growth) and asset prices. Because we use CC's habit model, we cannot escape the implication that consumption growth and equity excess returns remain excessively correlated. This correlation goes down substantially at longer horizons, at which point most models generate a correlation that is within or close to a two standard error band around the sample correlation of 19 percent. Nevertheless, the BEGE model generally produces an even higher correlation between consumption growth and equity returns than the Gaussian habits model. This is not surprising as the non-linear recession risk now contributes to asset returns. The dividend growth process is less correlated with asset returns in the BEGE model than it is in the original Gaussian CC formulation. However, this correlation is still much higher than in the quarterly data sample, but it is close to a two standard error band

in annual data.

In Panels C and D, we investigate conditional correlations between consumption growth and lagged price dividend ratios and equity return volatility. In the static models, the surplus ratio is the only state variable and therefore the price dividend ratio and volatility do not reflect recession risk. In the dynamic models, n_t is an additional state variable driving their dynamics and reflecting recession risk. The data confirm that such an additional state variable is necessary to capture the conditional distribution of consumption growth. In panel C, we look at quantile shifts (for the 5th, median, and 95th percentiles) for the consumption growth distribution when the price dividend ratio being above versus below its median value. When the price dividend ratio falls below its median value, the median of future consumption growth is little affected. Similarly, the 95th percentile remains essentially unchanged. However, the 5th percentile experiences a significant downward shift. The static models obviously do not generate any such conditional distribution shifts, but the dynamic BEGE models match them well, though they undershoot the shift in the left tail somewhat.

In Panel D, we condition the consumption growth distribution on equity return volatility and look at quantile shifts comparing observations when volatility in the previous quarter is above versus below its median level. We find a significant downward shift in the left tail when volatility is elevated, which only the dynamic BEGE models can deliver. However, the dynamic BEGE models also produces an increase in the right tail, which is not consistent with the sample data.

In Panel E, we report the contemporaneous correlation between excess equity returns and (physical) equity return volatility. This correlation is highly significantly negative at -0.38, reflecting the well-known asymmetric volatility phenomenon (see e.g. Wu, 2001). The static models have no real chance of matching this data moment, but the dynamic BEGE models, while generating substantially negative correlations, still fall somewhat short of the data moment.

IV.D Equity return predictability

In Table 10, we investigate the conditional distribution of equity returns. Panel A repeats the standard predictability regressions of equity returns on the price dividend ratio at various horizons, as also conducted in CC. We report the slope coefficients and R2 statistics, with bootstrapped standard errors. The betas are negative and generally significant and the per period magnitude is relatively constant over different horizons. The R^2 increases from 2 percent at the one quarter horizon to almost 30 percent at the 5 year horizon. The Gaussian model fits the slope coefficient pattern but fails to generate a meaningful increase in the R2 with the horizon. The R2 is a noisy statistic however, and we cannot reject the Gaussian model on this basis. The BEGE models mostly generate slightly larger (in absolute magnitude) slope coefficients, but, more importantly, they match the increase in R2's with horizon that we observe in the data as well. The static BEGE models do as well as the dynamic BEGE models on this score.

In Panel B, we investigate the conditional distribution of equity returns conditioning on physical lagged volatility[11]. As we did for consumption growth in Table 9, we report the 5th, 50th, and 95th percentile shifts for periods with above median versus below median lagged physical volatility. The median return is 1.90 percent higher in high physical volatility periods, a fact matched by all models within a two standard error range. The distribution of consumption growth conditional on high equity market volatility is substantial different than the distribution conditional on low equity return volatility: it is wider, has a lower left tail (significant at the 5% level) and a higher right tail (significant at the 10% level). The Gaussian model can generate such a shift, but it is too weak. The BEGE models all produce the right pattern and are generally not rejected in statistical terms, generating quantile shifts well within a two standard error band of the data. The only exception is the standard dynamic BEGE model which generates an upward shift in the positive tail that is too strong.

[11] It would also be interesting to look at the relationship with risk neutral volatility, but this severely restricts the sample length and thus also statistical power.

V Conclusion

To date, the consumption based asset pricing literature has mostly focused on matching unconditional features of asset returns: the equity premium, the low risk free rate, and the variability of equity returns and dividend yields. In terms of conditional dynamics, a great deal of attention has been paid to time variation in the expected excess return on equities. A number of models in addition to that of CC have emerged that can claim some empirical success along these dimensions. Bansal and Yaron (2004), while working with different preferences due to Epstein and Zin (1989), generate realistic asset pricing dynamics by introducing "long-run risk" and time-varying uncertainty in the consumption growth process. Another recent strand of the literature that also focuses on the technology rather than preferences has rekindled the old Rietz (1990) idea that fear of a large catastrophic event may induce a large equity premium (see Barro (2006)). Importantly, in such "disaster risk" frameworks, there is no time variation in risk premiums unless the probability of the crash is assumed to vary through time, see Wachter (2010).

In this article, we introduce a flexible model for consumption growth, the BEGE model, which, in theory, can accommodate jumps and other non-linearities. Fitting this model to US postwar consumption growth data, the model matches various salient features of recession dynamics that Gaussian models do not match. For example, after a consumption slowdown, the lower tail of consumption growth decreases significantly, and consumption growth also exhibits longer left tails when equity volatility is high and the price-dividend ratio is lower than its unconditional median value. The model also delivers an unconditional left-skewed and leptokurtic distribution where bad consumption growth states are much more likely than under a normal distribution. Yet, these bad states are mostly consistent with persistent recession behavior rather than extreme, relatively transitory disaster shocks, as in the "disaster risk" articles.

The non-Gaussianities in consumption growth deliver first-order variance risk premiums and help match other features of the risk neutral distribution of stock returns, in addition

to the standard salient features of asset returns. There have been some other recent articles attempting to match such dynamics, using alternative mechanisms. Drechsler and Yaron (2011) introduce jumps in the conditional mean of consumption growth in a long-run risk model, but therefore fail to match the non-Gaussian consumption dynamics the BEGE model matches. Likewise, the models in Bollerslev, Tauchen and Zhou (2009; stochastic volatility of the volatility) and Gabaix (2012; disaster risk with a time-varying probability) are not likely to match the recession dynamics that we document.

Our use of a habits model does have some disadvantages. First, even though our agent is less risk averse than an agent not confronted with BEGE dynamics would have to be to match the Sharpe ratio in equity data, she still exhibits implausibly large levels of risk aversion. Second, at least at the quarterly level, there remains a tight link between fundamentals and asset returns that is absent in the data. If we were willing to concede that preference shocks need not necessarily follow the strict parameterization of the CC model and can be imperfectly correlated with fundamentals, we should be able to break both of these implications. In macro-economics it is quite common to accommodate preferences shocks (e.g. Justiniano, Primiceri and Tambalotti, 2009), and in finance, there is, albeit controversial, evidence of sentiment shocks that may not be tightly correlated with fundamentals (Baker and Wurgler, 2007). Building on Bekaert, Engstrom and Grenadier (2010), it is straightforward to formulate a BEGE model where risk aversion is stochastic and heteroskedastic but may not be perfectly correlated with consumption shocks. With the extra flexibility, we can formulate the model such that it delivers closed form solutions for asset prices. We defer a full analysis of such a model to future work.

Appendices

A Calculation of Variance Premiums

1. Generic Linear-Gaussian case

Suppose that m_{t+1} and ret_{t+1} both have linear dependence on a multivariate conditionally Gaussian state vector, X_{t+1}:

$$\begin{aligned} X_{t+1} &\sim N(0, \Sigma_t) \\ m_{t+1} &= m_0 + \beta_m X_{t+1} \\ ret_{t+1} &= r_0 + \beta_r X_{t+1} \end{aligned} \tag{30}$$

We show that under these conditions the physical variance variance of returns equals the risk neutral variance. To begin, note that the moment generating function for returns under the physical measure is defined as

$$\begin{aligned} mgf_t^P(ret_{t+1}; \nu) &= E_t[\exp(\nu\, ret_{t+1})] = E_t[\exp(\nu r_0 + \nu \beta_r' X_{t+1})] \\ &= \exp\left(\nu r_0 + \frac{1}{2}\nu^2 \beta_r' \Sigma_t \beta_r\right). \end{aligned} \tag{31}$$

Calculating the first derivative of $mgf_t^P(ret_{t+1}; \nu)$ with respect to v and evaluating at $v = 0$ yields the conditional mean of returns:

$$E_t[ret_{t+1}] = r_0. \tag{32}$$

Calculating the second derivative of $mgf_t^P(ret_{t+1}; \nu)$ with respect to v and evaluating at $v = 0$ yields the uncentered second moment, $E_t\left[ret_{t+1}^2\right]$. Using the definition of variance, this leads trivially to the expected expression for variance under the physical measure:

$$VAR_t[ret_{t+1}] = E_t\left[ret_{t+1}^2\right] - E_t[ret_{t+1}]^2 = \beta_r' \Sigma_t \beta_r \tag{33}$$

Next, to calculate risk-neutral conditional moments, we need to evaluate the risk-neutral moment generating function, which is defined as:

$$mgf_t^Q(ret_{t+1}; \nu) = \frac{E_t[\exp(m_{t+1} + \nu\, ret_{t+1})]}{E_t[\exp(m_{t+1})]} \tag{34}$$

This simplifies to

$$\begin{aligned}\frac{E_t\left[\exp\left(m_{t+1}+\nu\, ret_{t+1}\right)\right]}{E_t\left[\exp\left(m_{t+1}\right)\right]} &= E_t\left[\exp\left(m_0+\nu r_0+\left(\beta_m'+\nu\beta_r'\right)X_{t+1}\right)\right] \\ &= \exp\left(\nu r_0+\frac{1}{2}\nu^2\beta_r'\Sigma_t\beta_r+\frac{1}{2}\beta_m'\Sigma_t\nu\beta_r+\frac{1}{2}\beta_r'\Sigma_t\nu\beta_m\right) \quad (35)\end{aligned}$$

We proceed, again, by evaluating successive derivatives of $mgf_t^Q\left(ret_{t+1};\nu\right)$ with respect to v and evaluating at $\nu=0$. Eventually, we arrive at:

$$VAR_t^Q\left[ret_{t+1}\right]=\beta_r'\Sigma_t\beta_r \quad (36)$$

so that $VAR_t^Q\left[ret_{t+1}\right]=VAR_t^P\left[ret_{t+1}\right]$.

2. *Quadratic approximation of the CC model*

This computation is also accomplished by evaluating the moment generating function for returns. Suppressing terms that have no random component, which do not affect variances under any measure, returns and the pricing kernel are:

$$\begin{aligned}ret_{t+1} &= \left(\sigma_{dc}+\kappa_1\lambda_t+2\kappa_2\overline{s}_t\lambda_t\right)\sigma_{cc}\varepsilon_{t+1}+\kappa_2\lambda_t^2\sigma_{cc}^2\varepsilon_{t+1}^2+\sigma_{dd}v_{t+1} \\ m_{t+1} &= a_{cc,t}\,\varepsilon_{t+1}\end{aligned}$$

where $a_{cc,t}=-\gamma\sigma_{cc}\left(1+\lambda_t\right)$, the sensitivity of the pricing kernel to the consumption shock under the original CC formulation and $\overline{s}_t=E_t\left[s_{t+1}\right]$. In this case, the moment generating function is the expectation of an exponential-quadratic function of Gaussian random variables. We evaluate this expectation using results from Ang and Liu (2004). Let ϵ be a $K\times 1$ vector where $\epsilon\sim N\left(0,\Sigma\right)$, A a $K\times K$ matrix, and Ω a symmetric $K\times K$ matrix. If $\left(\Sigma^{-1}-2\Sigma\Omega\right)$ is strictly positive definite, then

$$E\left[\exp\left(A\epsilon+\epsilon'\Omega\epsilon\right)\right]=\exp\left(-\frac{1}{2}\ln\det\left(I-2\Sigma\Omega\right)+\frac{1}{2}A'\left(\Sigma^{-1}-2\Omega\right)^{-1}A\right) \quad (37)$$

We need to evaluate first the moment generating function under the physical measure,

$$E_t\left[\exp\left(\nu\, ret_{t+1}\right)\right]=E_t\left[\exp\left(A_1\varepsilon_{t+1}+A_2v_{t+1}+\Omega_{11}\varepsilon_{t+1}^2\right)\right]$$

where $A_1=\nu\left(\sigma_{dc}+\kappa_1\lambda_t+2\kappa_2\overline{s}_t\lambda_t\right)\sigma_{cc}$, $A_2=\nu\sigma_{dd}$ and $\Omega_{11}=\nu\kappa_2\lambda_t^2\sigma_{cc}^2$ and recognizing that $\Omega_{12}=\Omega_{22}=0$ and $\Sigma=I$. This allows us to calculate the moment generating function

for ret_{t+1} under both the physical and risk neutral measures as defined above. Direct computation and lengthy algebra leads to the expression for physical variance shown in the text. For the risk-neutral variance, we repeat the calculation except that we replace A_1 with $A_1' = \nu\left(\sigma_{dc} + \kappa_1\lambda_t + 2\kappa_2\overline{s}_t\lambda_t\right)\sigma_{cc} + a_{cc,t}$. We also require calculation of $E_t\left[\exp\left(m_{t+1}\right)\right] = \frac{1}{2}a_{cc,t}^2$.

3. Linear approximation of the BEGE model

This computation is also accomplished by evaluating the moment generating function for returns. Suppressing terms that have no random component, which do not affect variances under any measure, returns and the pricing kernel are:

$$\begin{aligned} ret_{t+1} &= \left(\sigma_{dp} + \kappa_1\lambda_t\sigma_{cp}\right)\omega_{p,t+1} + \left(-\sigma_{dn} - \kappa_1\lambda_t\sigma_{cn}\right)\omega_{n,t+1} \\ m_{t+1} &= a_{p,t}\omega_{p,t+1} + a_{n,t}\omega_{n,t+1} \end{aligned} \tag{38}$$

In this case, the moment generating function is

$$E_t\left[\exp\left(\nu\, ret_{t+1}\right)\right] = \exp\left(-pf\left(\nu\left(\sigma_{dp} + \kappa_1\lambda_t\sigma_{cp}\right)\right) - nf\left(\nu\left(-\sigma_{dn} - \kappa_1\lambda_t\sigma_{cn}\right)\right)\right) \tag{39}$$

which leads directly to the conditional variance reported in the text. For the risk-neutral moment generating function,

$$\begin{aligned} E_t\left[\exp\left(\nu\, ret_{t+1} + m_{t+1}\right)\right] &= \exp\begin{pmatrix} -pf\left(\nu\left(\sigma_{dp} + \kappa_1\lambda_t\sigma_{cp}\right) + a_{p,t}\right) \\ -nf\left(\nu\left(-\sigma_{dn} - \kappa_1\lambda_t\sigma_{cn}\right)\right) + a_{n,t} \end{pmatrix} \\ E_t\left[\exp\left(m_{t+1}\right)\right] &= \exp\left(-pf\left(a_{p,t}\right) - nf\left(a_{n,t}\right)\right) \end{aligned} \tag{40}$$

which leads directly to the expression for risk-neutral variance that is reported in the text.

4. General Quadratic-Gaussian case

Suppose that m_{t+1} and ret_{t+1} are both quadratic in a multivariate conditionally Gaussian state vector, X_{t+1}:

$$\begin{aligned} X_{t+1} &\sim N(0, \Sigma_t) \\ m_{t+1} &= m_0 + \beta_m X_{t+1} + X'_{t+1} \Omega_m X_{t+1} \\ ret_{t+1} &= r_0 + \beta_r X_{t+1} + X'_{t+1} \Omega_r X_{t+1} \end{aligned} \tag{41}$$

We seek to show that if $\Omega_r = \Omega_m = 0$ then the physical variance variance of returns equals the risk neutral variance. This result is considerably more difficult to derive, but has the same economic interpretation as the linear case: unless the return process or the pricing kernel is conditionally non-Gaussian, the variance premium will be zero. A proof is available upon request from the authors.

B Solving for the sensitivity function

In CC, the expression for the risk free rate is

$$rrf_t = -\ln\beta + \gamma\overline{g} - \gamma(1-\phi)(s_t - \overline{s}) - \frac{\gamma^2\sigma_{cc}^2}{2}\left[1 + \lambda(s_t)\right]^2$$

To eliminate the dependence of rf_t on s_t, which is a necessary condition for the short rate to be constant, one can simply solve the quadratic expression above to yield,

$$\lambda(s_t) = \begin{cases} \frac{1}{\overline{S}}\sqrt{1 - 2(s_t - \overline{s})} - 1 & s_t \leq s_{\max} \\ 0 & s_t \geq s_{\max} \end{cases}$$

where $s_{\max} = \overline{s} + \frac{1}{2}\left(1 - \overline{S}^2\right)$ and $\overline{S} = \exp(\overline{s})$.

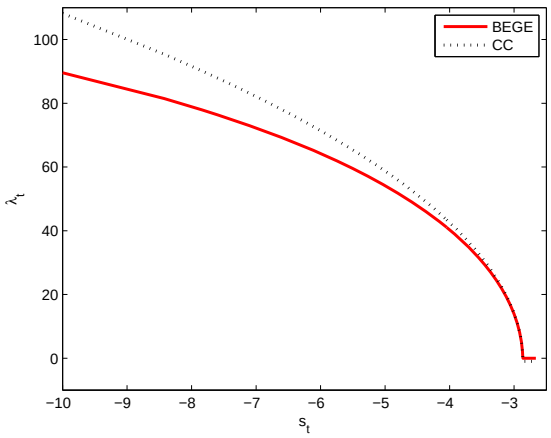
Under the static BEGE specification, the problem of finding $\lambda(s_t)$ is similar. The expression for the risk free rate is

$$rrf_t = -\ln\beta + \gamma\overline{g} + \gamma(1-\phi)(\overline{s} - s_t) + p\ f(a_{p,t}) + n_t\ f(a_{n,t})$$

where, as noted in the text,

$$\begin{aligned} a_{p,t} &= -\gamma\left(1+\lambda_t\right)\sigma_{cp} \\ a_{n,t} &= +\gamma\left(1+\lambda_t\right)\sigma_{cn} \\ f\left(x\right) &= x+\ln\left(1-x\right) \end{aligned}$$

In this case, even if n_t is constant, there is no closed-form solution for $\lambda\left(s_t\right)$ such that the short rate does not depend on s_t. However, this value is easily evaluated numerically. In the below figure, the red solid line shows the values for $\lambda\left(s_t\right)$ for the BEGE(s1) model. For comparison, the dashed black line shows the values for $\lambda\left(s_t\right)$ under the Gaussian CC model. In this exercise, all of the model parameters are held identical in the two cases, as is the volatility of consumption growth. The only difference is the assumed distribution for consumption growth.



Note that, for more negative values of s_t, the BEGE sensitivity function is substantially lower than its Gaussian counterpart. In the case of dynamic BEGE, the sensitivity variable depends on both state variables, $\lambda = \lambda\left(s_t, n_t\right)$, but the numerical solution procedure is identical.

C The Maximum Sharpe Ratio

We start from the general expression for the maximum available Sharpe ratio:

$$\begin{aligned}\max_{\{\text{all assets}\}} \frac{E_t\left[R^e_{t+1}\right]}{\sigma_t\left[R^e_{t+1}\right]} &= \frac{\sigma_t\left[\exp\left(m_{t+1}\right)\right]}{E_t\left[\exp\left(m_{t+1}\right)\right]} = \sqrt{\frac{E_t\left[\exp\left(2m_{t+1}\right)\right]}{E_t\left[\exp\left(m_{t+1}\right)\right]^2} - 1} \\ &= \sqrt{\left[\frac{1-a_{p,t}}{\left(1-2a_{p,t}\right)^{1/2}}\right]^{2p}\left[\frac{1-a_{n,t}}{\left(1-2a_{n,t}\right)^{1/2}}\right]^{2n} - 1} \qquad (42)\end{aligned}$$

where, as defined above,

$$\begin{aligned}a_{p,t} &= -\gamma\left(1+\lambda_t\right)\sigma_{cp} \\ a_{n,t} &= +\gamma\left(1+\lambda_t\right)\sigma_{cn} \qquad (43)\end{aligned}$$

This expression is fairly messy. Some approximations are helpful for gaining intuition

$$\begin{aligned}\frac{\sigma_t\left[\exp\left(m_{t+1}\right)\right]}{E_t\left[\exp\left(m_{t+1}\right)\right]} &= \sqrt{\exp\left\{2p\ln\left[\frac{1-a_{p,t}}{\left(1-2a_{p,t}\right)^{1/2}}\right] + 2n\ln\left[\frac{1-a_{n,t}}{\left(1-2a_{n,t}\right)^{1/2}}\right]\right\} - 1} \\ &\approx \sqrt{2p\ln\left[\frac{1-a_{p,t}}{\left(1-2a_{p,t}\right)^{1/2}}\right] + 2n\ln\left[\frac{1-a_{n,t}}{\left(1-2a_{n,t}\right)^{1/2}}\right]} \\ &\approx \sqrt{p\left(a^2_{p,t}+2a^3_{p,t}\right) + n\left(a^2_{n,t}+2a^3_{n,t}\right)} \qquad (44)\end{aligned}$$

D The Bates Filter

Bates (2006) develops a direct filtration-based maximum likelihood methodology for estimating the parameters and realizations of latent affine processes. Filtration is conducted using characteristic functions and employing a version of Bayes' rule for recursively updating the joint characteristic function of latent variables and the data conditional upon past data.

For the dynamic BEGE model:

$$\begin{aligned}
\Delta c_{t+1} &= g + \sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1} \\
n_{t+1} &= (1-\rho_n)\,\overline{n} + \rho_n n_t + \rho_n \sigma_{nn}\omega_{n,t+1} \\
\omega_{p,t+1} &\sim \widetilde{\Gamma}(p,1) \\
\omega_{n,t+1} &\sim \widetilde{\Gamma}(n_t,1)
\end{aligned} \tag{45}$$

the key required condition is that the conditional joint characteristic function of Δc_{t+1} and n_{t+1} be an affine exponential function of n_t. That is, the Bates filter requires that

$$E_t\left[\exp\left(i\phi\Delta c_{t+1} + i\theta n_{t+1}\right)\right] = \exp\left(\psi + \delta n_t\right) \tag{46}$$

for any ϕ and θ and where the parameters ψ and δ are generally functions of the model parameters. Fortunately, the gamma distributions on which the BEGE process is based admit such a representation. Straightforward calculations starting from the characteristic function of the gamma distribution lead to:

$$\begin{aligned}
\psi &= i\phi g + i\theta\left(1-\rho_n\right)\overline{n} - pf\left(i\phi\sigma_{cp}\right) \\
\delta &= i\theta\rho_n - f\left(-i\phi\sigma_{cn} + i\theta\rho_n\sigma_{nn}\right)
\end{aligned} \tag{47}$$

To estimate the model, we follow the steps provided in Bates (2006). That the procedure only approximates maximum likelihood owes to the fact that in each step of the filtration, the posterior density for the latent process is assumed to adhere to a particular parametric density function, which may not be true. At each step, we assume that the posterior distribution for n_t is gamma, with the parameters of the conditional posterior distribution being informed by the filtered conditional moments for n_t. This assumption is recommended by Bates (2006) for a latent variable that is assumed to be strictly positive, as is the case for n_t.

The procedure for using the Bates filter for the Gaussian stochastic volatility model is analogous.

E Numerical Solution for Equity Prices

We closely follow the procedure of Wachter (2005). We first note that the price dividend ratio is the present value of future discounted dividends:

$$\frac{P_t}{D_t} = E_t \left[\sum_{i=1}^{\infty} \exp \left(\sum_{j=1}^{i} \left(m_{t+j} + \Delta d_{t+j} \right) \right) \right] \tag{48}$$

Because the surplus ratio formulation contains a significant non-linearity, no closed-form solutions are available for this expectation beyond the $i = 1$ term. To implement a numerical solution of Equation (48), we recursively solve for the individual terms in the summation. That is,

$$\frac{P_t}{D_t} = \sum_{i=1}^{\infty} F_i \left(s_t, n_t \right) \tag{49}$$

where

$$F_i \left(s_t, n_t \right) = E_t \left[\exp \left(\sum_{j=1}^{i} \left(m_{t+j} + \Delta d_{t+j} \right) \right) \right] \tag{50}$$

and

$$F_i \left(s_t, n_t \right) = E_t \left[\exp \left(m_{t+1} + \Delta d_{t+1} \right) F_{i-1} \left(s_{t+1}, n_{t+1} \right) \right] \tag{51}$$

The latter recursion is solved numerically on a grid suggested by Wachter. In Wachter and our static models, the only state variable is s_t. However, under our dynamic framework, prices depend on both s_t and n_t As such, we define a two-dimensional grid over these variables. Let $\{s_0, n_0\}$ represent the grid. We define the grid for s_0 spanning the range of s_t in a simulation of $100,000$ periods, and using 400 intervals. For n_t we derive grid values using the same procedure. For each point on the resulting two dimensional grid, we find λ_0, the sensitivity parameter, that sets the short term risk free rate to the desired constant value, although, as noted above and as in CC and Wachter, for some extreme points on the grid, this short rate cannot be obtained. For these points on the grid, we set $\lambda_0 = 0$.

To start the algorithm, we calculate $F_1 \left(s_0, n_0 \right) = E \left[\exp \left(m + \Delta d \right) | \left(s_0, n_0 \right) \right]$. This can be evaluated analytically in the BEGE framework as it is just the expectation of an

exponential-affine function of the state vector.

To continue the sequence, we must rely on numerical methods. Let $\{s_1, n_1\}$ denote the set of random values that s and n may take on following $\{s_0, n_0\}$. Under the model, this set is continuous and unbounded, a function of the realizations of $\omega_{n,1}$ and $\omega_{p,1}$, which, the reader will recall, are independent. For ease of numerical integration, we discretize the distributions $\omega_{n,1}$ and $\omega_{p,1}$. Specifically, we identify a set of 10 points, denoted $\widehat{\omega}_{n,1}(n_0)$ and associated probabilities, $p(\widehat{\omega}_{n,1}(n_0))$ such that this discrete distribution exactly matches the first 10 moments of the continuous distribution for $\omega_{n,1}$, and similarly for ω_p although the distribution of $\omega_{p,1}$ is invariant across the grid, since p is constant. This yields associated finite sets for all of the other endogenous variables: :

$$\begin{aligned}
\Delta\widehat{c_1} &= \overline{g} + \sigma_{cp}\widehat{\omega}_{p,1} - \sigma_{cn}\widehat{\omega}_{n,1} \\
\Delta\widehat{d_1} &= \overline{g} + \sigma_{dp}\widehat{\omega}_{p,1} - \sigma_{dn}\widehat{\omega}_{n,1} \\
\widehat{s_1} &= (1-\phi)\,\overline{s} + \phi s_0 + \lambda_0\left(\Delta\widehat{c_1} - \overline{g}\right) \\
\widehat{n_1} &= (1-\rho_n)\,\overline{n} + \rho_n n_0 + \rho_n\sigma_{nn}\widehat{\omega}_{n,1} \\
\widehat{m_1} &= \ln(\beta) - \gamma\Delta\widehat{c_1} - \gamma\left(\widehat{s_1} - s_0\right)
\end{aligned} \tag{52}$$

Now, we can evaluate $F_2(s_0, n_0)$ using,

$$\begin{aligned}
F_2(s_0, n_0) &= E\left[\exp(m_1 + \Delta d_1)\, F_1(s_1, n_1) \mid (s_0, n_0)\right] \\
&\approx \sum_{\widehat{\omega}_{p,1}}\sum_{\widehat{\omega}_{n,1}} \exp\left(\widehat{m_1} + \Delta\widehat{d_1}\right) F_1\left(\widehat{s_1}, \widehat{n_1}\right) p\left(\widehat{\omega}_{p,1}\right) p\left(\widehat{\omega}_{n,1}\right)
\end{aligned} \tag{53}$$

This involves many calculations. Our base grid, $\{s_0, n_0\}$ has 16,000 points, every one of which must be operated on every one of the 100 points on the $\{\widehat{\omega}_{p,1}, \widehat{\omega}_{n,1}\}$ grid. Still, these are straightforward calculations except for the term, $F_1\left(\widehat{s_1}, \widehat{n_1}\right)$, which is complicated by the fact that the only previously evaluated values for F_1 are on the grid, (s_0, n_0). Following Wachter and CC, we use a (log-linear) interpolation from the points $F_1(s_0, n_0)$ to evaluate the points $F_1\left(\widehat{s_1}, \widehat{n_1}\right)$. Of course, not all of the points in $\{\widehat{s_1}, \widehat{n_1}\}$ lie in the interior of $\{s_0, n_0\}$, so some points must be extrapolated, not just interpolated. Wachter and CC do this also. We then iterate, solving for $F_i(s_0, n_0)$ for $i = 3$ through 1000, after which we find that the

contribution to prices is negligibly small.

References

Ang, A., Liu, J., 2004, How to Discount Cashflows with Time-Varying Expected Returns, Journal of Finance, 59(6), 2745-2783.

Baker, M, Wurgler, J., 2007, Investor Sentiment in the Stock Market, Journal of Economic Perspectives, 21(2), 129-152.

Bakshi, G. and D. Madan, 2006, A Theory of Volatility Spreads, Management Science, 52, 1945-1956.

Bakshi, G. Kapadia, N., and D. Madan, 2003, Stock Return characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options, Review of Financial Studies, 16, 101-143.

Bansal, R., Yaron, A., 2004, Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles, Journal of Finance, 59, 1481.

Barro, R.J., 2006, Rare Disasters and Asset Markets in the Twentieth Century, Quarterly Journal of Economics, 121, 823-866.

Bates, David S., Maximum likelihood estimation of latent affine processes, Review of Financial Studies, Fall 2006, 909-965.

Bekaert, G., Engstrom, E., Grenadier, S., 2010, Stock and Bond Returns with Moody Investors, Journal of Empirical Finance, 17, 867-894.

Bollerslev, T., Tauchen, G. E., and Zhou, H., 2009, Expected Stock Returns and Variance Risk Premia, Review of Financial Studies, 22, 4463-4492.

Bollerslev, T. Gibson, M., and H. Zhou, 2011, Dynamic Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities, Journal of Econometrics, 160, 235-245.

Campbell, J. Y., Cochrane, J. H., 1995, By Force of Habit: A consumption based explanation of aggregate stock market behavior, NBER working paper.

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

Carr, P., Wu, L., 2008, Variance Risk Premiums, Review of Financial Studies, 22, 1311-1341.

Corsi, F., 2009, A Simple Approximate Long-Memory Model of Realized Volatility, Journal of Financial Econometrics, 7, 174-196.

Drechsler, I., and Yaron, A., 2011, What's Vol Got to Do With It, Review of Financial Studies, 24, 1-45.

Epstein, L.G., Zin, S.E., 1989, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, Econometrica, 57(4), 937-969.

Figlewski, S., 2008, Estimating the Implied Risk-Neutral Density of the U.S. Market Portfolio, Volatility and Time Series Econometrics (eds. T. Bollerslev, J. Russell, and M. Watson), Oxford University Press.

Gabaix, X., 2012, Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance, Quarterly Journal of Economics, 127(2), 645-700.

Glosten, L., Jagannathan, R., and D. Runkle, 1993, On the Relation Between the Expected Value and the Volatility of Nominal Excess Returns on Stocks, Journal of Finance, 48, 1779-1801.

Jarque, C., Bera, A., 1987, A test for normality of observations and regression residuals,. International Statistical Review 55(2), 163-172.

Justiniano, A., Primiceri, G., Tambalotti, A., 2010, Investment Shocks and Business Cycles, Journal of Monetary Economics, 57(2), 132-145.

Newey, W., West, K., 1987, A Simple Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, Econometrica, 55, 703-708.

Rietz, T.A., 1988, The Equity Risk Premium: A Solution, Journal of Monetary Economics, 22, 117-131.

Wachter, J., 2005, Solving models with external habit, Finance Research Letters 2:210-226.

Wachter, J., 2006, A consumption-based model of the term structure of interest rates, Journal of Financial Economics 79:365-399.

Wachter, J., 2010, Can Time-varying Risk of Rare Disasters Explain Stock Market Volatility? working Paper.

Working, H., 1960, Note on the correlation of first differences of averages in a random chain, Econometrica, 28, October 1960, 916-918.

Wu, G., 2001, The Determinants of Asymmetric Volatility, Review of Financial Studies, 14, 837-859

Table 1: Maximum likelihood estimation of static models of consumption growth

Gauss(s)		BEGE(s1)	
g	0.0047	g	0.0047
	(0.0003)		(0.0003)
σ	0.0049	$\overline{p}$	3.27
	(0.0002)		(2.29)
		$\overline{n}$	0.73
			(0.54)
		σ_{cp}	0.0018
			(0.0006)
		σ_{cn}	0.0043
			(0.0018)
$loglike$	1005.5		1017.1
$JBpval$	0.0010		0.5000

Notes: Statistics for real, per capita, quarterly consumption growth in for U.S. nondurables and services from 1947Q2-2011Q4. Data are pre-whitened to remove time-aggregation effects using an AR(1) model. The left column shows ML parameter estimates for an i.i.d. Gaussian model, and the right column shows ML parameter estimates for an i.i.d. BEGE model:

$$\begin{aligned} \Delta c_t &= \sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1} \\ \omega_{p,t+1} &\sim \widetilde{\Gamma}\left(\overline{p},1\right), \qquad \omega_{n,t+1} \sim \widetilde{\Gamma}\left(\overline{n},1\right) \end{aligned}$$

where $\widetilde{\Gamma}\left(a,b\right)$ denotes the centered gamma distribution with shape parameter a and scale parameter b. The row $JBpval$ presents the p-value from a specification test. Specifically, the p-value is from a Jarque-Bera test for normality of the transformed series, $\widehat{x}_t$, where

$$\widehat{x}_t = \Phi_G^{-1}\left[\Phi_{\text{mod}}\left(x_t\right)\right]$$

where Φ_G^{-1} denotes the inverse cdf function for the standard Gaussian distribution, and Φ_{mod} denotes the cdf under the model for each column. For the left column,Φ_{mod} is $N\left(g,\sigma\right)$, and for the right column it is, $BEGE\left(g,\overline{p},\overline{n},\sigma_{cp},\sigma_{cn}\right)$.

Table 2: Approximate maximum likelihood estimation of stochastic factor models of consumption growth

Gauss(d)		BEGE(d)	
g	0.0048	g	0.0048
	(0.0003)		(0.0003)
		$\overline{p}$	4.05
			(4.09)
		$\overline{n}$	1.98
			(1.70)
σ_{cs}	0.0039	σ_{cp}	0.0014
	(0.0005)		(0.0007)
σ_{cv}	0.0022	σ_{cn}	0.0029
	(0.0011)		(0.0012)
ρ_v	0.91	ρ_n	0.93
	(0.04)		(0.03)
σ_{vv}	0.51	σ_{nn}	0.36
	(0.49)		(0.21)
appx loglike	1004.9		1030.3
JBpval	0.0027		0.5000

Notes: Statistics for real, per capita, quarterly consumption growth in for U.S. nondurables and services from 1947Q2-2011Q4. Consumption data are pre-whitened to remove time-aggregation effects using an AR(1) model. The left column shows parameter estimates for the model,

$$\begin{aligned} \Delta c_{t+1} &= g + \sigma_{cs}\varepsilon_{t+1}^1 - \sigma_{cv}\sqrt{\overline{v}_t}\varepsilon_{t+1}^2 \\ v_t &= \overline{v}_{t-1} + \sigma_{vv}\sqrt{\overline{v}_{t-1}}\varepsilon_{t+1}^2 \\ \overline{v}_{t-1} &= \overline{v} + \rho_v\left(\overline{v}_{t-1} - \overline{v}\right) \end{aligned}$$

where ε_t^1 and ε_t^1 are independent $N(0,1)$. The right column shows parameter estimates for the BEGE model:

$$\begin{aligned} \Delta c_t &= g + \sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1} \\ \omega_{p,t+1} &\sim \widetilde{\Gamma}\left(\overline{p},1\right), \qquad \omega_{n,t+1} \sim \widetilde{\Gamma}\left(\overline{n_t},1\right) \\ n_t &= \overline{n}_{t-1} + \sigma_{nn}\omega_{n,t+1} \\ \overline{n}_{t-1} &= \overline{n} + \rho_n\left(n_{t-1} - \overline{n}\right) \end{aligned}$$

where $\widetilde{\Gamma}(a,b)$ denotes the centered gamma distribution with shape parameter a and scale parameter b. The row $JBpval$ presents the p-value from a specification test. Specifically, the p-value is from a Jarque-Bera test for normality of the transformed series, $\widehat{x}_t$, where$\widehat{x}_t = \Phi_G^{-1}\left[\Phi_{\text{mod}}\left(x_t\right)\right]$, Φ_G^{-1} denotes the inverse cdf function for the standard Gaussian distribution, and Φ_{mod} denotes the cdf of the model for Δc_{t+1} for each column. For the left column, $N\left(g, \sigma_{cs}^2 + \sigma_{cv}^2 v_t\right)$, and for the right column, $BEGE\left(g, \overline{p}, \overline{n_t}, \sigma_{cp}, \sigma_{cn}\right)$.

Table 3: Maximum likelihood estimation of GJR-GARCH models of consumption growth

Panel A

Gauss(d^{GJR})		BEGE(d^{GJR})	
g	0.0048	g	0.0048
	(0.0003)		(0.0003)
		$\overline{p}$	4.01
			(2.29)
		n_0	0.0086
			(0.0712)
σ_{cs}	0.0013	σ_{cp}	0.0015
	(0.0023)		(0.0006)
σ_{cv}	0.0011	σ_{cn}	0.0029
	(0.0005)		(0.0012)
ρ_v	0.85	ρ_n	0.78
	(0.02)		(0.07)
ϕ_{v1}	25.4	ϕ_{n1}	9.30
($\times$0.0001)	(38.2)	($\times$0.0001)	(9.38)
ϕ_{v2}	81.1	ϕ_{n2}	10.0
($\times$0.0001)	(75.2)	($\times$0.0001)	(12.0)
loglike	1026.6		1033.0
JBpval	0.0050		0.5000

Panel B

Gauss(d^{alt})			BEGE(d^{alt})		
$\overline{v}$	ρ_v	σ_{vv}	$\overline{n}$	ρ_n	σ_{nn}
19.6	0.97	0.88	1.87	0.95	0.54
(12.8)	(0.10)	(3.56)	(1.43)	(0.06)	(0.27)

Notes: Statistics for real, per capita, quarterly consumption growth in for U.S. nondurables and services from 1947Q2-2011Q4. Consumption data are pre-whitened to remove time-aggregation effects using an AR(1) model. In Panel A, the left column shows parameter estimates for the model,

$$\begin{aligned} \Delta c_{t+1} &= g + \sigma_{cs}\varepsilon^1_{t+1} - \sigma_{cv}\sqrt{v_t}\varepsilon^2_{t+1} \\ v_t &= v_0 + \rho_v v_{t-1} + \phi_{v1}\ (\Delta c_t - g)^2 + \phi_{v2}\ (\Delta c_t - g)^2 \cdot I_{\Delta c<g} \end{aligned}$$

where ε_t^1 and ε_t^1 are$N(0,1)$. The right column shows parameter estimates for the BEGE model:

$$\begin{aligned}
\Delta c_t &= g + \sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1} \\
\omega_{p,t+1} &\sim \widetilde{\Gamma}(\overline{p},1), \qquad \omega_{n,t+1} \sim \widetilde{\Gamma}(n_t,1) \\
n_t &= n_0 + \rho_r n_{t-1} + \phi_{n1}\ (\Delta c_t - g)^2 + \phi_{n2}\ (\Delta c_t - g)^2 \cdot I_{\Delta c<g}
\end{aligned}$$

where $\widetilde{\Gamma}(a,b)$ denotes the centered gamma distribution with shape parameter a and scale parameter b. The row $JBpval$ presents the p-value from a specification test. Specifically, the p-value is from a Jarque-Bera test for normality of the transformed series, $\widehat{x}_t$, where

$$\widehat{x}_t = \Phi_G^{-1}\left[\Phi_{\text{mod}}(x_t)\right]$$

where Φ_G^{-1} denotes the inverse cdf function for the standard Gaussian distribution, and Φ_{mod} denotes the cdf of the model for Δc_{t+1} for each column. For the left column, $N(g, \sigma_{cs}^2 + \sigma_{cv}^2 v_t)$, and for the right column, $BEGE(g, \overline{p}, n_t, \sigma_{cp}, \sigma_{cn})$.

Panel B reports on the estimated properties for the latent factor for each model. For the Gaussian model, the unconditional mean, autocorrelation, and volatility of the fitted v_t process are used to estimate the parameters for the model

$$v_t = \overline{v} + \rho_v(v_{t-1} - \overline{v}) + \sigma_{vv}\sqrt{v_{t-1}}\varepsilon_{t+1}^v; \quad \varepsilon_{t+1}^v \sim N(0,1)$$

For the BEGE model, the unconditional mean, autocorrelation, and volatility of the fitted n_t process are used to estimate the promoters for the BEGE model estimated in Table 2.

Table 4: Summary of models: unconditional moments of consumption and dividend growth

Panel A, Quarterly data

	sample	Gauss(s)	BEGE(s1)	BEGE(s2)	Gauss(d)	BEGE(d)
$vol\,(\Delta c)$	0.0098^{***}	0.0098	0.0098	0.0098	0.0098	0.0096
	(0.0010)					
$skw\,(\Delta c)$	-0.32^{**}	0	−0.60	−0.32	0	−0.57
	(0.18)					
$xkur\,(\Delta c)$	1.53^{***}	0	2.83	1.53	0.66	2.48
	(0.46)					
$vol\,(\Delta d)$	0.091^{***}	0.091	0.091	0.091	0.091	0.092
	(0.0028)					
$skw\,(\Delta d)$	−0.07	0	−1.26	−1.12	0	−0.87
	(0.28)					
$xkur\,(\Delta d)$	4.18^{***}	0	2.61	2.24	1.57	1.81
	(1.58)					

Panel B, Annual data

	sample	Gauss(s)	BEGE(s1)	BEGE(s2)	Gauss(d)	BEGE(d)
$vol\,(\Delta c)$	0.0090^{***}	0.0080	0.0081	0.0081	0.0083	0.0081
	(0.0014)					
$skw\,(\Delta c)$	-0.41^{*}	0.00	−0.32	−0.14	−0.28	−0.67
	(0.28)					
$xkur\,(\Delta c)$	0.33	0.05	0.66	0.28	0.56	1.97
	(0.77)					
$vol\,(\Delta d)$	0.0600^{***}	0.076	0.075	0.076	0.075	0.076
	(0.0083)					
$skw\,(\Delta d)$	−0.78	0.02	−0.59	−0.57	−0.71	−0.72
	(0.58)					
$xkur\,(\Delta d)$	2.94	0.00	0.64	0.53	1.90	1.40
	(1.71)					

Notes. In the column to the left of the vertical line, sample statistics with block-bootstrapped standard errors are reported in parentheses. One, two, and three asterisks denote significance at the 10, 5, and 1 percent levels, respectively.

The first two columns to the right of the vertical line report simulated results for the properties of consumption growth under the models estimated in Table 1. The third column, labeled "BEGE(s2)" reports statistics for an additional "BEGE-static" calibration designed to match exactly the sample volatility, skewness, and kurtosis of consumption growth. The calibrated parameters of this model are: $\overline{p} = 5.760$, $\overline{n} = 1.1090$, $\sigma_{cp} = 0.00147$, and $s_{cn} = 0.00322$. The columns labeled "Gauss(d)" and "BEGE(d)" report on the properties of consumption and dividends under the stochastic factor models of Table 2. Volatility statistics are reported at an annual rate.

Table 5: The conditional distribution of quarterly consumption growth

		sample	Gauss(d)	BEGE(d)
Slowdown statistics				
freq		0.18***	0.15	0.15
		(0.03)		
avg. dur.		3.83***	2.37	2.56
		(0.61)		
Conditional distribution				
	no $(t-1)$ *slowdown*	−0.0089*	−0.0096	−0.0117
		(0.0050)		
q^5	$(t-1)$ *slowdown*	−0.0281***	−0.0134	−0.0257
		(0.0058)		
	$difference$	−0.0192***	−0.0038	−0.0140
		(0.0075)		
	no $(t-1)$ *slowdown*	0.0192***	0.0192	0.0193
		(0.0016)		
q^{50}	$(t-1)$ *slowdown*	0.0160	0.0192	0.0207
		(0.0060)		
	$difference$	−0.0033	0.0000	0.0013
		(0.0060)		
	no $(t-1)$ *slowdown*	0.0486	0.0481	0.0468
		(0.0042)		
q^{95}	$(t-1)$ *slowdown*	0.0523	0.0516	0.0563
		(0.0064)		
	$difference$	0.0036	0.0036	0.0095
		(0.0073)		

Notes. All panels report statistics for the distribution of real, per capita, quarterly consumption growth in for U.S. nondurables and services from 1947Q2-2011Q4. Consumption growth data are annualized and pre-whitened to remove time-aggregation effects using an AR(1) model. Sample results are reported in the left column. The upper rows in each section report on the distribution conditional on being not being in a slowdown in the previous period, with bootstrapped standard errors below. The second row reports on the distribution conditional on being in a slowdown in the previous period. The third row reports on the difference in the statistics reported in the top two rows. q^5, q^{50}, and q^{95} denote the 5th, 50th, and 95th percentiles of the distribution, respectively. One, two, and three asterisks denote statistical significance at the 10 5, and 1 percent level.

The column headed, "Gauss(d)" reports results for the dynamic Gaussian model estimated in Table 2. The finalcolumn reports results from the "BEGE(d)" model in Table 2.

Table 6: Asset pricing model calibration and implied risk aversion

Panel A, parameterization

Model	Gauss(s)	BEGE(s1)	BEGE(s2)	BEGE(d)	BEGE(d)'
Δc *dynamics*	Gauss(s)	BEGE(s1)	BEGE(s2)	BEGE(d)	BEGE(d)
γ	2.8	1.7	1.9	1.9	8.0
ϕ	0.97	0.97	0.97	0.97	0.99

Panel B, distribution of utility curvature

Model	Gauss(s)	BEGE(s1)	BEGE(s2)	BEGE(d)	BEGE(d)'
5%	41.2	30.6	32.1	32.3	43.9
25%	47.6	34.5	36.2	36.8	50.0
50%	57.9	42.0	44.2	44.9	56.6
75%	77.7	59.2	62.1	62.6	66.1
95%	148.2	137.1	139.7	136.5	85.8

Panel B, distribution of risk aversion

Model	Gauss(s)	BEGE(s1)	BEGE(s2)	BEGE(d)	BEGE(d)'
5%	171.8	87.3	102.5	91.6	107.5
25%	203.7	96.3	114.5	102.5	123.4
50%	258.1	114.1	138.1	122.6	140.9
75%	369.9	156.1	194.0	167.8	166.5
95%	819.1	359.2	459.5	370.5	219.7

All models use CC preferences and are calibrated by varying γ to match the unconditional Sharpe ratio of equity returns in quarterly data from 1947-2011 (0.35). Model BEGE(d)' is further calibrated by adjusting ϕ to match the unconditional variance of equity returns. Panel B reports on the unconditional distribution of local curvature of utility for each model.

$$local\ curvature_t\ = \gamma \exp\left(-s_t\right)$$

Panel C reports on the unconditional distribution of global risk aversion for models:

$$risk\ aversion\ = local\ curvature\ \cdot \frac{PC_t}{PCS_t}$$

where PC_t the price-dividend ratio of the asset with dividend equal to the consumption stream, and PCS_t is the price-dividend ratio of an asset with dividends equal to $C_t S_t$.

Table 7: Effects of shifts in s_t and n_t under the BEGE(d)' model

Panel A, Endogenous variables

Variable	BEGE(d)'		
	baseline	*low s*	*high n*
λ_t	6.59	8.71	4.22
$a_{p,t}$	−0.085	−0.109	−0.059
$a_{n,t}$	0.170	0.218	0.117
$E_t[m_{t+1}]$	−0.040	−0.065	−0.040
$VAR_t^P[m_{t+1}]$	0.070	0.115	0.073
$VAR_t^Q[m_{t+1}]$	0.084	0.148	0.088
$MaxSharpe_t$	0.299	0.406	0.304
$\sigma^P[\Delta d_{t+1}]$	0.085	0.085	0.117
$\sigma^Q[\Delta d_{t+1}]$	0.090	0.092	0.127

Panel B, Asset prices

Variable	BEGE(d)'		
	baseline	*low s*	*high n*
pd_t	3.13	2.83	3.04
$E_t[xret_{t+1}]$	0.051	0.079	0.075
$pvol_t$	0.15	0.16	0.20
$qvol_t$	0.18	0.20	0.23
$vprm_t$	0.03	0.04	0.03

This table reports the values of various endogenous variables and asset prices at three points in the state space:

variable	*baseline*	*low s*	*high n*
s_t	−3.2	−4.0	−3.2
n_t	1.4	1.4	4.3

The values in the above table reflect the 50th and 10th percentiles of the unconditional values of the s_t distribution, and the 50th and 90th percentile values of the distribution of n_t in simulated data. In Panel A, λ_t is the sensitivity function. In panel B, $E_t[xret_{t+1}]$ is the conditional one-period ahead equity risk premium, expressed at an annual rate. All volatilities are reported at an annual rate.

Table 8: Performance of Models: Univariate Asset Price Statistics

	Sample	Model-Implied				
	stats	Gauss(s)	BEGE(s1)	BEGE(s2)	BEGE(d)	BEGE(d)'
$E[xret_t]$	0.055	0.045	0.060	0.055	0.073	0.056
	(0.021)					
$\sigma[xret_t]$	0.16	0.13	0.18	0.17	0.21	0.16
	(0.010)					
$skew[xret_t]$	−0.92	0.11	−1.74	−1.32	−0.87	−1.10
	(0.15)					
$xkurt[xret_t]$	1.71	−0.03	5.94	3.81	3.27	2.87
	(0.63)					
$E[pd_t]$	3.44	3.31	3.01	3.03	2.63	3.05
	(0.14)					
$\sigma[pd_t]$	0.43	0.14	0.27	0.25	0.31	0.26
	(0.09)					
$acorr[pd_t]$	0.98	0.96	0.97	0.96	0.96	0.98
	(0.03)					
$E[pvol_t]$	0.14	0.13	0.17	0.16	0.22	0.15
	(0.007)					
$\sigma[pvol_t]$	0.047	0.01	0.036	0.032	0.092	0.043
	(0.007)					
$E[vprm_t]$	0.052	0.00	0.038	0.026	0.028	0.027
	(0.016)					
$\sigma[vprm_t]$	0.027	0.00	0.026	0.017	0.017	0.008
	(0.011)					
$E[qskew_t]$	−1.71	−0.03	−2.20	−1.77	−1.73	−1.76
	(0.15)					
$\sigma[qskew_t]$	1.37	0.02	0.14	0.12	0.61	0.67
	(0.24)					
$E[qxkur_t]$	3.47	0.03	6.77	4.46	4.78	4.94
	(0.62)					
$\sigma[qxkur_t]$	1.53	0.01	0.88	0.58	4.06	4.26
	(0.26)					

The left-most column reports sample statistics and block-bootstrapped standard errors. $xret_t$, pd_t, $pvol_t$, and $vprm_t$ refer to excess equity returns, the log price-dividend ratio, physical conditional volatility (annualized), and the volatility premium (annualized). All statistics are reported at an annual rate, except autocorrelation, which is reported for the quarterly frequency. The columns to the right of the vertical line report simulated values for the asset price statistics under various models. Columns under the heading

"Campbell-Cochrane preferences" use the preference specification from Campbell and Cochrane (1999),

$$\begin{aligned} m_{t+1} &= \ln(\beta) - \gamma \Delta c_{t+1} - \gamma \Delta s_{t+1} \\ s_{t+1} &= \overline{s} + \rho_s (s_{t+1} - \overline{s}) - \lambda_t (\Delta c_{t+1} - g) \\ \lambda_t &: \quad rf_t = -\ln(E_t[\exp(m_{t+1})]) = 0.0025 \end{aligned}$$

and the parameter, γ, is chosen to match the sample unconditional Sharpe ratio for equity.

The column subheadings denote the consumption technology for each model. "Gauss(s)" refers to i.i.d. Gaussian consumption growth using the parameters in Table 3. "BEGE(s1) uses the BEGE consumption process estimated in Table 1. "BEGE(s2)" uses the alternative static BEGE model as reported in Table 4. "BEGE(d)" refers to models using the dynamic BEGE consumption process presented in Table 2.

Table 9: Performance of Models: Cross-moments between asset prices and fundamentals

	Sample Stats	Model-Implied Gauss(s1)	BEGE(s1)	BEGE(s2)	BEGE(d1)	BEGE(d)'
Panel A, correlations in quarterly data						
$[\Delta c_t, xret_t]$	0.18	0.70	0.82	0.81	0.83	0.70
	(0.06)					
$[\Delta d_t, xret_t]$	0.04	0.82	0.63	0.67	0.64	0.81
	(0.04)					
$[\Delta c_t, \Delta d_t]$	0.09	0.20	0.20	0.20	0.20	0.20
	(0.07)					
Panel B, correlations in annual data						
$[\Delta c_t, xret_t]$	0.19	0.45	0.51	0.46	0.51	0.47
	(0.13)					
$[\Delta d_t, xret_t]$	0.18	0.61	0.48	0.45	0.44	0.57
	(0.14)					
$[\Delta c_t, \Delta d_t]$	0.35	0.20	0.20	0.04	0.19	0.19
	(0.11)					
Panel C, the price-dividend ratio and the conditional distribution of Δc						
Δq^5	-0.0245^{**}	0	0	0	-0.0121	-0.0109
	(0.0080)					
Δq^{50}	0.0037	0	0	0	0.0014	0.0010
	(0.0027)					
Δq^{95}	-0.0003	0	0	0	0.0076	0.0067
	(0.0079)					
Panel D, conditional equity volatility and the conditional distribution of Δc						
Δq^5	-0.0197^{*}	0	0	0	-0.0144	-0.0153
	(0.0080)					
Δq^{50}	-0.0045	0	0	0	0.0019	0.0021
	(0.0033)					
Δq^{95}	-0.0068	0	0	0	0.0090	0.0095
	(0.0061)					
Panel E, asymmetric volatility: $corr\,[xret_t, pvol_t]$						
	-0.38^{***}	-0.05	-0.07	-0.08	-0.14	-0.21
	(0.05)					

The left-most column reports sample correlation statistics and block-bootstrapped standard errors. The columns to the right of the vertical line report simulated values for the asset price statistics under various models. In panels C and D, quantile moment shifts are reported. In Panel C, these are defined as the shift in the conditional quantile measures of the distribution of consumption growth when the price-dividend ratio of equity is below its unconditional median value. In Panel D, reported are the shifts in quantile measures when the conditional volatility is above its unconditional median value. In Panels C and D, one, two, and three asterisks denote statistical significance at the 10, 5, and 1 percent level.

Table 10: Performance of models: the conditional distribution of equity returns

Panel A: Predictability of excess equity returns w.r.t price-dividend ratio

	$j=1$		$j=4$		$j=12$		$j=20$	
	β	R^2	β	R^2	β	R^2	β	R^2
Sample	-0.023^{**}	0.013	-0.024^{**}	0.056	-0.023^{*}	0.188	-0.022^{*}	0.297
	(0.020)	(0.016)	(0.019)	(0.055)	(0.014)	(0.113)	(0.011)	(0.140)
Gauss(s)	-0.043	0.006	-0.044	0.024	-0.039	0.057	-0.034	0.075
BEGE(s1)	-0.051	0.023	-0.048	0.087	-0.042	0.215	-0.037	0.297
BEGE(s2)	-0.049	0.020	-0.047	0.075	-0.041	0.190	-0.036	0.264
BEGE(d)	-0.058	0.028	-0.054	0.102	-0.047	0.255	-0.040	0.352
BEGE(d)'	-0.030	0.010	-0.029	0.037	-0.026	0.099	-0.024	0.147

Panel B: Conditional distribution of excess returns w.r.t. $pvol_t$

	Sample	Model-Implied				
	Stats	Gauss(s)	BEGE(s1)	BEGE(s2)	BEGE(d)	BEGE(d)'
Δq^{5}	-0.324^{**}	-0.007	-0.154	-0.121	-0.374	-0.188
	(0.129)					
Δq^{50}	0.085	0.032	0.103	0.086	0.127	0.051
	(0.039)					
Δq^{95}	0.161^{*}	0.088	0.265	0.247	0.513	0.290
	(0.078)					

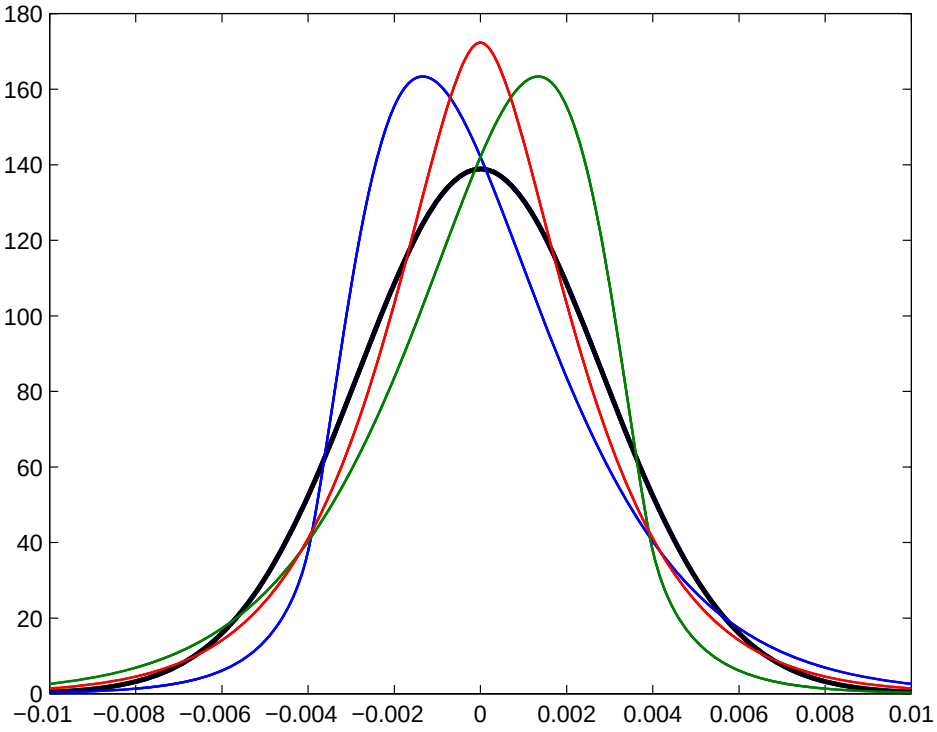
In Panel A, the upper rows report sample regression statistics with bootstrapped standard errors for the regression model

$$\frac{1}{j}\sum_{i=1}^{j} xret_{t+j} = a + \beta \cdot pd_t + u_{t+j}$$

where j is measured in quarters. The rows below report simulated values for the same regression statistics under the various models.

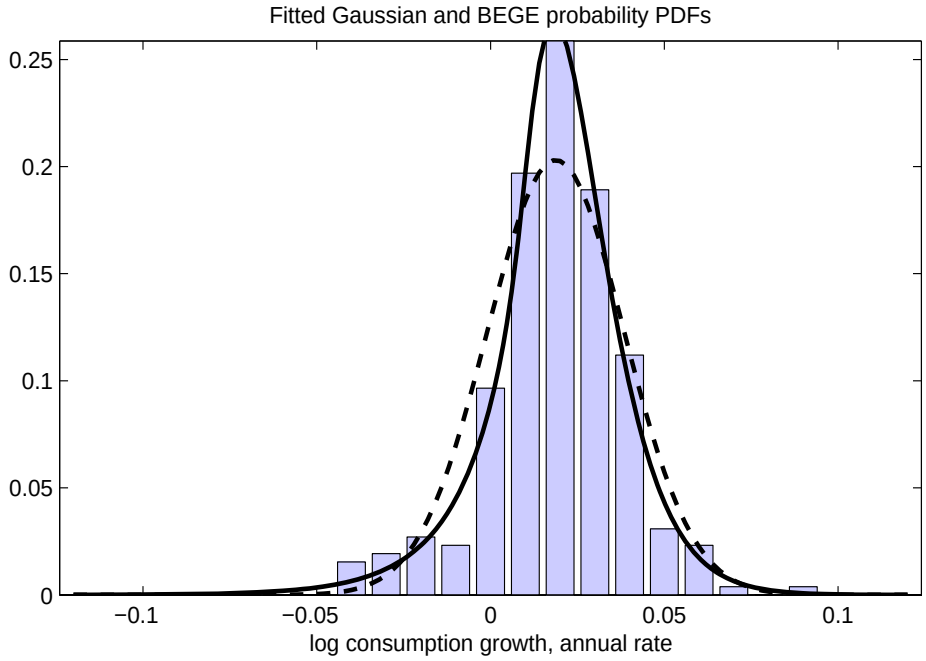
In panel B, quantile moment shifts are reported. These are defined as the shift in the conditional quantiles for one-period ahead excess equity returns when lagged physical conditional volatility of equity returns moves above its unconditional median value. One, two, and three asterisks denote statistical significance at the 10, 5, and 1 percent level.

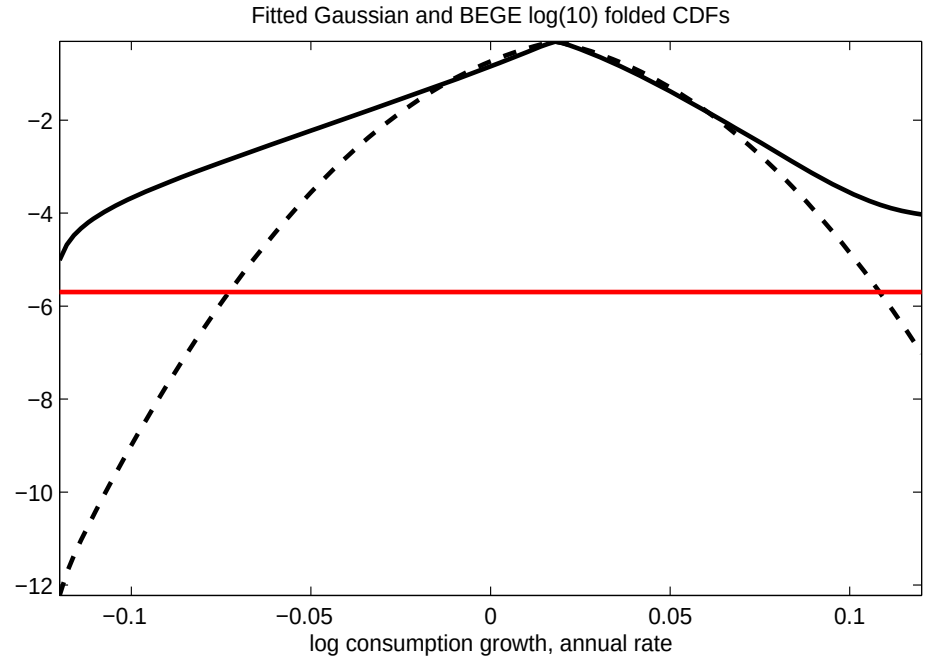
Figure 1:Examples of the BEGE Distribution



This figure plots BEGE densities under various configurations for p_t, n_t, σ_{cp} and σ_{cn}. All the distributions have zero mean and standard deviation 0.0029. The parameter configurations for the lines are as follows.

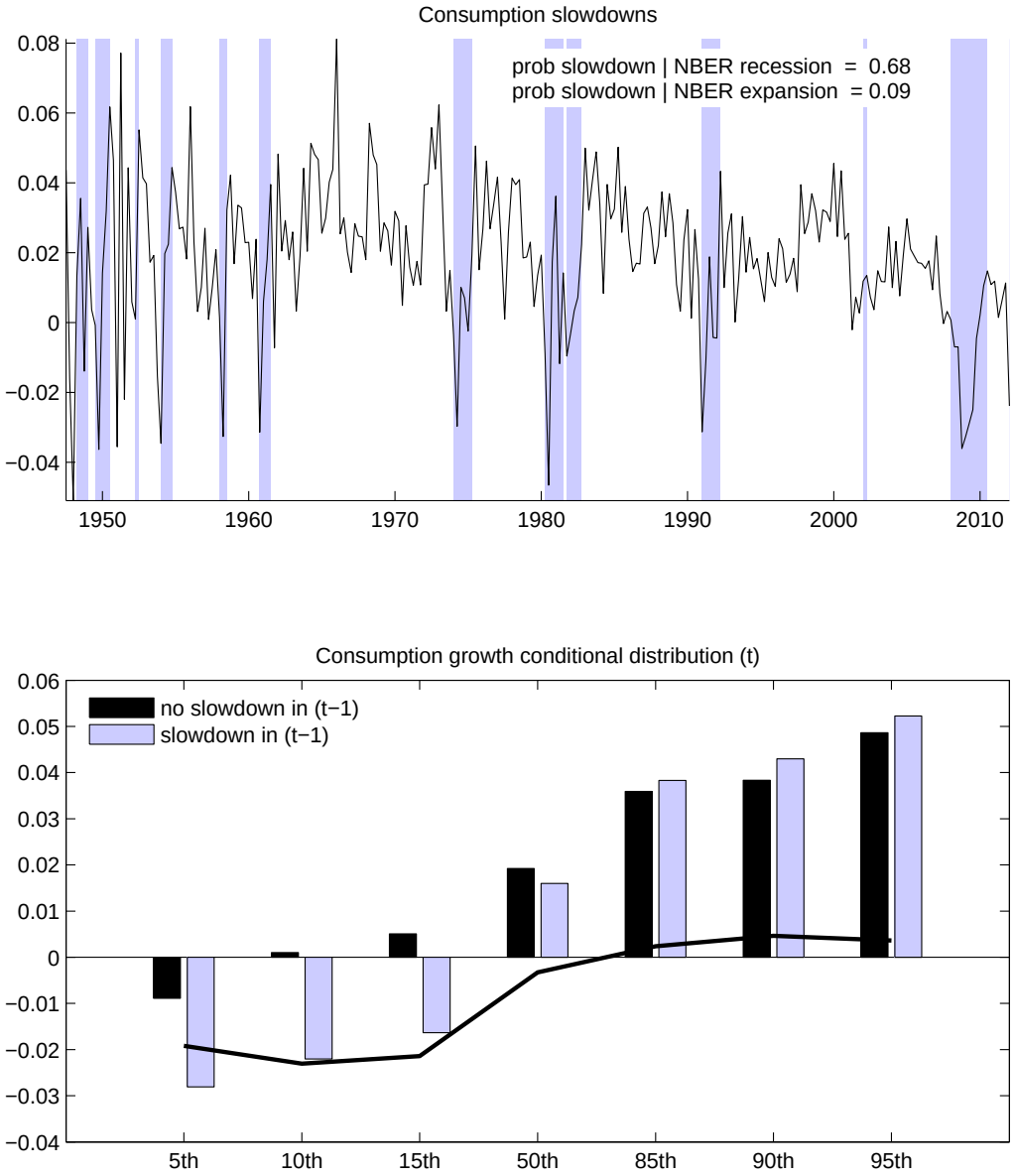
	p_t	n_t	σ_{cp}	σ_{cn}
black	40	40	0.0003	0.0003
red	2	2	0.0014	0.0014
green	.4	3	0.0016	0.0016
blue	3	.4	0.0016	0.0016





Notes: The top panel shows the empirical histogram of real, quarterly consumption growth in for U.S. nondurables and services from 1947Q2-2011Q4. The data have be pre-whitened to remove aggregation effects using an AR(1) model. The panel also shows fitted i.i.d. Gaussian (dashed) and BEGE (solid) probability densities functions for distributions that are fitted to the sample consumption data. The parameters for these distributions are reported in Table 1. The bottom panel shows the base10 log of the fitted folded Gaussian (dashed) and BEGE (solid) cumulative distribution functions. (the folded cdf of a variable, x, is defined as max(cdf(x),1-cdf(x))). For reference, the horizontal red line indicates the log_{10} of $2e^{-6}$, the estimated probability of a catastrophic asteroid Earth strike in any given year.

Figure 3: Conditional distribution of consumption growth



Notes: The top panel plots the time series for real, quarterly consumption growth in for U.S. nondurables and services from 1947Q2-2011Q4. The data have been pre-whitened to remove aggregation effects using an AR(1) model. The shaded regions represent "consumption slowdowns," which are defined as periods in which the four-quarter moving average of consumption growth is less than 1 percent (annual rate). The inset text in the figure describes the overlap between consumption slowdowns and NBER-defined recessions. The bottom panel shows the distribution of consumption growth in period (t) conditional on period $(t-1)$ having been a slowdown period or not. The solid line plots the differences between the quantiles of the two distributions.