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

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

Geert Bekaert and Eric Engstrom

2015-053

Please cite this paper as:
Bekaert, Geert and Eric Engstrom (2015). "Asset Return Dynamics under Habits and Bad-Environment Good-Environment Fundamentals," Finance and Economics Discussion Series 2015-053. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.053.

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

Geert Bekaert

Columbia University and NBER

Eric Engstrom

Federal Reserve Board of Governors†

July 3, 2015

*We appreciate comments from Andrey Ermolov, Nancy R. Xu, and seminar participants at the Massachusets Institute of Technology, the Wharton School, Northwestern University, New York University, and over 20 other 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 consumption growth and fits standard salient features of asset prices including the means and volatilities of equity returns and a low risk free rate. BEGE dynamics additionally allow the model to generate realistic properties of equity index options prices, and their comovements with the macroeconomic outlook. In particular, when option-implied volatility is high, as measured for instance by the VIX index, the distribution of consumption growth is more negatively skewed.

I Introduction

A number of consumption-based asset pricing models have emerged that can claim empirical success in matching salient features of macroeconomic and asset return data. For example, Campbell and Cochrane (1999) develop an external habit framework in which a representative agent's counter-cyclical risk aversion is the driver of asset return dynamics. Campbell and Cochrane keep the exogenous process for consumption growth deliberately simple, yet the model fits well the means of the equity risk premium and the risk free short-term interest rate, the variability of equity returns and price-dividend ratios, and long horizon return predictability. These feats are accomplished by assuming that habit moves slowly in response to consumption shocks, which are assumed to follow a time-invariant Gaussian distribution.

However, as shown in Figure 1, data for post-WWII U.S. consumption growth contradict the assumption of a time-invariant Gaussian distribution The top panel plots estimated values for the conditional variance of consumption growth from a standard GARCH-type model. Sharp peaks are witnessed periodically, usually around the time of NBER recessions. Moreover, estimates of the conditional skewness of consumption growth from a reduced-form model, shown in the bottom panel, also vary significantly over time, and the conditional skewness, especially unscaled skewness, appears to be negatively correlated with the conditional variance. For a risk-averse investor, the prospect that the skewness of consumption may become more negative at the same time that its volatility increases represents a risk that may have substantial implications for asset prices. (We will return to the details and statistical significance of these features in Section IV).

Moreover, the Great Recession of 2008 and 2009 witnessed stress in financial markets that was particularly evident in option markets. For example, the CBOE's VIX index, a measure of option-implied volatility, showed extreme variation and rose to unprecedented heights. The VIX index measures the option-implied expected stock market variance for the U.S. S&P 500 stock price index, and it is computed from a panel of options prices

across different strikes. Well-known as a "fear index" for asset markets (Whaley, 2000), the VIX is a positive function of uncertainty regarding stock market returns (the "physical" expected volatility) and a"variance premium," which is also the expected premium from selling a claim with a payoff equal to the realized future stock market variance (see Bollerslev, Gibson and Zhou (2011), Carr and Wu (2008)). Because during periods of stress the realized stock market variance generally increases, the variance premium tends to be positive. As depicted in Figure 2, movements in the VIX are empirically related to the conditional distribution of consumption growth. Displayed are shifts in the distribution of one-quarter ahead consumption growth depending on prior realizations of the VIX. Specifically, the "high VIX" ("low VIX") symbols indicate the consumption growth quantiles when the lagged value of VIX is in the highest (lowest) decile of its distribution over the sample period. When the VIX is relatively high, the consumption growth distribution becomes more negatively skewed, and in the left tail the quantile shifts are statistically significant. These empirical facts pose a challenge for the standard formulation of the Campbell and Cochrane model. While the standard model generates time-variation in risk aversion that may potentially propel the VIX to high values following bad shocks, the informativeness of the level of the VIX for the macroeconomic outlook is precluded by the assumption of a time-invariant distribution for consumption growth.

In this article, we propose a simple extension to the model of Campbell and Cochrane by assuming a stochastic process for consumption growth that has a time-varying and non-Gaussian distribution, following what we call a "Bad Environment-Good Environment" process, or "BEGE" for short. Each period, consumption growth 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 two types of shocks varies through time, there are "good times" where the positively-skewed distribution dominates, and "bad times" where the negatively-skewed distribution dominates and recession risk is high.

We show that appending the Campbell and Cochrane model with BEGE dynamics for consumption and dividend growth allows it to match the evidence in Figures 1 and 2 as well as several other features of macro economic data and data for asset and option prices. Regarding the latter, one can derive the option-implied higher order moments for stock returns from option prices across a range of strike prices. A prominent feature of these data is that the conditional moments of returns show time-varying (but generally negative) skewness and time-varying, but reliably fatter-than-Gaussian, kurtosis (see Figlewski, 2008). Our extension allows the model of Campbell and Cochrane to match these empirical features relatively well. The model also endogenously generates realistic asymmetric volatility, the tendency for equity returns to be negatively correlated with return variance.

As we discuss in Section V, our model combines stochastic risk aversion with a time-varying non-Gaussian distribution for consumption growth to fit salient features of asset returns and options data. Alternative mechanisms have been proposed in the literature. Within a long-run risk model, attempting to fit features of the VIX and the variance premium, Drechsler and Yaron (2011) introduce jumps in the conditional mean of consumption growth whereas Bollerslev, Tauchen and Zhou (2011) let volatility of the stochastic volatility process for consumption growth follows a separate stochastic process. Because these models are both conditionally Gaussian, they cannot generate the non-Gaussian consumption dynamics documented in Figures 1 and 2.[1] Another recent strand of the literature that also focuses on the technology rather than preferences has rekindled the idea in Rietz (1988) 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 (2013). Because the BEGE model delivers a flexible conditional distribution of consumption growth, it can potentially accommodate disaster risk. However, our estimation using U.S. data delivers a process that is consistent with persistent recession behavior rather than extreme

[1] However, it is possible for these models to generate non-Gaussian dynamics for consumption growth over longer horizons than the frequency at which the models are specified.

but transitory disaster shocks as in most disaster risk articles.

The remainder of the article is organized as follows. Section II introduces the model and Section III discusses the implied asset price solutions. Section IV outlines the calibration and estimation of the model parameters. As part of this exercise, we confront the BEGE model with actual consumption data, showing that it matches time-varying non-Gaussianities in the consumption growth process. Section V analyzes the fit of the model with respect to asset prices. The final section offers some concluding remarks.

II The Bad Environment-Good Environment (BEGE) Habit Model

In this section, we introduce the representative agent model. We begin with a discussion of the assumed data generating process for consumption and dividend growth, and then describe our assumptions regarding the preferences of the agent.

II.A The BEGE distribution for consumption growth

Our model for consumption growth is specified at the monthly frequency and is described 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 of nondurables and services, and $\overline{g}$ is the mean rate of consumption growth, which is assumed to be constant. The parameters σ_{cp} and σ_{cn} are both positive. The key point of departure for our model from the formulation of Campbell and Cochrane is that the shocks, $\omega_{p,t+1}$ and $\omega_{n,t+1}$, do not follow invariant Gaussian distributions, but instead are assumed to be drawn from distributions that are non-Gaussian with potentially time-varying moments. Specifically, we assume that the two shocks are independent and governed by centered gamma

distributions. We use the notation,

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

to denote that $\omega_{p,t+1}$ follows a centered gamma distribution with shape parameter p_t (using a subscript to denote that p_t may vary over time) and a unit scale parameter. The probability density function for $\omega_{p,t+1}$, denoted $\phi(w_{p,t+1})$, is given by the formula:

$$\phi(\omega_{p,t+1}) = \frac{1}{\Gamma(p_t)} (\omega_{p,t+1} + p_t)^{p_t - 1} \exp(-\omega_{p,t+1} - p_t) \tag{3}$$

for $\omega_{p,t+1} > -p_t$ and with $\Gamma(\cdot)$ representing the gamma function. Unlike the standard gamma distribution, the centered gamma distribution has a mean of zero. Similar results hold for the distribution of $\omega_{n,t+1}$, with n_t governing the shape of the distribution. The conditional moments of the two shocks are easily derived, for instance, by using the moment generating function (see the Appendix A) to obtain:

$$\begin{array}{lcc} \hline\hline & -1 \cdot \omega_{n,t+1} & +1 \cdot \omega_{p,t+1} \\ \hline var_t & n_t & p_t \\ skew_t & -2/\sqrt{n_t} & +2/\sqrt{p_t} \\ xkurt_t & 6/n_t & 6/p_t \\ \hline\hline \end{array} \tag{4}$$

Consider again the $\omega_{p,t+1}$ shock. The conditional variance of the distribution is equal to the shape parameter, p_t, but both conditional skewness ($skew_t$) and excess kurtosis ($xkurt_t$), while positive, are declining in p_t. Moreover, the distribution approaches the Gaussian distribution as p_t tends towards infinity. This is visibly apparent in the top right panel of Figure 3, in which the distribution of $\omega_{p,t+1}$ is plotted for various values of p_t. In each case, the distribution starts at $-p_t$ and then has a long right hand side tail. As p_t increases, the distribution appears more and more Gaussian. As can be seen in the upper left panel of Figure 3, the distribution for $-1 \cdot \omega_{n,t+1}$ is the mirror image of that for $\omega_{p,t+1}$.

To flesh out the properties of consumption growth that are implied by these two com-

ponent shocks, we begin by noting that the shape parameters, p_t and n_t, together govern the higher-order moments for consumption growth:

$$\begin{aligned} var_t\left(\Delta c_{t+1}\right) &= \sigma_{cp}^2 p_t + \sigma_{cn}^2 n_t \\ skew_t\left(\Delta c_{t+1}\right) &= 2\left(\sigma_{cp}^3 p_t - \sigma_{cn}^3 n_t\right) / var_t\left(\Delta c_{t+1}\right)^{3/2} \\ xkurt_t\left(\Delta c_{t+1}\right) &= 6\left(\sigma_{cp}^4 p_t + \sigma_{cn}^4 n_t\right) / var_t\left(\Delta c_{t+1}\right)^{4/2} \end{aligned} \tag{5}$$

The table shows that both p_t and n_t contribute positively to the conditional variance of consumption. They differ, however, in their implications for the conditional skewness of consumption. The numerator in the expression for skewness is positive when p_t is relatively large, and negative when n_t is large. This is the essence of the BEGE model: the bad environment refers to an environment in which the $\omega_{n,t+1}$ shocks dominate; in the good environment the $\omega_{p,t+1}$ shocks dominate. Of course, the shocks are always zero in expectation, 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_t and n_t given values for σ_{cp} and σ_{cn}. For the conditional excess kurtosis of consumption growth, both p_t and n_t contribute positively to both the numerator and denominator of the expression, so that the dependence of excess kurtosis on p_t and n_t depends on their respective levels as well as on σ_{cp} and σ_{cn}. However, both skewness and excess kurtosis for Δc_{t+1} eventually tend towards zero for larger values of p_t and n_t.

The bottom panel of Figure 3 plots four examples of BEGE densities under various combinations for p_t, n_t, 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_t and n_t. This distribution very closely approximates the Gaussian distribution. The blue line plots a BEGE density with smaller, but still equal values for p_t and n_t. This density is more peaked and has fatter tails than the Gaussian distribution. The red line plots a BEGE density with large p_t but small n_t and is duly right-skewed. Finally, the green line plots a density with large n_t and small p_t, and is left-skewed. This demonstrates the flexibility of the BEGE distribution and illustrates the

role of p_t as the "good environment" variable and n_t as the "bad environment" variable. In essence, the BEGE model is a stochastic volatility model with "good" and "bad" volatilities in which changes in volatility also change the shape of the distribution.

An astute reader may note that the mode of the "good environment" ("bad environment") distribution is shifted to the left (right), and that the total probability of negative shocks is actually higher in the good environment than in the bad environment distribution. Yet, the monikers "good and bad environment" are economically well-founded insofar as a typical risk averse agent should prefer facing an $\omega_{p,t}$ shock (which has a bounded left tail and unbounded upside) to an $\omega_{n,t}$ shock of equal variance (which has an unbounded left tail and a bounded upside).

The following example illustrates that a risk averse investor in a simple static setting would dislike exposure to bad-environment shocks more than she would Gaussian shocks with the same variance, and that she would prefer the good-environment shocks to the Gaussian case. For this example, we assume the agent's utility function has constant relative risk aversion over one-period ahead wealth:

$$U\left(\widetilde{W}\right) = \frac{\widetilde{W}^{1-\gamma} - 1}{1-\gamma} \tag{6}$$

where $\widetilde{W}$ is the stochastic value of one-period ahead wealth. We consider three different distributions for $\ln\left(\widetilde{W}\right)$. All three distributions have a common variance, σ_w^2, but they differ in their means and distributions of shocks:

distribution	mean	variance	skewness
Gaussian	μ_w	σ_w^2	0
$-1 \cdot \omega_n$	$\mu_w + \Delta_n$	σ_w^2	$-2/\sqrt{n}$
$+1 \cdot \omega_p$	$\mu_w + \Delta_p$	σ_w^2	$+2/\sqrt{p}$

(7)

where we have dropped the time subscripts for this static example. We find values for Δ_n and Δ_p such that the expected utility of all three distributions is the same, that is, equal to the expected utility under the log-Gaussian distribution for wealth. An exact formula

for these "certainty equivalent" losses or gains is derived in the appendix, but a third order Taylor series approximation yields the following intuitive expressions:

$$\begin{aligned} \Delta_n &\approx +\frac{1}{3n^{1/2}}\sigma_w^3\left(1-\gamma\right)^2 \\ \Delta_p &\approx -\frac{1}{3p^{1/2}}\sigma_w^3\left(1-\gamma\right)^2 \end{aligned} \tag{8}$$

The expression for Δ_n is unambiguously positive, indicating that the agent requires a positive boost to the mean of the $-1\cdot\omega_n$ distribution to be equally willing to bear the disutility of its long left tail and finite right tail. The certainty equivalent mean shift declines proportionately to $n^{1/2}$ because the distribution approaches a Gaussian distribution for larger values of n, and it is unsurprisingly increasing in both σ_w^2 and γ (assuming $\gamma > 1$). The certainty equivalent mean shift for the good-environment distribution, Δ_p, is unambiguously negative, so that the agent is willing to accept a lower level of expected gains in exchange for the long right tail, and it declines with $p^{1/2}$.

For illustrative purposes, Figure 4 plots the exact functions for Δ_n and Δ_p for $\gamma = 5$ and $\sigma_w^2 = 0.15^2$,[2] consistent with features of a US stock market gamble. At low values of n, the risk averse agent requires an 8 percent risk premium to be as happy with the BE shock as she is with the Gaussian shock. The equivalent reduction in premium she is willing to accept to face a GE shock is only 2 percent for low values of p. The asymmetry is due to the higher order terms not present in the Taylor approximation in Equation (8). (For the BE-shock all the higher order terms are positive, but for the GE shock the signs of the terms alternate, mitigating the total effect.)

II.B The BEGE distribution for dividend growth

For monthly dividend growth, Δd_t, we follow Campbell and Cochrane's specification in which dividend growth follows a process symmetric to that for consumption growth:

$$\Delta d_{t+1} = \overline{g} + \sigma_{dp}\omega_{p,t+1} - \sigma_{dn}\omega_{n,t+1} \tag{9}$$

[2]The mean μ_w is irrelevant to these calculations and can be set to 0 without loss of generality.

where Δd_t is real logarithmic dividend growth per capita. Note that the same shocks that affect consumption growth also drive dividend growth, as is true in the model of Campbell and Cochrane. As in Longstaff and Piazzesi (2004), the dividend process may exhibit stronger non-Gaussianity than consumption growth. Moreover, consumption and dividends follow two different unit root processes. An alternative assumption would be that consumption and dividends are cointegrated.

II.C Time variation in the BEGE distribution

For parsimony p_t is assumed to be constant at a level denoted by $\overline{p}$, but n_t varies over time, according to a simple, autoregressive process:

$$n_{t+1} = (1 - \rho_n)\,\overline{n} + \rho_n n_t + \sigma_{nn}\omega_{n,t+1} \tag{10}$$

where $\overline{n}$ is the unconditional mean of the process, ρ_n is the autocorrelation coefficient, and $\sigma_{nn} > 0$ governs the conditional volatility of the process. Therefore n_t is a state variable that the representative agent observes for the purposes of calculating asset prices. Note that the shock to n_{t+1}, $\omega_{n,t+1}$, is the same as the bad-environment shock for consumption, but it enters with a different sign; positive $\omega_{n,t+1}$, shocks raise n_{t+1} and simultaneously lower Δc_{t+1}. The conditional volatility of the shock to n_{t+1} is $\sigma_{nn}\sqrt{n_t}$ so that our specification is similar to other models in which the conditional volatility of the variable is proportional to the square root of its level.[3]

[3]The minimum possible value of $\omega_{n,t+1}$ is $-n_t$, an artifact of the one-sided nature of the gamma distribution. This implies that under the additional technical assumption that $\sigma_{nn} < \rho_n$, the n_t process can never become negative, even in a discrete time setting. This is a desirable property for a volatility process, which is not shared by standard Gaussian stochastic volatility processes.

II.D Preferences

We adhere to the preference structure of the representative agent in Campbell and Cochrane (1999) in which the agent maximizes the utility function,

$$E_t \sum_{j=0}^{\infty} \delta^j \frac{(C_t - H_t)^{1-\gamma} - 1}{1-\gamma}. \tag{11}$$

where C_t is aggregate consumption and H_t is the habit stock with $C_t > H_t$. The parameter δ is the subjective time discount factor. We also define the consumption surplus ratio S_t as,

$$S_t = \frac{C_t - H_t}{C_t}. \tag{12}$$

In Campbell and Cochrane, H_t is an external subsistence or habit level. Hence, the local curvature of the utility function, denoted Q_t, is given by

$$Q_t \equiv -C_t \frac{\partial^2 U_t}{\partial C_t^2} / \frac{\partial U_t}{\partial C_t} = \gamma \frac{C_t}{C_t - H_t} = \frac{\gamma}{S_t} \tag{13}$$

As the surplus ratio goes to zero, local curvature, and thus the risk aversion of the representative agent, both go to infinity.

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:

$$m_{t+1} = \ln \delta - \gamma \Delta c_{t+1} - \gamma \left(s_{t+1} - s_t \right) \tag{14}$$

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

The process for s_t is specified as:

$$s_{t+1} = (1-\phi)\,\overline{s} + \phi s_t + \lambda_t \left(\Delta c_{t+1} - \overline{g} \right) \tag{15}$$

where $\overline{s}$, and ϕ are scalar parameters. The variable λ_t is the so-called sensitivity function. The function is positive so that negative shocks to consumption growth decrease the surplus ratio and increase risk aversion, which generates countercyclical risk premiums. Campbell and Cochrane choose the functional form of the sensitivity function to control the volatility of interest rates. We proceed similarly and defer a full characterization of the sensitivity function in our framework until our derivation of the short term interest rate in the next

section.

III Asset Prices

In this section, we present various features of asset prices implied by the model. We begin by defining two useful parameters which capture the sensitivity of the pricing kernel to the two shocks in the model. (A table in the Appendix catalogs the main variables and parameters to which we refer in the text.)

$$\begin{aligned} a_{p,t} &\equiv \frac{\partial m_{t+1}}{\partial \omega_{p,t+1}} = -\gamma\left(1+\lambda_t\right)\sigma_{cp} \\ a_{n,t} &\equiv \frac{\partial m_{t+1}}{\partial \omega_{n,t+1}} = \gamma\left(1+\lambda_t\right)\sigma_{cn} \end{aligned} \tag{16}$$

Because positive $\omega_{p,t+1}$ shocks raise consumption growth and surplus consumption, their contribution to marginal utility is negative so $a_{p,t} < 0$. Conversely, because $\omega_{n,t+1}$ shocks lower both consumption and the surplus ratio, they raise marginal utility so that $a_{n,t} > 0$.

III.A The risk free rate and the sensitivity function

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

$$\exp\left(rf_t\right) = E_t\left[\exp\left(m_{t+1}\right)\right]^{-1} \tag{17}$$

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

$$\begin{aligned} rf_t = & -\ln\delta + \gamma\overline{g} + \gamma\left(1-\phi\right)\left(\overline{s}-s_t\right) \\ & +\overline{p}\ f\left(a_{p,t}\right) + n_t\ f\left(a_{n,t}\right) \end{aligned} \tag{18}$$

where we use the notation-simplifying function, $f\left(x\right) = x + \ln\left(1-x\right)$. In evaluating the conditional expectation in (17), we use the moment generating function of demeaned gamma distributions, which is a power function (See Appendix A for details). The solution is isomorphic to the well-known result for log-normally distributed variables. The first line in (18) simply represents the negative of the conditional mean of m_{t+1} and the second line

is a non-Gaussian equivalent of one-half times its conditional variance.

Economically, the first line in the solution for rf_t reflects 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 be higher in equilibrium to dissuade them from doing so (short term bonds having zero net supply). The terms on the second line capture precautionary savings effects: Higher uncertainty exerts downward pressure on risk free rates. 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 the precautionary savings terms:

$$rf_t \approx \begin{pmatrix} -\ln\delta + \gamma\overline{g} + \gamma(1-\phi)(\overline{s} - s_t) \\ + \left(-\frac{1}{2}a_{p,t}^2 - \frac{1}{3}a_{p,t}^3\right)\overline{p} \\ + \left(-\frac{1}{2}a_{n,t}^2 - \frac{1}{3}a_{n,t}^3\right)n_t \end{pmatrix} \tag{19}$$

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 terms of order higher than two are zero. The cubic terms above 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. Economically, 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 savings 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 attributable to volatility. Through this mechanism our model may generate the kind of extremely low risk free rates witnessed in the 2007-2009 crisis period.

In Campbell and Cochrane, λ_t is modeled as a shifted square root function:

$$\lambda_t = \left\{ \begin{array}{ll} \frac{1}{\overline{S}}\sqrt{1-2\left(s_t-\overline{s}\right)}-1 & s_t < s_{\max} \\ 0 & s_t \geq s_{\max} \end{array} \right\} \tag{20}$$

where $s_{\max}$ is simply defined as the value of s at which λ_t runs into zero. With this specification, negative consumption shocks increase λ_t. Campbell and Cochrane then parameterize the function to impose conditions consistent with an interpretation of habit persistence (see below) and to control the volatility of the short rate. The parameter $\overline{S}$ is chosen as

$$\overline{S} = \sqrt{\sigma_c^2 \frac{\gamma}{1-\phi-b/\gamma}} \tag{21}$$

where $\overline{s} = \ln\left(\overline{S}\right)$ and σ_c^2 is the variance of the consumption shock. The specifications for λ_t and $\overline{S}$ produce a short rate that follows:

$$r f_t = -\ln\delta + \gamma\overline{g} - \frac{\gamma\left(1-\phi\right)-b}{2} + b\left(\overline{s}-s_t\right). \tag{22}$$

In their main specification, Campbell and Cochrane further impose $b = 0$ so that the utility-smoothing and precautionary savings demands cancel out and the short rate is rendered constant. Wachter (2006) and Campbell and Cochrane (1995) demonstrate how having a nonzero value for b enables habit models to fit some salient features of the term structure of interest rates.

We proceed with an analogous specification of the sensitivity function. We first define an analog to $\overline{S}$, which we denote by $\overline{S}_t$ to signify that it is a time-varying function of n_t.

$$\overline{S}_t = \sqrt{\left(\sigma_{cp}^2\overline{p} + \sigma_{cn}^2 n_t\right)\frac{\gamma}{1-\phi-b/\gamma}} \tag{23}$$

Note that this specification merely inserts the time-varying variance of consumption growth under the BEGE model into Campbell and Cochrane's formulation of $\overline{S}$ in Equation (21). Our sensitivity function is then defined as,

$$\lambda_t = \left\{ \begin{array}{ll} \frac{1}{\overline{S}_t}\sqrt{1-2\left(s_t-\overline{s}\right)}-1 & s_t < s_{t,\max} \\ 0 & s_t \geq s_{t,\max} \end{array} \right\} \tag{24}$$

where $\overline{s}$ is the unconditional mean of s_t (see below) and noting, with some abuse of notation,

that the maximum level of s_t that keeps λ_t positive, $s_{t,\max}$, is time-varying under our model. That is, $s_{t,\max} = \overline{s} + \frac{1}{2}\left(1 - \overline{S_t^2}\right)$

In addition, $\overline{s}$, the unconditional mean of s_t, is set to $\ln\left(\overline{S}\right)$ where $\overline{S}$ is $\overline{S}_t$ when evaluated at $n_t = \overline{n}$. As in Campbell and Cochrane, this restriction ensures that the parameterization obeys two conditions they set out as necessary for a habit interpretation of the preference specification.[4] We note that because λ_t depends on n_t through the specification of $\overline{S}_t$, s_t cannot be written purely as a function of past consumption shocks but also depends on the history of n_t. From that perspective, the model is not a pure habit model.

Using again the third order approximation in Equation (19) to gain intuition (we use the exact solution in our empirical work), the risk free rate simplifies to

$$rf_t = -\ln\delta + \gamma\overline{g} - \frac{\gamma(1-\phi) - b}{2} + b\left(\overline{s} - s_t\right) - \frac{1}{3}a_{p,t}^3\overline{p} - \frac{1}{3}a_{n,t}^3 n_t. \tag{25}$$

Except for the final two terms, this expression is identical to that of Campbell and Cochrane in which the short rate is a linear function of s_t and if b is positive, a lower surplus ratio (higher risk aversion) increases the short rate. In the model of Wachter (2006) , this mechanism ensures that the intertemporal smoothing effect of s_t on interest rates dominates the precautionary savings effect. This, in turn, generates an upward sloping real yield curve and positive bond risk premiums.

The final two terms are novel features of our model The term, $-\frac{1}{3}a_{p,t}^3\overline{p}$, is positive, and increasing in λ_t. In economic terms, the presence of the positively skewed shock raises the risk free rate, more so when sensitivity is high (large λ_t). The term, $-\frac{1}{3}a_{n,t}^3 n_t$, is unambiguously negative and its magnitude increases with both λ_t and n_t. Intuitively, the risk free rate falls when the variance of the negatively skewed shock is high (large n_t) and

[4] These conditions are as follows. At $s_t = \overline{s}$ and $n_t = \overline{n}$

$$\frac{dh_t}{dc_t} = 0$$

and

$$\frac{d}{ds_t}\left[\frac{dh_t}{dc_t}\right] = 0$$

where $h_t = \ln(H_t)$ is the logarithm of the habit stock.

when the sensitivity of risk aversion to shocks is high (large λ_t). In our model n_t and λ_t are correlated, with positive $\omega_{n,t}$ shocks increasing both variables. (Positive $\omega_{p,t}$ shocks decrease λ_t and do not affect n_t.)

III.B Equity prices

To solve for the price of the equity claim, we rely on the no-arbitrage condition that any asset return, R_{t+1}, must satisfy,

$$1 = E_t \left[M_{t+1} R_{t+1} \right] \tag{26}$$

For equity returns, which we model as a claim to the real dividend stream, let P_t denote the ex-dividend price of the claim. In this case, Equation (26) can be written as

$$P_t = E_t \left[M_{t+1} \left(P_{t+1} + D_{t+1} \right) \right] \tag{27}$$

One can solve forward for P_t and after applying a transversality condition, arrive at the expression

$$PD_t = E_t \sum_{i=1}^{\infty} \exp \left[\sum_{j=1}^{i} \left(m_{t+j} + \Delta d_{t+j} \right) \right] \tag{28}$$

where PD_t is the price-to-dividend ratio. Let $F_i \left(s_t, n_t \right)$ denote i^{th} term in the inner summation above,

$$F_i \left(s_t, n_t \right) = E_t \exp \left[\sum_{j=1}^{i} m_{t+j} + \Delta d_{t+j} \right] \tag{29}$$

Then substitution yields the recursion,

$$F_i \left(s_t, n_t \right) = E_t \left[\exp \left(m_{t+1} + \Delta d_{t+1} \right) F_{i-1} \left(s_{t+1}, n_{t+1} \right) \right] \tag{30}$$

Iterating on Equation (30) produces $F_i \left(s_t, n_t \right)$ for all i and we then sum terms in Equation (28) (see the "series" method in Wachter (2005)). We are not aware of a closed-form solution for the price-dividend ratio in either the case of Gaussian- or BEGE-distributed innovations, so we solve the system numerically for a grid of values for s_t and n_t. We chose a grid of 100 equally spaced points for both s_t and n_t (10,000 points in total), with the grid edges chosen as the minimum and maximum values for s_t and n_t from a single simulation of the system for

100,000 periods. Adding points to the grid or expanding the grid area does not materially alter our solution. At each step of the recursion, we must evaluate $F_{i-1}(s_{t+1}, n_{t+1})$ at points (s_{t+1}, n_{t+1}) that are, in general, not on the grid. These points are then approximated by interpolating between grid points.

Solving for expectations such as in Equation (30) requires numerical integration. We use standard numerical quadrature techniques to accomplish the integration. In particular, we replace the continuous conditional distributions for $\omega_{n,t+1}$ and $\omega_{p,t+1}$ with discrete distributions having 10 points each with the discrete points and their associated probabilities chosen to exactly match the first 10 moments of their continuous counterparts (see, for instance, Miller and Rice, 1983). The moments of the distribution for $\omega_{n,t+1}$ depend on the values for n_t on the grid.

We define log equity returns, ret_t, in the usual manner

$$ret_{t+1} = \Delta d_{t+1} + \ln(1 + PD_{t+1}) - \ln(PD_t) \tag{31}$$

Log excess returns are defined as

$$xret_{t+1} = ret_{t+1} - rf_t \tag{32}$$

and the equity risk premium is defined as the one-period ahead conditional expectation of log excess returns, $E_t[xret_{t+1}]$.

The model must not only generate a sizable equity premium, but also a sizable Sharpe ratio for equity. In the Campbell and Cochrane 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:

$$\max_{\{\text{all assets}\}} \frac{E_t\left[XRET^e_{t+1}\right]}{\sigma_t\left[XRET^e_{t+1}\right]} = \frac{\sigma_t[M_{t+1}]}{E_t[M_{t+1}]} \approx \sqrt{a^2_{c,t}} \tag{33}$$

where $XRET^e_{t+1}$ is the excess return of a generic asset, and $a_{c,t} = -\gamma\sigma_c(1 + \lambda_t)$ is the derivative of the pricing kernel with respect to the consumption shock under the Campbell and Cochrane model. The Sharpe ratio is also easy to compute in our framework, given the presence of two independent gamma processes, and we derive an approximation to the

Sharpe ratio under the BEGE-habit model in the Appendix:

$$\begin{aligned}\frac{\sigma_t\left[M_{t+1}\right]}{E_t\left[M_{t+1}\right]} &= \sqrt{\exp\left\{2\overline{p}\ln\left[\frac{1-a_{p,t}}{\left(1-2a_{p,t}\right)^{1/2}}\right]+2n_t\ln\left[\frac{1-a_{n,t}}{\left(1-2a_{n,t}\right)^{1/2}}\right]\right\}-1} \\ &\approx \sqrt{\overline{p}\left[a_{p,t}^2+2a_{p,t}^3\right]+n_t\left[a_{n,t}^2+2a_{n,t}^3\right]} \qquad (34)\end{aligned}$$

Absent the third order terms, this expression looks very much like Equation (33) except with two sources of uncertainty. Intuitively, the third order effect of n_t is to raise the Sharpe ratio (recall that $a_{n,t}>0$). This further highlights the role of n_t as governing "bad" volatility.

III.C The variance premium

We follow the common approach of defining the variance premium as the difference between the conditional variance of logged equity returns under the "risk neutral" and physical measures. Specifically,

$$vprem_t = qvar_t - pvar_t \qquad (35)$$

where

$$\begin{aligned}pvar_t &= E_t\left[ret_{t+1}^2\right]-E_t\left[ret_{t+1}\right]^2 \\ qvar_t &= E_t\left[M_{t+1}ret_{t+1}^2\right]/E_t\left[M_{t+1}\right]-E_t\left[M_{t+1}ret_{t+1}\right]^2/E_t\left[M_{t+1}\right]^2 \qquad (36)\end{aligned}$$

Intuitively, in a discrete state economy, the physical variance is computed using the actual state probabilities whereas the risk-neutral variance makes use of probabilities that are adjusted for the pricing of risk using the pricing kernel. It therefore assigns more weight to high marginal utility states.

The square of the CBOE's VIX index essentially provides a direct reading on the risk neutral variance. The VIX calculation measures 30-day expected option-implied volatility of the S&P 500 Index. It reflects a weighted average of the prices of near- and next-term put and call options over a range of strike prices, with maturities of more than 23 days and less than 37 days to expiration (see CBOE (2014) for more details). Importantly, this estimate

is model free and does not rely on an option pricing model. The computation of the actual VIX index relies on theoretical results showing that option prices can be used to replicate any bounded payoff pattern; in fact, they can be used to replicate Arrow-Debreu securities (Breeden and Litzenberger (1978), Bakshi and Madan (2000)). Britten-Jones and Neuberger (2000) and Bakshi, Kapadia and Madan (2003) show how to infer the "risk-neutral" expected variance for a stock index from option prices. We rely on their results to similarly infer risk neutral skewness and kurtosis from a panel of option prices as well.

While the risk neutral variance can be measured from option prices, the variance premium also requires an estimate of the physical expected variance. It is typical in the literature to do so by projecting future "realized" monthly variances (computed using high frequency returns) onto a set of current instruments (see Bekaert and Hoerova, 2014, for a discussion). The variance premium can also be viewed as the expected payoff to selling variance in a variance swap. A variance swap is a financial derivative in which the buyer of the swap pays an amount based upon the realized variance of the price changes of the underlying product. Conventionally, the realized variance is computed as the sum of squared daily log returns, based upon the closing price. The seller of the swap pays a fixed amount, quoted at the deal's inception, the variance swap rate. A return variance swap has zero net market value at inception. At maturity, the payoff to the short side of the swap is equal to the difference between the variance swap rate and the realized variance over the life of the contract. No-arbitrage conditions dictate that the variance swap rate equals the risk-neutral expected value of the realized variance, that is, $qvar_t$, or the squared value of the VIX. The expected premium on such a contract varies through time but is reliably positive. This is economically logical, as the buyer in the swap receives a positive pay-off when the variance spikes, which tends to happen in bad times. Moreover, during periods of stress, put option prices tend rise more than call option prices of equal "moneyness" (that is, having a strike price with the same distance from the contemporaneous futures price), an asymmetry that leads the VIX to be higher than the physical expected volatility.

In equilibrium, this asymmetry can be achieved through risk aversion and/or an asymmetric distribution for returns. 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 distribution of (log) returns distribution is Gaussian, then the volatility spread is zero. Analogously, the habit mechanism in Campbell and Cochrane (1999) will only succeed in generating a meaningful variance premium if it induces non-Gaussianities in endogenous returns. This occurs in the BEGE version of the Campbell – Cochrane model because returns "inherit" the non-Gaussianities from the fundamentals. We evaluate this conjecture quantitatively in Section V.

IV Parameter Calibration and Estimation

In this section, we estimate the model in Equations (1)-(10) using consumption data. Given the presence of an unobserved state variable (n_t) this is a non-trivial undertaking. Section IV.A describes the data, Section IV.B describes the estimation methodology, and Section IV.C reports the parameter estimates and their implications for consumption growth dynamics. We also discuss the calibration of the dividend growth process. Section IV.D describes the calibration of the preference parameters.

IV.A Data

We use U.S. post-WWII consumption growth data from the U.S. National Income and Product Accounts (published by the Bureau of Economic Analysis (BEA)). We first add together quarterly, nominal, seasonally adjusted consumption expenditures for nondurables and services over the period 1958Q1-2013Q4. According to the BEA, data before 1958

was subject to substantially higher measurement error, which could spuriously affect our estimates of the conditional volatility of consumption growth.[5] We deflate the resulting series using a weighted average of the deflators for nondurables and services, also available from the BEA, 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. The consumption based asset pricing literature often uses annual data going back to the Great Depression. The non-Gaussianities in these (annual) consumption data are much stronger than in our sample, but consumption growth is substantially more variable in the pre-WWII data. A dramatic downshift in volatility could lead the model estimates to indicate spurious non-Gaussian behavior. We therefore focus attention on post-1958 data to present a conservative view of potential non-Gaussianities in the consumption data.

Quarterly consumption growth data features a relatively high first-order autocorrelation coefficient (0.36 with a standard error of 0.09), which is inconsistent with the model in Equation (1). Such autocorrelation may well be due, in part, to temporal aggregation bias (see Working, 1960). We therefore formulate the model at the monthly frequency (as is also done in Campbell and Cochrane, 1999), assuming monthly consumption growth to be serially uncorrelated. This implies that quarterly consumption growth will have an autocorrelation coefficient of 0.21, which is within 2 standard errors of the moment in the data.

IV.B Estimation methodology

To estimate the model parameters, we use the simulated Classical Minimum Distance (CMD) technique described in Wooldridge (2002), which relies on the matching of sample statistics. A first set of statistics to be matched comprises the unconditional moments of consumption growth, which are reported in Table 1, Panel A, with bootstrapped standard

[5] See "Reliability and Accuracy of the Quarterly Estimates of GDP" in the October 1993 edition of the BEA publication, Survey of Consumer Business.

errors in parentheses.[6] Note that unconditional skewness ($skew$) is negative but not statistically significant, but excess kurtosis ($xkurt$), at 0.75, is highly statistically significant. When we aggregate to the annual frequency (Panel B), that situation is reversed: skewness is significantly negative, but excess kurtosis is not significant. However, we do not use the annual statistics in the estimation of the model.

Our next set of moments captures time variation in the conditional variance of consumption growth. To generate such moments, we estimate an auxiliary model with the quarterly data. The model is a well-known asymmetric volatility model in the GARCH class (see Glosten, Jagannathan and Runkle, 1993), where we assume quarterly consumption growth to follow an MA(1) process, consistent with the temporal aggregation of the serially uncorrelated consumption growth process at the monthly frequency. The model can be consistently estimated using maximum likelihood assuming Gaussian error terms, even though the error terms are not Gaussian under the BEGE model. We describe the GARCH model and the estimation results in the Appendix. The advantage of the GARCH model is that it delivers an estimate of the conditional variance of consumption growth at each point in time (which is graphed in Figure 1). The GARCH model delivers three additional statistics to match: the standard deviation of the conditional variance of consumption growth (the "vol of vol" in finance parlance) which Table 1, Panel C, shows to be statistically significantly different from zero; the autocorrelation of quarterly consumption volatility (which is 0.60), and the correlation between consumption growth and its conditional volatility, which is significantly negative. The joint significance of the latter three statistics is very strong, with a p value of less than 0.0001.

Our final set of moments concern time-varying skewness in the conditional distribution of consumption growth. To generate these statistics, we again estimate a reduced-form model

[6] The block-bootstrap method draws samples of length 40 quarters with replacement which can start from any point in the sample. In each bootstrap, six such blocks are drawn to approximately match our sample length of 224 quarters. For every bootstrapped sample, we calculate the statistics in Panels A, B, and C of Table 1 exactly as we do for the data sample, including estimating the required auxiliary models. We use 10,000 bootstrapped samples.

using the quarterly consumption data that allows for time-variation in the conditional third moment. The conditional third moment is specified as a function of past innovations in a fashion analogous to that of the GARCH model. To avoid imposing any distributional assumptions, we estimate this model by least squares. The exact methodology is described in the Appendix. This model provides five statistics: the unconditional mean, volatility, and autocorrelation of the conditional third moment, and its correlations with the conditional volatility of consumption (from the GARCH estimation) and consumption growth itself. If consumption growth were conditionally Gaussian, all of these moments would be zero. Results from this estimation are reported in Panel D of Table 1. We find that although the unconditional mean of the third moment is not statistically different from zero, its volatility, autocorrelation and correlation with consumption growth are significantly positive. The correlation between the volatility and third moment is estimated to be negative, but not significantly so. A joint test for the significance of these five moments strongly rejects that they are all zero, with a p-value of 0.0016. A time series for the conditional third moment (and for scaled skewness) are plotted in Figure 1.

Panels A, C, and D of Table 1 (excluding the annual statistics) provide 13 statistics, denoted by v, which serve to identify seven parameters in the model. Denote the true model parameters by the vector, θ_0; the 7 parameters to be estimated are,

$$\theta = [\overline{g}, \sigma_{cp}, \sigma_{cn}, \overline{p}, \overline{n}, \rho_n, \sigma_{nn}]'. \tag{37}$$

Under the null hypothesis that our model is true,

$$v_0 = h(\theta_0) \tag{38}$$

where $h(\theta)$ is a vector-valued function that maps the structural parameters into the reduced-form statistics. The $h(\theta)$ statistics are computed using a long (100,000 months long) simulation of the model. Given the simulated monthly series, we first aggregate the data to the quarterly frequency. We then calculate all the fitted statistics, $h(\theta)$, including those from the auxiliary models (which we also re-estimate in the simulated data). To calculate

estimates for the parameters, $\widehat{\theta}$, we minimize an objective function of the form,

$$\min_{\theta \in \Theta} \{\widehat{v} - h(\theta)\}' \widehat{W} \{\widehat{v} - h(\theta)\} \tag{39}$$

where $\widehat{W}$ is a symmetric, positive semi-definite, data-based weighting matrix. Efficient CMD suggests setting $\widehat{W} = \widehat{V}^{-1}$, where $\widehat{V}$ is the estimated covariance matrix of v_0. However, we use a diagonal weighting matrix, $\widehat{W} = diag\left(\widehat{V}^{-1}\right)$, because some of the statistics are nearly linearly dependent, rendering $\widehat{V}$ nearly singular. Standard asymptotic arguments lead to a Gaussian limiting distribution of $\widehat{\theta}$, even under our nonstandard weighting matrix (which is further described in the appendix).

IV.C Parameter Estimates

The parameter estimates (with standard errors in parentheses) are reported in Table 2. The annual average growth rate of consumption is 1.8 percent (0.0015 times 1200). The shape parameter for the good environment variable is estimated to be 11.43. At this value for p, the "good environment" shock is only mildly non-Gaussian (with skewness and excess kurtosis both around 0.6). The n_t process is persistent (with an autocorrelation coefficient of 0.9) and its unconditional mean is much lower than that of the good environment shock. Therefore, the bad environment shock is much more non-Gaussian, with unconditional skewness equal to -1.60 and excess kurtosis equaling 3.85. Despite the very different magnitudes of the shape parameters, the relative size of the scale parameters is such that the relative contribution of "bad" variance (coming from the $\omega_{n,t}$ shock) and "good" variance (attributed to $\omega_{p,t}$) is, unconditionally, about even, with $\omega_{p,t}$ accounting for about 50 percent of the total variance.

To assess how well the estimated model fits the data, re-consider Table 1, where we report the BEGE model-implied statistics for the moments in rows labeled "BEGE." When they are bolded, they are within 1.96 standard errors of the data moment (relevant for a 5 percent two-sided test) and when they are under-scored they are only within 2.58 standard errors of the data statistic (relevant for a 1 percent two-sided test). The table reveals that for 12 of the 13 moments, the BEGE model delivers statistics that are within 1.96 standard errors

of their counterparts in the data, but the model under-predicts the volatility of the GARCH-derived volatility series, being slightly over 2 standard errors away from the data moment. A model in which monthly consumption growth follows an i.i.d Gaussian distribution (the rows labeled "Gaussian") generates some realistic correlation statistics for the conditional volatility and third moment of time-aggregated quarterly consumption growth. However, it generates no meaningful volatility in either the conditional volatility or the conditional third moment. An overidentification test is available for the CMD estimation, and while the BEGE model is only marginally rejected at the 5 percent level (with the test producing a p-value of 0.045), the Gaussian model is very strongly rejected (with a p-value of 0.0005). For the annual statistics in Panel B, which are not used in the estimation, the BEGE model produces volatility that is somewhat too low, skewness that is a bit less negative and kurtosis that is a bit stronger than the point estimates from the data. The Gaussian model fails to generate nontrivial skewness or kurtosis.

Time-variation in n_t induces changes in the conditional distribution of monthly consumption growth and Table 3 shows how its conditional volatility, skewness and excess kurtosis vary with n_t realizations. The n_t variable goes from 0.44 at the 1^{st} percentile of its distribution to 4.64 at its 99^{th} percentile, increasing the annualized volatility of consumption growth from 0.90 percent to 1.62 percent and lowering its skewness from 0.06 to -0.55. As n_t increases, it accounts for an increasingly larger part of the variance of consumption growth, reaching 76 percent at the 99th percentile of its distribution.

To better appreciate the differences between our fitted time-varying BEGE distribution and the static Gaussian distribution, Figure 5 shows the cumulative distribution function for quarterly consumption growth of our estimated BEGE distribution using a log base 10 scale for better visibility of the left tail. The vertical line is drawn at the minimum quarterly consumption growth observation, which is -4.3 percent (at an annual rate). Clearly, the BEGE model has substantial mass that goes beyond this and there is a 0.50 percent chance that consumption growth is lower than this minimum. Also plotted is a Gaussian distribution

matching the mean and standard deviation of the data. For this distribution, there is only a 0.01 percent chance that consumption growth is lower than the minimum observation in the sample. To provide further background on the relative reasonableness of the tail behavior of both distributions, consider the horizontal line. It indicates the 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 (at an annual rate) or worse 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 annual real consumption growth registered -9 percent in 1932, and annual real GDP fell by 7 percent in Greece in 2011. For the BEGE distribution instead, the likelihood of an asteroid strike is similar to experiencing (annualized) consumption growth of -13.5 percent, which would indeed appear unlikely. We conclude that the BEGE model displays more plausible tail behavior than the Gaussian model.

As emphasized by Campbell and Cochrane and many others, it is difficult to observe the properties of dividend growth at a frequency higher than annual due to its strong seasonality and because such statistics are sensitive to assumptions about timing and time-aggregation. For comparability to the existing literature, we closely follow Campbell and Cochrane in setting the parameters for dividend growth. Specifically, we assume that the correlation in monthly data between consumption growth and dividend growth is 0.20, and that the (annualized) volatility of dividend growth is 11.2 percent. These statistics are in quite good agreement with those in our data sample[7]. These two moments, in conjunction with the

[7]In our quarterly sample from 1953-2013, the correlation between quarterly seasonally adjusted real per capita dividend growth and real per capita consumption growth is 0.12 (with a standard error of 0.10). The volatlity of dividend growth in our sample is 13.0 percent at an annual rate (with a bootstrapped standard error of 3.0). In our calibration, we are thus assuming that dividends are less volatile than we actually observe (though well within one standard error). Importantly, our estimated sample volatility is strongly influenced by a one-time large dividend payment by Microsoft in 2004:Q4, which caused dividend growth to

parameters for consumption growth and the specification in Equation (9) are sufficient to identify σ_{dp} and σ_{dn} as -0.0055 and 0.0217, respectively.

IV.D Calibration of the Preference Parameters

We have four preference parameters to calibrate: γ, b, ϕ, and δ. We choose parameters that minimize the distance between moments of asset prices as implied by the model, which are calculated using a 100,000 month long simulation, and their counterparts in the data. However, because it is computationally expensive to solve for equity prices, it is not feasible to formally estimate these parameters. Instead, we use grid search and a two-step process. We begin each iteration of the grid search with $[\gamma^g, \phi^g]$, where the superscripts, g, denote elements of the grid. Conditional on $[\gamma^g, \phi^g]$, we find the values for $[\delta^g, b^g]$ that enable the BEGE model to match the mean and volatility of the short rate, $E[rf_t]$ and $\sigma[rf_t]$, exactly. Because the real short rate is unobserved, we use the results in Ang, Bekaert and Wei (2008) finding the unconditional mean and volatility of the real short rate to be 1.24 and 1.46 percent, respectively. This "inner optimization" is computationally inexpensive and reduces the dimensionality of the grid search while imposing realistic short rate properties. Having completed the inner optimization and thus armed with the four-tuple $[\gamma^g, b^g, \phi^g, \delta^g]$, we impose the restrictions described above to ensure that our parameterization meets the conditions for habit,[8] and solve for equity prices. We choose the point on the grid that solves,

$$\min_{grid[\gamma,\phi]} [m_{model} - \widehat{m}_{data}] \widehat{W} [m_{model} - \widehat{m}_{data}]' \tag{40}$$

where $\widehat{m}_{data}$ are estimated moments in the data, m_{model} are their counterparts under the long simulation of the BEGE model, and $\widehat{W}$ is a symmetric, positive definite weighting matrix. The search area is initially set such that $\gamma \in [2, 40]$ and $\phi \in [0.9900, 0.9999]$ and we identify

spike to nearly 50 percent (not at an annual rate). If we exclude that observation, the sample volatility of our dividend series drops to 10.0 percent. The value for dividend volatility that we use thus lies comfortably between these two estimates.

[8] In particular, $\overline{S}_t = \sqrt{\left(\sigma_{cp}^2 \overline{p} + \sigma_{cn}^2 n_t\right) \frac{\gamma}{1-\phi-b/\gamma}}$ and $\overline{s} = \ln\left(\overline{S}\right)$ where $\overline{S} = \overline{S}_t$ evaluated at $n_t = \overline{n}$.

γ to a tolerance of 0.01, and ϕ to a tolerance of 0.0001. The unconditional moments of asset prices that we seek to match are:

$$\widehat{m}_{data} = \left\{ \frac{E[xret_t]}{\sigma\left[xret_t\right]}, \sigma\left[xret_t\right], E\left[\ln\left(PD_t\right)\right], \sigma\left[\ln\left(PD_t\right)\right], ac\left[\ln\left(PD_t\right)\right] \right\} \tag{41}$$

where ac denotes the first order autocorrelation coefficient. For $\widehat{W}$, we begin with the inverse of $\widehat{V} = COV\left(\widehat{m}_{data}\right)$ as would be optimal in a classical minimum distance estimation setting, where $COV\left(\widehat{m}_{data}\right)$, is the estimated covariance matrix for the parameters, estimated using a bootstrap procedure.[9] However, following Campbell and Cochrane, we require the model to match one moment very closely: the unconditional Sharpe ratio for equity. To implement this within our minimization routine, we effectively divide the volatility of the sampling error of this moment by a factor of 100. Formally,

$$\widehat{W} = \left[\widetilde{D}\widehat{R}\widetilde{D}\right]^{-1} \tag{42}$$

where $\widetilde{D}$ is a diagonal matrix of the standard deviations of the moments, $\widehat{m}_{data}$, except that the standard deviation for $\frac{E[xret_t]}{\sigma[xret_t]}$ is divided by 100, and $\widehat{R}$ is the estimated correlation matrix of $\widehat{m}_{data}$.

The calibrated parameters are reported in Table 4, Panel A. The calibration procedure selects parameters that are quite similar across the Gaussian and BEGE models. As required by our procedure, both models fit the mean and volatility of the rate short rate, as estimated by Ang, Bekaert and Wei (2008), exactly. Because we heavily overweight the Sharpe ratio in our calibration, both models also match the value in our sample of 0.35 (at an annual rate) within 0.01. The surplus ratio is highly persistent and features an unconditional mean of about 22 percent (calculated as $\exp(\overline{s})$). The curvature coefficient is high, 11.44 (12.85) for the BEGE (Gaussian) model. In Panel B we report characteristics of the distribution of local curvature, $\gamma \exp\left(-s_t\right)$. Average curvature is high (around 55), but the BEGE model features a slightly lower average, standard deviation, and skewness compared to the Gaussian

[9] The block-bootstrap method draws samples of length 60 months with replacement which can start from any point in the sample. Each bootstrapped sample consists of 11 such blocks to approximate the length of our sample. We generate 10,000 bootstrapped samples and calculate statistics for each one. The matrix, $\widehat{V}$, is estimated as the covariance matrix of the statistics across bootstrapped samples.

model.

Figure 6 shows the dependence of the sensitivity function λ_t on n_t and the surplus ratio, s_t, at our calibrated parameter values. Sensitivity is generally higher when the surplus ratio is low, consistent with Campbell and Cochrane. However, λ_t falls at higher values of n_t. This is true because λ_t is decreasing in $\overline{S}_t$, which, in turn, is increasing in n_t (see Equation (23)). Because s_t and n_t are negatively correlated, this mechanism helps to reduce the volatility of the short rate to be consistent with the level in the data.

V Asset Price Results

In this section we confront the model with the data on asset returns.

V.A Data

Our asset pricing data is monthly, spanning from January 1958 through December 2013. For log nominal equity returns, ret_t, we use the log of returns to the value-weighted index in CRSP. To measure excess log returns, $xret_{t+1}$, we subtract the log nominal Treasury Bill return (using one-month bills). For the price-dividend ratio, PD_t, we use the month-end price for the CRSP index divided by trailing 12-month dividends.[10] We calculate the physical probability measure of the equity return conditional variance, $pvar_t$, in two steps. We begin with the monthly realized variance, $rvar_t^{mn}$, calculated as squared log daily returns (close-to-close, excluding dividends) over the month, as well as weekly and quarterly realized variances, $rvar_t^{wk}$ and $rvar_t^{qt}$, defined analogously. Then we project $rvar_t^{mn}$ onto one-month lags of the variables: $rvar_t^{mn}$, $rvar_t^{qt}$,and $rvar_t^{wk}$ using a robust regression.[11] The fitted values from this regression are used to measure $pvar_t$. This procedure is similar to that used

[10] We calculate an analagously defined PD_t ratio using trailing 12-month dividends in the denominator when comparing our simulated data to our sample data.

[11] We find that robust regression techniques produce substantially fewer negative fitted values for the conditional variance. The robust algorithm uses iteratively reweighted least squares with a bisquare weighting function to underweight outliers. This is the default procedure in the Matlab function "robustfit" (version 7.12.0.634).

by Corsi (2009), Drechsler and Yaron (2011) and others.

We calculate the risk-neutral conditional second, third and fourth moments of equity returns, $qvar_t$, $qskew_t$, and $qxkurt_t$ 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 obtain a panel of option prices across the moneyness spectrum for the S&P 500 index from 1996-2013 from OptionMetrics, and from 1990-1995 from DeltaNeutral. We use the option contracts that have maturity closest to one month, and filter out options that are likely to be illiquid as in Figlewski (2008). (See appendix for details.) The late start to the option price data is potentially problematic because the time period over which option prices are available may not be representative of our full sample from 1958. While we cannot observe option prices from the early period, we address this concern by using a simple algorithm to sample from the conditional distribution of those variables during the early period. Concretely, we first project (again using a robust regression algorithm), $qvar_t$, $qskew_t$, and $qxkurt_t$ contemporaneously onto the vector $\left[rvar_t^{mn}, rvar_t^{qt}, rvar_t^{wk}, \ln(PD_t)\right]$, which is observable over the entire sample period. To sample $qvar_t$, $qskew_t$, and $qxkurt_t$ during the early period, we use the predicted values from this regression plus a disturbance term that is drawn from a Gaussian distribution with variance equal to the estimated variance of the regression error term.

To conduct inference on asset price statistics and joint asset price-consumption statistics, we use a block bootstrap procedure using blocks of length 60 months, jointly sampling asset price and consumption data. Of course, many blocks have missing options price data. For bootstrapped samples in which some option price data are missing (which happens whenever a block is chosen from a period prior to 1990), we use the sampling procedure described for $qvar_t$, $qskew_t$, and $qxkurt_t$ inside the bootstrap procedure. For point estimates, we generally report the statistic from the actual data. However, because we are forced to sample (as opposed to directly observing) option price data from the early portion of the sample, there is

no available point estimate pertaining to the full sample. For statistics that are functions of option price data, we report the medians from the bootstrapped distributions as the "point estimate."

V.B Results

We now compare the fit of the BEGE and Gaussian models to a broad array of asset price statistics, and cross moments between asset prices and consumption growth.

V.B.1 Univariate statistics and model intuition

In Tables 5 and 6, we examine how the model matches univariate equity return statistics, using the Gaussian Campbell and Cochrane model as a benchmark. Table 5 focuses on the standard moments of (log) price-dividend ratios and excess equity returns. Both models generate price-dividend ratios that are somewhat too low, but generate standard deviations and autocorrelation coefficients that are within a two standard error band (See Panel A). These price-dividend ratio statistics were also used in the calibration of the two models, so the good fit is not surprising. In panel B, we focus on the standard moments for equity excess returns, but we add their skewness and kurtosis. The Gaussian habits model fits the equity premium of 5.2 percent observed in the data almost exactly whereas the BEGE mode overshoots slightly, but the statistic is still within a two standard error band of the sample moment. For volatility, we observe a similar pattern. Both models generate some negative serial correlation in equity returns, whereas it is slightly positive in the data, though not significantly so. The table shows that the standard Campbell and Cochrane model cannot generate non-Gaussianity in returns. However, the BEGE model overshoots on both skewness (too negative) and kurtosis (too fat tailed). These results are very sensitive to the parameter, $\overline{n}$. Making $\overline{n}$ slightly larger for example (one standard error larger than the point estimate) provides a better fit for all moments in Panel B, with skewness and kurtosis being comfortably within one standard error of the data moments.

In Table 6, we examine the performance of the models in matching features of the conditional higher-order moments of returns, both physical and risk-neutral. Because the Gaussian habit model does not endogenously generate non-Gaussianity in returns, it is no surprise that the Gaussian habits model fails to generate a variance premium, or meaningful risk neutral skewness and kurtosis. The BEGE model generates realistic levels and variability of the physical and risk neutral variance. The implied variance premium is only about 70 basis points, whereas it is 2 percent in the data; the variance premium's variability is also underestimated. Both moments are within 2.58 standard errors of the moments in the data however. The BEGE model generates negative risk neutral skewness that is somewhat too low on average, but its variability is well matched to the data. Similarly, the BEGE model overshoots the mean of risk neutral kurtosis a bit, but its variability is within a two standard error band of the data's moment.

Overall, the BEGE model can fit both standard salient features of asset return data and some key characteristics of the risk neutral distribution of equity returns. The fit of the BEGE model is far from perfect, but recall that none of the higher order moments where used in the calibration of the model. In Figure 7, we provide some intuition for how the model works and clarify the different roles of the bad environment and good environment shocks. We graph impulse responses 100 periods out to a one standard deviation negative shock to $\omega_{p,t}$ and a one standard deviation positive shock to $\omega_{n,t}$. Both shocks, naturally, drive down consumption and the $\omega_{n,t}$ shock drives up n_t. A negative consumption shock lowers the surplus ratio and thus, the habit features of the model apply to both shocks. However, the graphs overall show that asset prices react much more forcefully to the bad environment shock.

The short rate increases because the standard intertemporal smoothing effect dominates the precautionary savings effect (the b-coefficient in Equation (21) is positive, as is true in Wachter (2006)). The response to a bad-environment shock is, however, weaker because the precautionary savings effect is stronger (see the discussion above). This effect implies that

a bad environment shock may decrease equity prices less than a good environment shock; however, it is counterbalanced by a much sharper increase in the equity premium following a bad environment shock. Recall that the risk premium associated with bad environment shocks is larger than the premium associated with good environment shocks (See Figure 4). Moreover, dividend growth loads strongly on bad environment shocks making equity especially risky. While the equity premium increases sharply following a bad–environment shock, so does the physical variance of equity returns. Because p_t is assumed constant, the physical variance does not react to a good environment shock. Put together, the Sharpe ratio increases in bad times, whether they are driven by good environment or bad environment shocks, but the effect is larger following a bad environment shock. Finally, the effect of a bad environment shock on the risk neutral variance is larger than its effect on the physical variance implying that the variance premium sharply increases after a bad environment shock.

To bring home how BEGE dynamics and stochastic risk aversion interact to generate features of asset prices, Figure 8 shows two heat maps depicting the dependence of the price dividend ratio and the variance premium on both n_t and local curvature, $\gamma \exp(-s_t)$. As in Campbell and Cochrane, the price-dividend ratio, the price of a claim to a long-lived asset, is driven primarily by s_t, which has a relatively high degree of persistence. Deviations of n_t from its unconditional mean are more short-lived, and are less important for the price-dividend ratio. Not surprisingly, the variance premium depends positively on both variables. Moreover, the apparent curvature in the contour lines indicates an intuitive nonlinearity in the model. Absent substantial negative skewness from the "bad environment," (i.e. at low values of n_t) even a large amount of risk aversion cannot generate a meaningful variance premium. Similarly, even at high levels of n_t the variance premium is small when risk aversion is low. Given the convexity of the contour lines, variation in the variance premium is larger when risk aversion and n_t tend to move up and down together, which is a feature of the BEGE habit model.

V.B.2 Consumption growth and asset prices

In Table 7 (Panel A), we focus on the cross-correlation between consumption growth and asset prices. To do so, we align quarter-end values of our (monthly) asset price series with our quarterly consumption growth data. For returns, we use cumulative quarterly values. The correlation between consumption growth and equity excess returns for both the Gaussian and BEGE models is around 55 percent, whereas the sample correlation is only 20 percent. The BEGE model therefore does not escape the excessive correlation between consumption growth and equity returns that is typical for consumption-based asset pricing models. However, the data do show strong correlations between asset prices and consumption growth that have not been stressed previously. For example, we also report the correlation between consumption growth and the physical and risk neutral variance of equity returns, which is strongly negative in the data. The Gaussian model generates slightly negative values for these correlations, whereas the BEGE model delivers negative values within a two standard error band of the data moments.

The key contribution of the BEGE dynamics for the standard habit model is that it allows the conditional distribution of consumption growth to vary with changes in n_t, the shape parameter of the bad environment shocks. As n_t increases, consumption growth becomes more variable and left skewed, as demonstrated in Table 3. Figure 2, as already discussed in the introduction, shows that shifts in the conditional distribution of consumption growth are correlated with option prices as measured by the risk neutral conditional variance ("*qvar*"). We show the unconditional distribution using squares and two conditional distributions using arrows. The conditioning uses the risk neutral conditional variance of equity returns (*qvar*); down pointing (up pointing) arrows denote the distribution of consumption growth for quarters following a quarter in which *qvar* was in the upper (lower) decile of its sample distribution. In the data, it is apparent that the "high *qvar*" distribution not only has higher variance than the "low *qvar*" distribution but also much lower skewness. Statistically, we investigate shifts in the quantiles of the distribution of consumption growth and find that the

shifts in the lower tail from the high to low $qvar$ distribution are statistically significant at the 5 or 10 percent level. The bottom panel of the figure shows that the BEGE model mimics the shifts in the distribution rather well, with the lower tail shifting more downward than the upper tail shifts upward. In Table 8, (Panel B), we reproduce the statistics for the 5th and 95th percentiles. Of course, the Gaussian model does not produce any conditional shift in the distribution of consumption growth following "high $qvar$" or "low $qvar$" environments, by assumption. The BEGE model fits the shift in the 5 percent quantile, which amounts to -2.40 percent in the data, almost exactly. At the 95th percentile, the distribution in the data shifts downward in an insignificant manner, whereas the BEGE distribution shifts upward by about 50 basis points. However, this is still within a two standard error band of the data moment.

V.B.3 Asset price cross-moments

In Table 8, we focus on cross-correlations between different financial variables. First, in Panel A, we report the contemporaneous correlation between excess equity returns and the conditional physical and risk neutral equity return variance. This correlation is highly significantly negative at -0.49, respectively -0.46, reflecting the well-known asymmetric volatility phenomenon (see e.g. Wu, 2001). The BEGE model generates substantially negative correlations which are within a two standard error band of the data moments, whereas the Gaussian model generates an economically insignificant negative correlation.

In Panel B of Table 8, we repeat the standard predictability regressions of equity returns, ranging from one month to 120 month returns, on the log price-dividend ratio observed at the beginning of the holding period, as conducted in Campbell and Cochrane. The returns are expressed on a per-month basis. We report the slope coefficients and R^2 statistics, with bootstrapped standard errors. The betas are negative but they are not significantly different from zero. The R^2 increases from less than 0.5 percent at the one month horizon to 24 percent at the 10 year horizon, but even that large R^2 is barely statistically significantly

different from zero. The BEGE model delivers statistics closer to the point estimates and the observed R^2s than the Gaussian model, but the large standard errors in the data make this a weak test, and none of the statistics are outside of a two standard error band around the point estimates. We note that the R^2 does not increase with horizon in the model simulation statistics as it does in the data because we use a long simulation devoid of finite sample biases which are the source of the increases in R^2 in the sample.[12]

VI Conclusion

In this article, we introduce a model for consumption growth, the BEGE model, that can accommodate time-varying uncertainty, tail risk , and flexible time-variation in higher order moments. After inserting these features into a standard asset pricing model with preferences determined by habit persistence as in Campbell and Cochrane (1999), the model matches various salient features of consumption dynamics and asset prices better than versions with i.i.d. Gaussian shocks. For example, we document substantial time-variation in both the conditional variance and the conditional skewness of consumption growth, with these moments being negatively correlated. Moreover, consumption growth exhibits a longer left tail when risk-neutral equity volatility is high, evidence of a strong link between asset prices and the conditional distribution of macroeconomic outcomes. The model also delivers an unconditional left-skewed and leptokurtic distribution where bad consumption growth states are more likely than under a normal distribution. The BEGE model can accommodate some aspects of "disaster risk" as specified, for example in Gabaix (2012). However, at our point estimates, identified using post WWII U.S. data, the bad states that we identify are more consistent with persistent recession behavior rather than extreme, relatively transitory disaster shocks.

When coupled with the habit preference structure of Campbell and Cochrane (1999),

[12]Simulations of the BEGE model do reproduce the pattern of R^2 increasing with horizon when the simulated samples are generated with length equal to that of the data sample.

the non-Gaussianities in consumption growth deliver substantial, variable, 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. The model also suggests an intriguing new fact that links consumption growth to option prices: when the VIX is high, the distribution of consumption growth is more left-skewed.

There are many useful directions for future research. Building on Bekaert, Engstrom and Grenadier (2010), it is straightforward to formulate a BEGE model where risk aversion is stochastic and heteroskedastic but the price of risk does not vary through time as it does in Campbell and Cochrane. In such a model, we can derive closed-form solutions for equity prices and bonds with multi-period maturities, which greatly simplifies the numerical computation. We defer a full analysis of such a model to future work.

A Appendix

1. Symbol and variable quick lookup table

In order of first appearance:

parameters...	
...governing consumption and dividend growth	
$\overline{g}$	mean of log consumption growth
σ_{cp}	loading of consumption growth onto "good environment" shock
σ_{cn}	loading of consumption growth onto "bad environment" shock
σ_{dp}	loading of dividend growth onto "good environment" shock
σ_{dn}	loading of dividend growth onto "bad environment" shock
$\overline{n}$	unconditional mean of "bad environment" state variable
ρ_n	autocorrelation coefficient of "bad environment" state variable
σ_{nn}	conditional volatility parameter of "bad environment" state variable
$\overline{p}$	unconditional mean of "good environment" shape parameter
... governing preferences	
δ	subjective discount factor
γ	curvature parameter for utility function
$\overline{s}$	mean of log surplus ratio
ϕ	autocorrelation coefficient for the surplus ratio
b	parameter governing the dependence of the risk free rate on the surplus ratio

variables...	
Δc_t	log consumption growth (real, per capita)
Δd_t	log dividend growth
$\omega_{p,t}$	"good environment" shock
$\omega_{n,t}$	"bad environment" shock
p_t	"good environment" state variable
n_t	"bad environment" state variable
H_t	habit stock
Q_t	local curvature of the utility function, γ/S_t
$S_t\left(s_t\right)$	level (log) surplus consumption ratio
$s_{t,\max}$	maximum conditional level of surplus ratio
$M_t\left(m_t\right)$	(log) stochastic discount factor / pricing kernel
λ_t	sensitivity function
$a_{p,t}$	loading of pricing kernel on good environment shock
$a_{n,t}$	loading of pricing kernel of bad environment shock
rf_t	real (log) short rate
PD_t	price-dividend ratio for equity
$(x)\, ret_t$	log (excess) equity return
$vprem_t$	variance premium for equity returns
$pvar_t$	physical conditional variance of returns
$qvar_t$	risk-neutral conditional variance of returns
$qskw_t$	risk-neutral conditional skewness of returns
$qxkurt_t$	risk-neutral conditional excess kurtosis of returns

2. Moment generating functions for gamma and BEGE distributions

For a variable, x, following a centered gamma distribution with shape and scale parameters of k and θ, respectively, the probability density function, $g\left(x;k,\theta\right)$ is given by

$$g\left(x;k,\theta\right)=\frac{\left(x+k\theta\right)^{k-1}e^{-\frac{x+\kappa\theta}{\theta}}}{\theta^{k}\Gamma\left(k\right)} \tag{43}$$

The moment generating function, $M_x\left(t;k,\theta\right)$ is given by

$$M_x\left(t;k,\theta\right)=E\left[\exp\left(tx\right)\right]=\exp\left(-k\ln\left(1-\theta t\right)-k\theta t\right) \tag{44}$$

for $t<\frac{1}{\theta}$. Our "BEGE" model for consumption growth, Δc_t, takes the form,

$$\Delta c_t=g+\sigma_{cp}\omega_{p,t}-\sigma_{cn}\omega_{n,t} \tag{45}$$

where $\omega_{p,t}$ and $\omega_{n,t}$ each have centered gamma distributions with shape parameters p_{t-1} and n_{t-1}, unit scale parameters, and are independent. Solutions for asset prices under our model

generally require evaluation of expectations of the form,

$$E_t\left[\exp\left(a+b\Delta c_{t+1}\right)\right] \tag{46}$$

which, using the above result for $M_x\left(t;k,\theta\right)$, has the solution

$$\begin{aligned} E_t\left[\exp\left(a+b\Delta c_{t+1}\right)\right] &= \exp\left(a+bg\right)E_t\exp\left(b\sigma_{cp}\omega_{p,t+1}-b\sigma_{cn}\omega_{n,t+1}\right) \\ &= \exp\left(a+bg\right)\exp\left(-p_t\ln\left(1-b\sigma_{cp}\right)-p_tb\sigma_{cp}\right)\exp\left(-n_t\ln\left(1+b\sigma_{cp}\right)+n_tb\sigma_{cp}\right) \\ &= \exp\left[a+bg-p_tf\left(b\sigma_{cp}\right)-n_t\left(-b\sigma_{cn}\right)\right] \end{aligned} \tag{47}$$

where the function $f\left(x\right)=x+\ln\left(1-x\right)$.

3. Certainty-equivalent means for BE/GE distributions

Denote the Gaussian-distributed log-wealth distribution as $\widetilde{w}_G$ and the bad-environment-distributed version as $\widetilde{w}_{BE}$. The latter has a mean of $\left(\mu_w+\Delta_n\right)$ and a shock distributed as a negative centered gamma with shape parameter n, and scale parameter, $\sigma_w/\sqrt{n}$ (to set the variance equal to σ_w^2). We set the expected utility of the two shocks as equal and solve for Δ_n.

$$E\left[U\left(\exp\left(\widetilde{w}_G\right)\right)\right]=E\left[U\left(\exp\left(\widetilde{w}_{BE}\right)\right)\right] \tag{48}$$

$\Leftrightarrow$

$$E\left[\frac{\exp\left(\widetilde{w}_G\right)^{1-\gamma}-1}{1-\gamma}\right]=E\left[\frac{\exp\left(\widetilde{w}_{BE}\right)^{1-\gamma}-1}{1-\gamma}\right] \tag{49}$$

$\Leftrightarrow$

$$E\left[\exp\left(\widetilde{w}_G\right)^{1-\gamma}\right]=E\left[\exp\left(\widetilde{w}_{BE}\right)^{1-\gamma}\right] \tag{50}$$

$\Leftrightarrow$

$$\exp\left(\left(1-\gamma\right)\mu_w+\frac{1}{2}\sigma_w^2\left(1-\gamma\right)^2\right)=\exp\left(\left(1-\gamma\right)\left(\mu_w+\Delta_n\right)-n\cdot f\left(\sigma_w\left(1-\gamma\right)/n^{1/2}\right)\right) \tag{51}$$

$\Leftrightarrow$

$$\Delta_n=\frac{1}{1-\gamma}\left[\frac{1}{2}\sigma_w^2\left(1-\gamma\right)^2+n\cdot f\left(\sigma_w\left(1-\gamma\right)/n^{1/2}\right)\right] \tag{52}$$

where $f\left(\cdot\right)$ is the function defined in the text. A third order approximation of this function is presented in Equation (8). Results for Δ_p are calculated similarly.

4. The Maximum Sharpe Ratio

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

$$\begin{aligned}\max_{\{\text{all assets}\}} \frac{E_t\left[XRET^e_{t+1}\right]}{\sigma_t\left[XRET^e_{t+1}\right]} &= \frac{\sigma_t\left[\exp\left(m_{t+1}\right)\right]}{E_t\left[\exp\left(m_{t+1}\right)\right]} = \sqrt{\frac{E_t\left[\exp\left(2m_{t+1}\right)\right]}{E_t\left[\exp\left(m_{t+1}\right)\right]^2} - 1} \\ &= \sqrt{\left[\frac{1-a_{p,t}}{\left(1-2a_{p,t}\right)^{1/2}}\right]^{2p}\left[\frac{1-a_{n,t}}{\left(1-2a_{n,t}\right)^{1/2}}\right]^{2n} - 1} \qquad (53)\end{aligned}$$

where, as defined above,

$$\begin{aligned}a_{p,t} &= -\gamma\left(1+\lambda_t\right)\sigma_{cp} \\ a_{n,t} &= +\gamma\left(1+\lambda_t\right)\sigma_{cn} \qquad (54)\end{aligned}$$

Some approximations are helpful for gaining intuition.

$$\begin{aligned}\frac{\sigma_t\left[\exp\left(m_{t+1}\right)\right]}{E_t\left[\exp\left(m_{t+1}\right)\right]} &= \sqrt{\exp\left\{2p\ln\left[\frac{1-a_{p,t}}{\left(1-2a_{p,t}\right)^{1/2}}\right] + 2n\ln\left[\frac{1-a_{n,t}}{\left(1-2a_{n,t}\right)^{1/2}}\right]\right\} - 1} \\ &\approx \sqrt{2p\ln\left[\frac{1-a_{p,t}}{\left(1-2a_{p,t}\right)^{1/2}}\right] + 2n\ln\left[\frac{1-a_{n,t}}{\left(1-2a_{n,t}\right)^{1/2}}\right]} \\ &\approx \sqrt{p\left(a^2_{p,t}+2a^3_{p,t}\right) + n_t\left(a^2_{n,t}+2a^3_{n,t}\right)} \qquad (55)\end{aligned}$$

where the second line uses a Taylor approximation for the $exp\left(\right)$ function, and the third line uses Taylor approximations for the functions in square brackets.

5. Auxiliary models for consumption growth

We estimate a version of the GARCH model due to Glosten Jagannathan and Runkle (1993). Specifically, for quarterly consumption growth (for both actual and bootstrapped samples, as well for data simulated under the BEGE model), we begin by demeaning the process and removing MA(1) variation in the conditional mean that approximates the persistence induced by temporal aggregation of the underlying serially uncorrelated process, producing an MA(1) autocorrelation of 0.21. Denote this "pre-whitened" series as ε_t. We then estimate the model,

$$\begin{aligned}\varepsilon_{t+1} &\sim N\left(0, h_t\right) \\ h_t &= h_0 + \alpha_0 h_{t-1} + \alpha_1\varepsilon^2_t + \alpha_2 I_t\varepsilon^2_t \\ I_t &= \begin{cases} 1 & if\ \varepsilon_t < 0 \\ 0 & if\ \varepsilon_t \geq 0 \end{cases} \qquad (56)\end{aligned}$$

Estimation by maximum likelihood using the sample data for consumption growth (multiplied by 400 to produce units at and annual rate and in percentage terms) produces the

following results.

h_0	α_0	α_1	α_2
1.1169	0.3971	0.1665	0.2349
(0.5796)	(0.1916)	(0.1367)	(0.1151)

(57)

with bootstrapped standard errors in parentheses. We also estimate a model for time-variation of the conditional third moment:

$$\begin{aligned} \nu_t &= E_t\left(\varepsilon_{t+1}\right)^3 \\ \nu_t &= \nu_0 + \beta_0 \nu_{t-1} + \beta_1 \varepsilon_t^2 + \beta_2 I_t \varepsilon_t^2 \end{aligned} \tag{58}$$

This model is estimated by least squares. Using cubed residuals makes the estimation poorly behaved, so we use squared residuals. Conditional on the parameter vector, $\left[\widehat{\nu}_0, \widehat{\beta}_0, \widehat{\beta}_1, \widehat{\beta}_2\right]$, and an initial value of 0, a sequence $\{\widehat{\nu}_t\}_{t=1}^T$ is generated, and we calculate a quadratic loss function,

$$\sum_{t=2}^{T} \left(\varepsilon_{t+1}^3 - \widehat{\nu}_t\right)^2 \tag{59}$$

The point estimates are chosen to minimize the loss function. The resulting point estimates are

ν_0	β_0	β_1	β_2
−0.3905	0.5824	0.4370	−0.9775
(1.0849)	(0.2626)	(0.3578)	(0.6248)

(60)

where bootstrapped standard errors are reported in parentheses. A joint test that all four parameters above are equal to zero yields a p-value of 0.0017 using a Wald test and a bootstrapped covariance matrix for the parameters. A test for the significance of the time-variation in $\widehat{v}_t$, that is, the hypothesis that $[\beta_0, \beta_1, \beta_2] = 0$, is rejected with a p-value of 0.0224.

6. CMD Estimation Asymptotics

First note that the first order condition for our CMD optimization, described in Equation (39), is

$$\widehat{H}'\widehat{W}\left\{\widehat{v} - h\left(\widehat{\theta}\right)\right\} = 0. \tag{61}$$

where $\widehat{H} = \nabla_\theta h\left(\widehat{\theta}\right)$ is the Jacobian of $h\left(\theta\right)$ evaluated at $\widehat{\theta}$. Second, using a standard mean value expansion,

$$h\left(\widehat{\theta}\right) = h\left(\theta_0\right) + H_0\left(\widehat{\theta} - \theta_0\right). \tag{62}$$

where $H_0 = \nabla_\theta h\left(\theta_0\right)$ is the gradient of $h\left(\theta\right)$ at the true parameter values. Combining Equations (61) and (62), we have,

$$\widehat{H}'\widehat{W}H_0\left(\widehat{\theta} - \theta_0\right) = \widehat{H}'\widehat{W}\left(\widehat{v} - h\left(\theta_0\right)\right) \tag{63}$$

so that under the usual arguments, the limiting distribution of the structural parameters is,

$$\left(\widehat{\theta}-\theta_0\right) \sim N\left(0, \widehat{V}_\theta\right) \tag{64}$$

where $\widehat{V}_\theta = \left(\widehat{M}^{-1}\widehat{H}'\widehat{W}\widehat{V}\widehat{W}\widehat{H}\widehat{M}^{-1}\right)$, $\widehat{M} = \widehat{H}'\widehat{W}\widehat{H}$, and $\widehat{V}$ is the variance-covariance matrix of the statistics, $\widehat{v}$.

7. *Measuring option-implied conditional equity return moments*

We follow Bakshi, Kapadia and Madan (2003) to measure the risk-neutral moments of one-month ahead equity returns.. For each month in the sample, the risk neutral moments are calculated using a portfolio of options indexed by their strikes. Let the price of the underlying index at time t be denoted by S_t, and let the price of a European call (put) option with one month to expiration and strike price K be a function denoted by $C_t(K)$ ($P_t(K)$). Under some technical conditions outlined in the paper, the price of a contract with a stochastic payoff equal to the realized variance of returns is shown to be (Bakshi, Kapadia and Madan refer to this as a "volatility contract").

$$VV_t = \int_{S(t)}^{\infty} \frac{2\left(1-\ln\left(\frac{K}{S_t}\right)\right)}{K^2} C_t(K)\,dK + \int_0^{S(t)} \frac{2\left(1+\ln\left(\frac{K}{S_t}\right)\right)}{K^2} P_t(K)\,dK \tag{65}$$

Similar expressions are derived for the prices of portfolios of options with payoffs equal to the realized cubic and quartic variation of returns:

$$\begin{aligned} WW_t &= \int_{S(t)}^{\infty} \frac{6\ln\left(\frac{K}{S_t}\right) - 3\left[\ln\left(\frac{K}{S_t}\right)\right]^2}{K^2} C_t(K)\,dK - \int_0^{S(t)} \frac{6\ln\left(\frac{K}{S_t}\right) + 3\left[\ln\left(\frac{K}{S_t}\right)\right]^2}{K^2} P_t(K)\,dK \qquad (66)\\ XX_t &= \int_{S(t)}^{\infty} \frac{12\left[\ln\left(\frac{K}{S_t}\right)\right]^2 - 4\left[\ln\left(\frac{K}{S_t}\right)\right]^3}{K^2} C_t(K)\,dK - \int_0^{S(t)} \frac{12\left[\ln\left(\frac{K}{S_t}\right)\right]^2 + 4\left[\ln\left(\frac{K}{S_t}\right)\right]^3}{K^2} P_t(K)\,dK \end{aligned}$$

where WW_t and XX_t are the prices of the cubic and quartic contracts. Let the risk free nominal interest rate with one month maturity be nf_t. Risk-neutral skewness and kurtosis are calculated as

$$\begin{aligned} qvar_t &= \exp(nf_t)\,VV_t - nu_t^2 \qquad (67)\\ qskw_t &= \frac{\exp(nf_t)\,WW_t - 3\exp(nf_t)\,nu_t VV_t + 2nu_t^3}{qvar_t^{3/2}}\\ qxkurt_e &= \frac{\exp(nf_t)\,XX_t - 4\exp(nf_t)\,nu_t WW_t + 6\exp(nf_t)\,nu_t^2 VV_t - 3nu_t^4}{qvar_t^{4/2}} - 3 \end{aligned}$$

where $nu_t = \exp(nf_t) - 1 - \exp(nf_t)\,VV_t/2 - \exp(nf_t)\,WW_t/6 - \exp(nf_t)\,XX_t/24$.

To implement this approach, we use a panel of option prices from DeltaNeutral (1990-1995) and OptionMetrics (1996-2013) with the S&P 500 index as the underlying. We use

option price observations from dates at month end and with time-to-maturity closest to one month. We filter out observations for which the reported spot price of the S&P 500 index from DeltaNeutral or OptionMetrics differs from the published value of the S&P 500 index by more than 2.5 percent options or with prices below \$0.50 per contract. Given this set of option prices, we next approximate the functions $C_t(K)$ and $P_t(K)$ numerically by interpolation, which we carry out in implied-volatility space. We calculate the Black-Scholes implied volatility for each call and put option between the minimum and maximum observed strikes, $K_{\min}^{obs}$ and $K_{\max}^{obs}$, respectively. Then, separately for calls and puts, we interpolate the implied volatilities on a grid from $K_{\min}^{obs}$ to $K_{\max}^{obs}$ at equally spaced at intervals of 5 S&P 500 index points. Because the limits of the integrations in Equations (66) in principle run from 0 to infinity, we also need option prices with very low and very high strikes. We set values for implied volatility for strikes from $K = 5$ to $K = K_{\min}^{obs}$ (again at intervals of 5) to the implied volatility at $K_{\min}^{obs}$. Similarly, we set the implied volatilities over a range from from $K_{\max}^{obs}$ to $2S_t$ equal to the implied volatility at $K_{\max}^{obs}$. We then convert all options prices back into price space, yielding values for $C_t(K)$ and $P_t(K)$ over the entire grid of $K = 5$ to $K = 2S_t$. The integrations in Equations (66) are carried out over the grid of option prices using trapezoidal numerical integration.

References

Ang, A., Bekaert, G., Wei, M., 2008, The Term Structure of Real Rates and Expected Inflation, Journal of Finance, 63, 797-849.

Bakshi, G., and Madan, D., 2000, Spanning and Derivative-Security Valuation, Journal of Financial Economics, 55, 205-238.

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.

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

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

Bekaert, G., Hoerova, M., 2014, The VIX, the Variance Premium and Stock Market Volatility, Journal of Econometrics, 183, 181-192.

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.

Breeden, D., and Litzenberger, R., 1978, Prices of State-Contingent Claims Implicit in Option Prices, Journal of Business, 51, 621-651.

Britten-Jones, M., Neuberger, A., 2000, Option Prices, Implied Price Processes, and Stochastic Volatility, Journal of Finance, 55, 839-866.

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.

Chicago Board Options Exchange (CBOE), 2014, The CBOE Volatility Index - VIX, white paper.

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.

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.

Longstaff, F. and Piazzesi, M., 2004, Corporate Earnings and the Equity Premium, Journal of Financial Economics, 72, 401-421.

Miller, A., Rice, T., 1983, Discrete Approximations of Probability Distributions, Management Science, 29(3), 352-362.

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., 2013, Can Time-varying Risk of Rare Disasters Explain Stock Market Volatility?, Journal of Finance, 68, 987-1035.

Whaley, R., 2000, The Investor Fear Gauge, Journal of Portfolio Management, 26, 12-17.

Wooldridge, J., 2002, Econometric Analysis of Cross Section and Panel Data, MIT Press.

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: Consumption Growth Statistics

Panel A: Unconditional quarterly statistics

	mean	*std*	*skew*	*xkurt*
point est	1.6635	0.9512	−0.3135	0.7475
	(0.2422)	(0.0530)	(0.2266)	(0.3027)
BEGE	**1.6668**	**0.9477**	**−0.3992**	**1.0403**
Gaussian	**1.6672**	**0.9354**	**−0.0008**	−0.0052

Panel B: Unconditional annual statistics

	mean	*std*	*skew*	*xkurt*
point est	1.6720	1.2714	−0.5777	0.2244
	(0.2458)	(0.1207)	(0.2228)	(0.5923)
BEGE	**1.6667**	0.9242	**−0.5323**	**1.0481**
Gaussian	**1.6665**	0.9107	−0.0493	**0.0689**

Panel C: GARCH statistics

	(1)	(2)	(3)	(4)	(5)
	$mean\left(\widehat{h}_t^{1/2}\right)$	$std\left(\widehat{h}_t^{1/2}\right)$	$ac\left(\widehat{h}_t^{1/2}\right)$	$\rho\left(\Delta c_t, \widehat{h}_t^{1/2}\right)$	$pval\ (2)-(4)$
point est	1.8025	0.4700	0.6019	−0.4311	< 0.0001
	(0.0904)	(0.1171)	(0.1881)	(0.2081)	
BEGE	**1.8430**	0.2219	**0.7477**	**−0.4930**	
Gaussian	**1.8275**	0.0067	0.0075	**−0.4760**	

Panel D: Conditional third moment statistics

	(1)	(2)	(3)	(4)	(5)	(6)
	$mean\left(\widehat{\nu}_t\right)$	$std\left(\widehat{\nu}_t\right)$	$ac\left(\widehat{\nu}_t\right)$	$\rho\left(\Delta c_t, \widehat{\nu}_t\right)$	$\rho\left(\widehat{\nu}_t, \widehat{h}_t^{1/2}\right)$	$pval\ (1)-(5)$
point est	−1.5950	4.4430	0.7266	0.7574	−0.6079	0.0016
	(1.5903)	(1.5976)	(0.2575)	(0.3921)	(0.4200)	
BEGE	**−2.1870**	**1.8672**	**0.7459**	**0.7478**	**−0.9820**	
Gaussian	0.0286	0.2589	**0.5402**	**0.7044**	**−0.7058**	

Panel E: Overidentifying tests

BEGE	0.0453
Gaussian	0.0005

Table 1 notes. The data are real, per capita, consumption expenditures for nondurables and services, quarterly, from 1958Q1-2013Q4. We present the same statistics under a BEGE model of consumption growth, described in Table 2, that is specified at the monthly frequency and aggregated to the quarterly frequency, and also a monthly i.i.d. Gaussian model. Panel A reports unconditional univariate statisics. The mean is multiplied by 400 and the standard deviation is multiplied by 200. Panel B repeats these statistics for annual consumption growth (not used in estimation). The statistics reported in the Panel C are obtained by fitting a Gaussian , MA(1), GJR-GARCH(1,1) model to the consumption growth data (multiplied by 400) by maximum likelihood, and those in Panel C are from an auxiliary model to estimate the conditional skewness of the process. We assume that the MA(1) coefficient is 0.21. The series, $\widehat{h}_t$, corresponds to the resulting fitted variance series, and $\widehat{\nu}_t$ is the estimated conditional third moment. The column headings, *mean*, *std*, *skew*, and *xkurt* refer to unconditional means, standard deviations, skewness, and excess kurtosis, respectively. The column $\rho\left(\Delta c_t, \widehat{h}_t^{1/2}\right)$ denotes the unconditional correlation between consumption growth and $\widehat{h}_t^{1/2}$, and similarly for the correlations in Panel D. The columns labels *pval* refer to a chi-squared test that the statistics in the indicated columns are jointly different from zero. The numbers in parentheses are bootstrapped standard errors. Bolded figures denote that the model-implied statistics are a distance of less than 1.96 standard errors from the corresponding point estimate, and underbars denote a distance of more than 1.96 but less than 2.58 standard errors. The GJR-GARCH and conditional skewness models are described in detail in the appendix. Panel E reports the p-value for a test of the over-identifying restrictions.

Table 2: BEGE Model Parameter Estimates

Parameter	Estimate
g	0.0015
	(0.0002)
$\overline{p}$	11.4314
	(1.6125)
$\overline{n}$	1.5599
	(1.6186)
ρ_n	0.9051
	(0.0733)
σ_{nn}	0.3169
	(0.3660)
σ_{cp}	0.00067
	(0.00015)
σ_{cn}	0.0019
	(0.00010)

Table 2 notes. The BEGE model for consumption growth is described by the equations:

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

where Δc_t iis consumption growth at the monthly frequency. Standard errors are reported in parentheses. The model is estimated by simulated CMD, where the moments being matched are those reported in Table 1.

Table 3: The Conditional Distribution of Consumption Growth

percentile	n_t	$std_t(\Delta c_{t+1})$	$skew_t(\Delta c_{t+1})$	*fraction "bad var"*
1	0.44	0.90	0.06	0.24
5	0.55	0.93	−0.02	0.27
10	0.65	0.95	−0.08	0.31
25	0.89	1.00	−0.21	0.38
50	1.33	1.09	−0.35	0.48
75	1.97	1.21	−0.46	0.58
90	2.76	1.35	−0.52	0.65
95	3.33	1.43	−0.54	0.70
99	4.64	1.62	−0.55	0.76

Table 3 notes. This table reports, for the model and parameter estimates described in Table 2, the conditional distribution of monthly consumption growth conditional on various values for n_t. The column with heading n_t corresponds to the unconditional distribution of n_t in a simulation for 100,000 periods. The columns, $std_t(\Delta c_{t+1})$ and $skew_t(\Delta c_{t+1})$ correspond to the conditional standard deviation (percent, annual rate) and skewness of monthly consumption growth at the corresponding value for n_t. The column "*fraction bad var*" reports values of $\sigma_{cn}^2 n_t / \left(\sigma_{cn}^2 n_t + \sigma_{cp}^2 \overline{p}\right)$ using the parameters in Table 2.

Table 4: Asset Pricing Model Calibrations

Panel A: Parameters

	Gaussian	*BEGE*
δ	0.9999	0.9999
ϕ	0.9964	0.9964
γ	12.85	11.42
b	0.0088	0.0099
$\overline{s}$	−1.5237	−1.5583

Panel B: Distribution of local risk aversion

	mean	*std*	*skew*	$p10$	$p50$	$p90$	$p99$	$p99.9$	$p99.99$
BEGE	53.17	8.18	0.99	45.89	53.90	65.91	80.34	91.62	97.91
Gaussian	59.91	8.64	1.06	50.26	58.50	71.19	87.18	102.11	106.89

Table 4 notes: In Panel A, for both models, the parameters $[\delta, \phi, \gamma, b, \overline{s}]$ are chosen to match the mean and volatility of the real short term risk free rate exactly, and the unconditional Sharpe ratio, and volatility of excess equity returns, and the unconditional mean, volatility, and autocorrelation of the log price-dividend ratio according to a quadratic loss function. The model "Gaussian" refers to a model with i.i.d. Gaussian consumption growth. The model "BEGE" refers to the model for consumption growth described in Table 2. In Panel B, we report the simulated percentiles of the distribution of $\gamma \exp(-s_t)$ for each model.

Table 5: Univariate Equity Price Statistics

Panel A: Price-dividend ratios

moment	*data*	*Gaussian*	*BEGE*
$E\left[\ln\left(PD_t\right)\right]$	3.5517	3.0847	2.9379
	(0.1109)		
$\sigma\left[\ln\left(PD_t\right)\right]$	0.3893	**0.2689**	**0.2911**
	(0.0633)		
$ac\left[\ln\left(PD_t\right)\right]$	0.9912	**0.9865**	**0.9851**
	(0.0071)		

Panel B: Equity returns

moment	*data*	*Gaussian*	*BEGE*
$E\left[xret_{t+1}\right]$	5.2051	**5.2207**	**5.9797**
	(1.8275)		
$\sigma\left[xret_{t+1}\right]$	14.8367	**14.9259**	<u>16.9997</u>
	(0.8386)		
$ac\left[xret_{t+1}\right]$	0.0629	**−0.0033**	**−0.0161**
	(0.0400)		
$skew\left[xret_{t+1}\right]$	−0.7011	0.0014	−1.5494
	(0.2441)		
$xkurt\left[xret_{t+1}\right]$	2.5288	<u>0.0157</u>	**4.4243**
	(1.1600)		

Table 5 notes: Asset price data are observed monthly, 1958-2013. $\ln\left(PD_t\right)$ is the price-dividend ratio, defined as the log ratio of quarter-end price per share to 12-month trailing dividends-per-share for the S&P 500 index, $xret_t$ is the monthly log excess return on the S&P 500 index for which the mean is multiplied by 1200, and the standard deviation is multiplied by sqrt(12)*100 to express statistics in percentage terms at an annual rate. Bootstrapped standard errors are shown in parentheses. The model "Gaussian" refers to a model with i.i.d. Gaussian consumption growth. The model "BEGE" refers to the model for consumption growth described in Table 2. Bolded figures denote a distance of less than 1.96 standard errors from the corresponding point estimate, and underbars denote a distance of more than 1.96 but less than 2.58 standard errors.

Table 6: Equity Return Conditional Higher-Order Moments

Panel A: Conditional variance

	data	*Gaussian*	*BEGE*
$E\left[pvart_t\right]$	0.0186	**0.0221**	0.0285
	(0.0035)		
$\sigma\left[pvar_t\right]$	0.0231	0.0004	**0.0125**
	(0.0085)		
$E\left[qvar_t\right]$	0.0376	<u>0.0221</u>	**0.0351**
	(0.0073)		
$\sigma\left[qvar_t\right]$	0.0374	<u>0.0004</u>	**0.0146**
	(0.0144)		
$E\left[vprem_t\right]$	0.0206	0.0000	<u>0.0066</u>
	(0.0057)		
$\sigma\left[vprem_t\right]$	0.0223	0.0000	<u>0.0023</u>
	(0.0087)		

Panel B: Conditional higher-order moments

	data	*Gaussian*	*BEGE*
$E\left[qskw_t\right]$	-1.3534	0.0098	<u>-1.7399</u>
	(0.1481)		
$\sigma\left[qskw_t\right]$	0.4280	0.0013	**0.4834**
	(0.0980)		
$E\left[qxkurt_t\right]$	2.7676	0.0006	<u>4.8360</u>
	(0.9018)		
$\sigma\left[qxkurt_t\right]$	2.1507	0.0001	**2.6729**
	(0.2690)		

Table 6 notes: Asset price data are generally monthly, 1958-2013. $pvar_t$ is the conditional variance of log equity returns under the physical measure, $qvar_t$ is the conditional variance of log returns under the risk-neutral measure, and $vprem_t$ is the difference between the two. All variance measures are reported at an annual rate. The variables $qskw_t$ and $qxkurt_t$ are risk-neutral conditional skewness and excess kurtosis, respectively. Bootstrapped standard errors are shown in parentheses. The model "Gaussian" refers to a model with i.i.d. Gaussian consumption growth. The model "BEGE" refers to the model for consumption growth described in Table 2. Bolded figures denote a distance of less than 1.96 standard errors from the corresponding point estimate, and underbars denote a distance of more than 1.96 but less than 2.58 standard errors.

Table 7: Cross-moments between asset prices and consumption growth

Panel A: Consumption growth and return volatility

	data	*Gaussian*	*BEGE*
$[xret_t, \Delta c_t]$	0.1989	0.5349	0.5538
	(0.0540)		
$[pvar_t, \Delta c_t]$	−0.3248	−0.0360	**−0.4574**
	(0.1040)		
$[qvar_t, \Delta c_t]$	−0.2499	−0.0329	**−0.4528**
	(0.0973)		

Panel B: Consumption quantile shifts conditional on $qvar_t$

	data	*Gaussian*	*BEGE*
5% quantile shift	−2.4072	−0.1676	**−2.5714**
	(1.1010)		
95% quantile shift	−0.8675	**−0.1425**	**0.5686**
	(1.1544)		

Table 7 notes: Asset price data are monthly, 1958-2013. $xret_t$ is the monthly log excess return on the S&P 500 index and we use cumulative quarterly returns to compare with the consumption data, $pvar_t$ is the conditional variance of log equity returns under the physical measure, $qvar_t$ is the conditional variance of log returns under the risk-neutral measure. These monthly variables are measured at quarter ends to align them with the quarterly consumption data. Bootstrapped standard errors are shown below in parentheses. The model "Gaussian" refers to a model with i.i.d. Gaussian consumption growth. The model "BEGE" refers to the model for consumption growth described in Table 2. Bolded figures denote a distance of less than 1.96 standard errors from the corresponding point estimate, and underbars denote a distance of more than 1.96 but less than 2.58 standard errors. In Panel B, we report quantile shifts, the difference between the 5th (95th) percentile of the distribution of consumption growth conditional on $qvar_t$ being "high" (in the upper 10th percentile of its distribution) less the 5th (95th) percentile for consumption conditional on $qvar_t$ being "low" (in the lower 10th percentile of its unconditional distribution).

Table 8: Asset Price Cross Moments

Panel A: Asymmetric volatility correlations

	data	*Gaussian*	*BEGE*
$[pvar_{t+1}, xret_{t+1}]$	−0.4879	−0.0125	**−0.6056**
	(0.0895)		
$[qvar_{t+1}, xret_{t+1}]$	−0.4643	−0.0110	**−0.5889**
	(0.1311)		

Panel B: Predictability with respect to dividend yield

	data		*Gaussian*		*BEGE*	
Horizon	*coef*	R^2	*coef*	R^2	*coef*	R^2
1	−6.34	0.002	**−3.10**	**0.000**	**−6.80**	**0.001**
	(5.71)	(0.004)				
3	−6.92	0.008	**−3.01**	**0.001**	**−6.40**	**0.003**
	(5.81)	(0.011)				
12	−7.44	0.033	**−3.04**	**0.003**	**−5.62**	**0.010**
	(5.67)	(0.039)				
36	−5.94	0.076	**−2.94**	**0.009**	**−4.69**	**0.024**
	(4.68)	(0.092)				
60	−5.63	0.131	**−2.94**	**0.015**	**−4.40**	**0.036**
	(4.16)	(0.111)				
120	−5.48	0.241	**−2.70**	**0.026**	**−3.86**	**0.060**
	(3.18)	(0.131)				

Table 8 notes: Asset price data are generally monthly, January 1958- December 2013. $xret_t$ is the monthly log excess return on the S&P 500 index $pvar_t$ is the conditional variance of log equity returns under the physical measure, $qvar_t$ is the conditional variance of log returns under the risk-neutral measure. Bootstrapped standard errors are shown in parentheses. Panel B reports results from regressions of the form

$$\frac{1}{K}\sum_{k=1}^{K} xret_{t+k} = a + b \cdot \ln(PD_t) + u_{t,t+k}$$

where $\ln(PD_t)$ is the log price-dividend ratio. Bootstrapped standard errors are shown in parentheses.

Figure 1: Consumption growth conditional moments

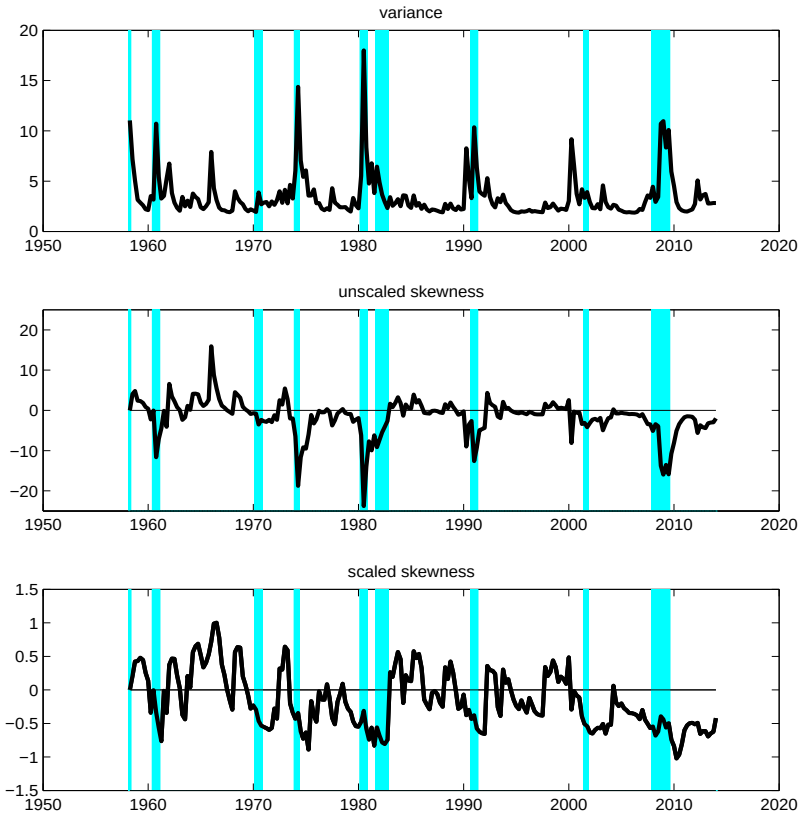

Figure 1 notes: Data are quarterly, real, per-capita consumption growth from 1958-2013. Data are demeaned prewhitened to remove MA(1) variation in the conditional mean. For the resulting series, ε_t, in the top panel we present the conditional variance, h_t:

$$\begin{aligned} \varepsilon_{t+1} &\sim N(0, h_t) \\ h_t &= h_0 + \alpha_0 h_{t-1}^2 + \alpha_1 \varepsilon_t^2 + \alpha_2 I_t \varepsilon_t^2 \\ I_t &= \begin{cases} 1 & if\ \varepsilon_t < 0 \\ 0 & if\ \varepsilon_t \geq 0 \end{cases} \end{aligned}$$

In the middle panel, we show the conditional skewness estimated using the following model

$$\begin{aligned} \nu_t &= E_t(\varepsilon_{t+1})^3 \\ \nu_t &= \nu_0 + \beta_0 \nu_{t-1} + \beta_1 \varepsilon_t^2 + \beta_2 I_t \varepsilon_t^2 \end{aligned}$$

where ν_t is the conditional third moment of the residuals. Shading denotes recession periods, as defined by the NBER.

Figure 2: Consumption growth quantile shifts conditional on $qvar_t$

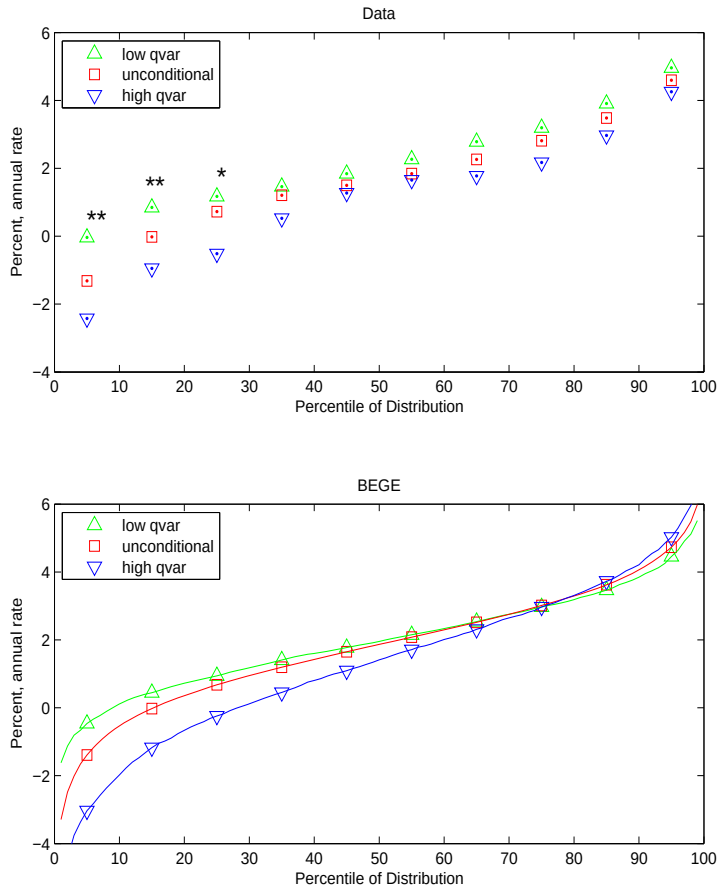

Figure 2 notes: Squares denote the unconditional distribution of quarterly consumption growth. Down-pointing triangles denote the distribution of consumption growth for quarters following a quarter in which a high level of the *qvar*, the risk neutral conditional variance of equity returns, was observed, defined as the *qvar* being in the upper decile of its sample distribution. Upward-point triangles denote the distribution of consumption growth conditional on *qvar* being in the lowest decile of is distribution. The upper panel plots results for U.S. quarterly consumption growth from 1958-2013. One and two asterisks denote that the difference between the upward- and downward-point triangles is significant at the 10 and 5 percent levels. The bottom panel plots results for data simulated under the BEGE model..

Figure 3: The BEGE Distribution

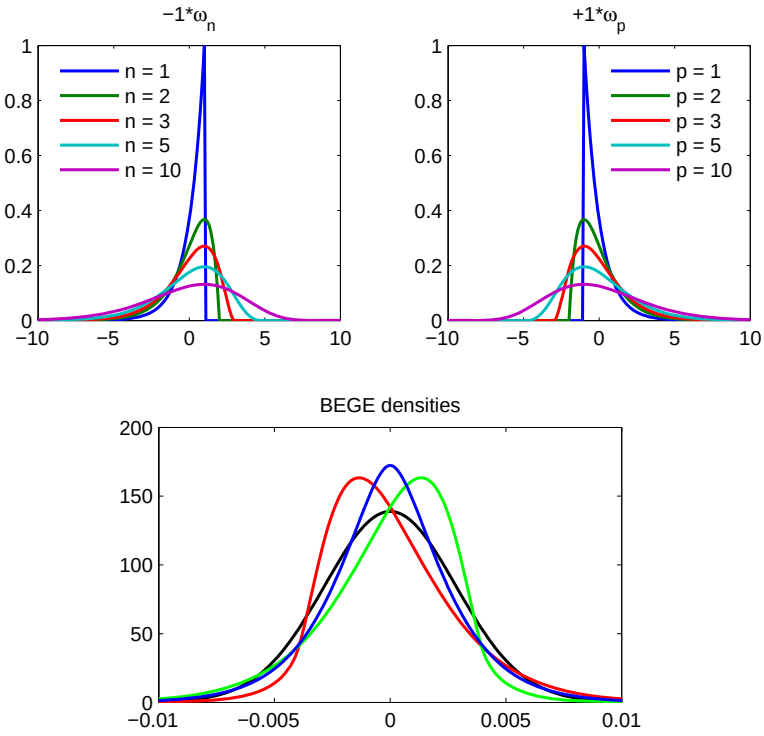

Figure 3 notes: The top panels display the centered gamma distribution for a range of values of the shape parameter and a unit scale parameter. The bottom panel displays four examples of the BEGE distribution. In the bottom panel, the black line plots the density for $[p_t, n_t] = [40, 40]$, the blue line, $[p_t, n_t] = [2, 2]$,the green line, $[p_t, n_t] = [0.4, 3]$, and the red line, $[p_t, n_t] = [3, 0.4]$. For all four lines, $\sigma_{cp} = \sigma_{cn}$ are both set to maintain a constant variance for all four distributions.

Figure 4: Pricing bad- and good-environment shocks

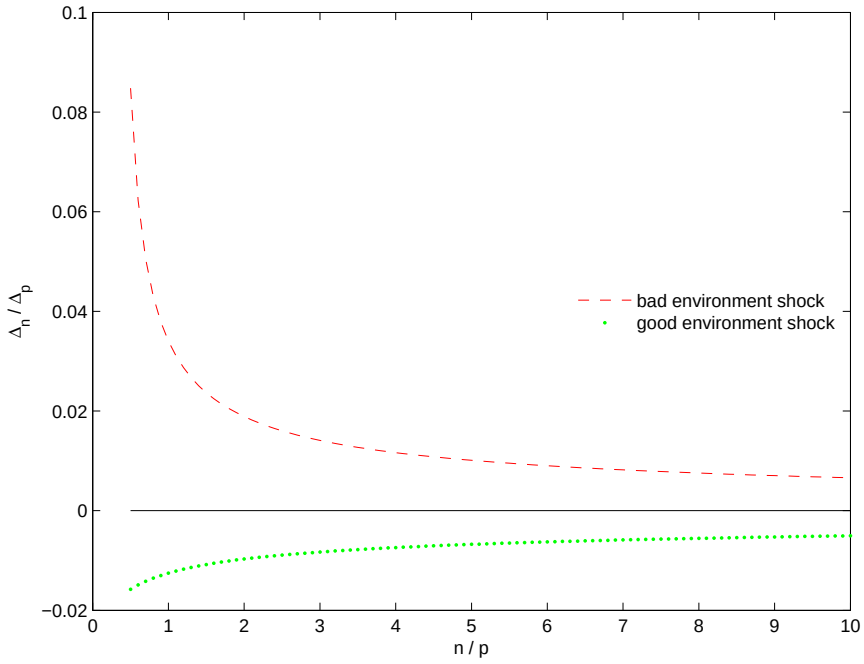

Figure 4 notes: We assume the agent's utility function has constant relative risk aversion over one-period ahead wealth, $\widetilde{W}$:

$$U\left(\widetilde{W}\right)=\frac{\widetilde{W}^{1-\gamma}-1}{1-\gamma}$$

where $\widetilde{W}$ is the stochastic value of one-period ahead wealth. We consider three different distributions for $\ln\left(\widetilde{W}\right)$. All three distributions have a common variance, σ_w^2, but they differ in their means and distributions of shocks:

distribution	mean	variance	skewness
Gaussian	μ_w	σ_w^2	0
$-1\cdot\omega_n$	$\mu_w+\Delta_n$	σ_w^2	$-2/\sqrt{n}$
$+1\cdot\omega_p$	$\mu_w+\Delta_p$	σ_w^2	$+2/\sqrt{p}$

We plot values for Δ_n and Δ_p such that the expected utility of all three distributions is the same, with $\gamma=5$, $\sigma_w^2=0.15^2$, where the latter is similar to the annual variance of U.S. equity returns in our sample. An exact formula for Δ_n and Δ_p is derived in the appendix..

Figure 5: Consumption growth tail probabilities

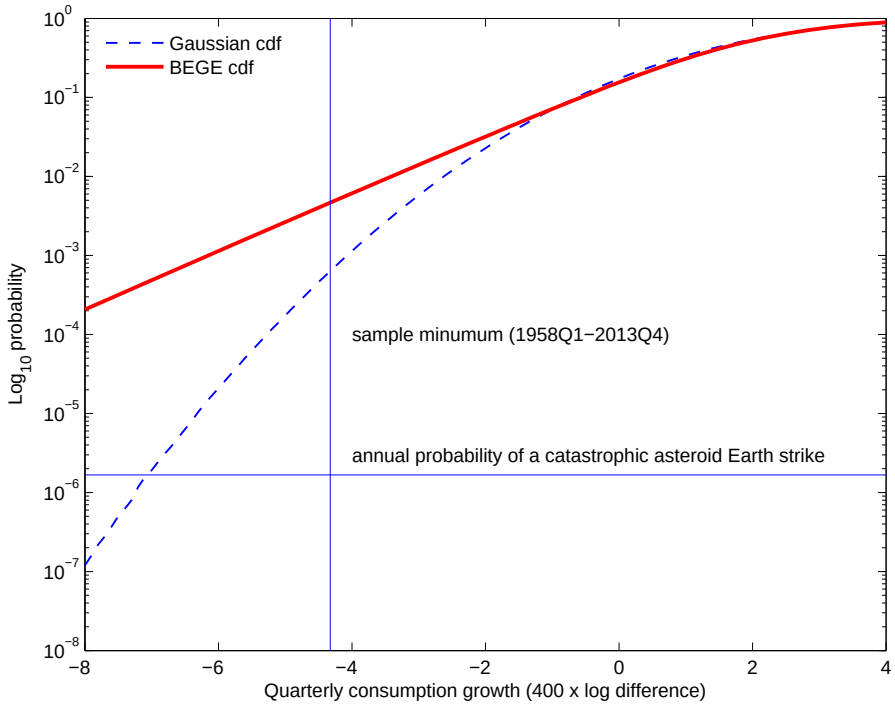

Figure 5 notes: The dashed line plots the cdf of a Gaussian distribution fit to the quarterly log growth rate of real per capita consumption on nondurables and services from 1958Q1 to 2013Q4. The mean and standard deviation of the series are 1.68 and 0.96 percent respectively and at an annual rate. The solid line plots a BEGE cdf fit to the same data using the model and point estimates described in Table 2. The horizontal line is placed at 2×10^{-6}, an estimate of the annual probability of a catastrophic asteroid Earth strike. The vertical line at -4.3 plots the minimum observed sample value of consumption growth (at an annual rate)...

Figure 6: Dependence of the sensitivity function on n and s

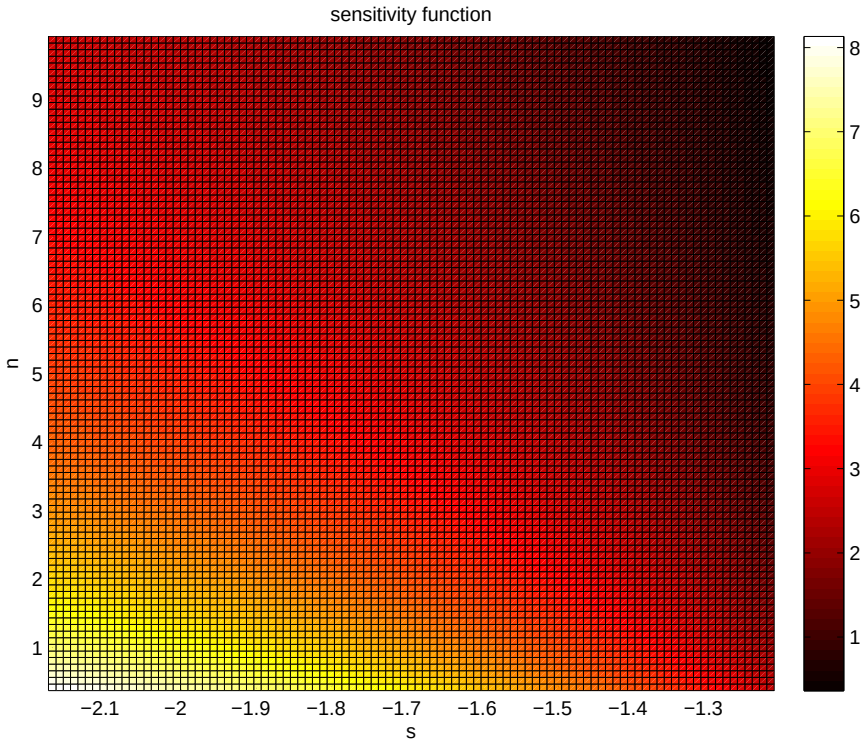

Figure 6 notes: The plot shows a heat map indicating the dependence of the sensitivity function, λ_t on s_t and n_t at the parameters values in Tables 2 and 4, (BEGE model).

Figure 7: Impulse response of asset prices to BEGE shocks

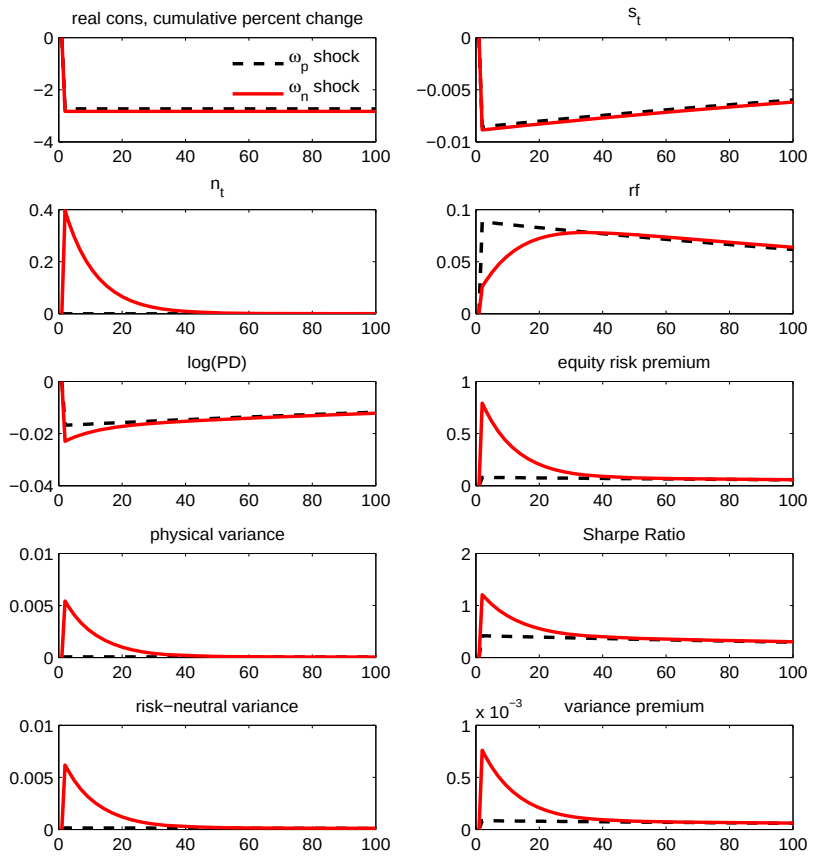

Figure 7 notes: These panels display the responses of various endogenous variables in the context of the BEGE model presented in Section 1. For each variable, the system is assumed to start at steady state so that $s_t = \overline{s}$ and $n_t = \overline{n}$. In period 0, the system is shocked by either a one standard deviation negative realization of $\omega_{p,t}$ or a one standard deviation positive realization of $\omega_{n,t}$. The units for all asset prices are percent, at an annual rate (squared, for variances).

Figure 8: Dependence of asset prices on n and s

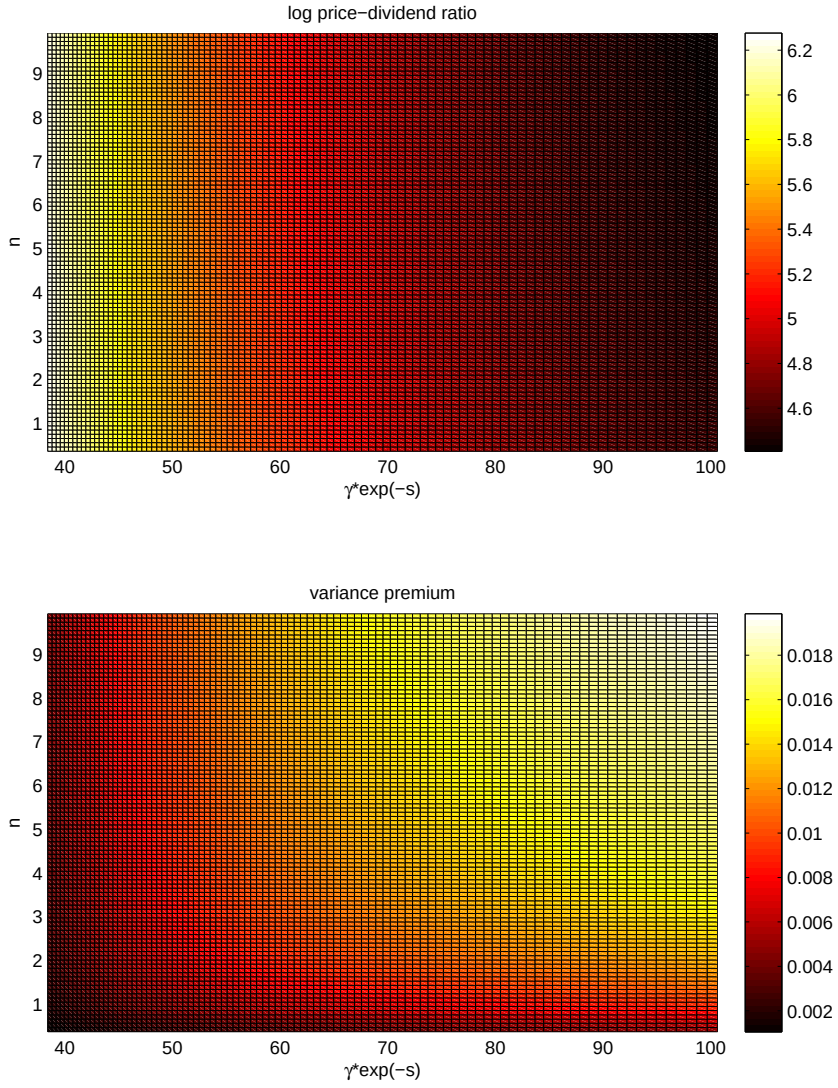

Figure 8 notes: The upper plot is a heat map showing the dependence of the log price-dividend ratio, $\ln(PD_t)$ on both n_t and local curvature, $-\gamma \exp(s_t)$. Dividends in the denominator are measured on a monthly basis, not trailing-twelve months. The units of the variance premium are percent annual rate, squared. The lower plot similarly depicts the dependence of the variance premium, $vprem_t = qvar_t - pvar_t$.