Asset Return Dynamics under Bad Environment-Good Environment Fundamentals

Geert Bekaert[a,*], Eric Engstrom[b‡]

[a] Columbia University and NBER; [b] Federal Reserve Board of Governors

First Draft: July 2009; This Draft: June 2010

Abstract

We introduce a "bad environment-good environment" technology for consumption growth in a consumption-based asset pricing model. Using the preference structure from Campbell and Cochrane (1999), the model generates realistic non-Gaussian features of fundamentals while still permitting closed-form solutions for asset prices. The model not only fits standard salient asset prices features including means and volatilities for equity returns and risk free rates, but also generates realistic features of the "risk-neutral" conditional density of equity returns, including the variance premium.

Keywords: Equity premium, variance premium, Countercyclical risk aversion, Economic Uncertainty, Dividend yield, Return predictability

JEL classification: G12, G15, E44

*Corresponding author:gb241@columbia.edu

†We appreciate comments from seminar participants at the 2009 CEF conference the University of Technology, Sydney, Australia, the University of Maryland (Finance), the University of Texas (Dallas), the University of Amsterdam, the London School of Economics and the London Business School. All errors are the sole responsibility of the authors.

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

1. Introduction

To date, the consumption based asset pricing literature has mostly focused on matching unconditional features of asset returns: the equity premium, the low risk free rate, and the variability of equity returns and dividend yields. In terms of conditional dynamics, a great deal of attention has been paid to time variation in the expected excess return on equities. A number of models have emerged that can claim some empirical success along these dimensions. Campbell and Cochrane (1999, CC henceforth) develop an external habit framework where time-varying risk aversion is the essential driver of asset return dynamics. CC keep the exogenous technology for consumption growth deliberately simple and linear. Bansal and Yaron (2004, BY henceforth), while working with different preferences due to Epstein and Zin (1989), generate realistic asset pricing dynamics by introducing "long-run risk" and time-varying uncertainty in the consumption growth process. Another recent strand of the literature that also focuses on the technology rather than preferences has rekindled the old Rietz (1990) idea that fear of a large catastrophic event may induce a large equity premium (see Barro (2006)). It is important to realize that in such a framework, there is no time variation in risk premiums unless the probability of the "crash" is assumed to vary through time (see Gourio (2010), and Wachter (2009)).

At the same time, a voluminous literature has focused on explaining the volatility dynamics of stock returns and the joint distribution of stock returns and option prices [see Chernov, Gallant, Ghysels and Tauchen (2003)]. This literature is largely reduced-form in nature, assuming stochastic processes for stock return dynamics and then testing how well such dynamics fit the data on both stock returns and option prices. Seminal articles in this vein include Chernov and Ghysels (2000) and Pan (2002). The current state-of-the art models are very complex, featuring stochastic volatility and jumps in both prices and volatility (see, for instance, Broadie, Chernov and Johannes (2007)).

From one perspective, the distinct development of these two literatures in dynamic asset pricing is surprising. Successfully modeling volatility and option price dynamics from a more structural perspective would appear not only economically important, but also statistically very informative. The empirical evidence on volatility dynamics is very strong, and many features of the data are without controversy, which is very different from the large uncertainty surrounding the evidence on return predictability (see e.g. Ang and Bekaert (2007), Goyal and Welch (2008) and Campbell and Thompson (2008)). From another perspective, however, this dichotomy is not surprising at all: every single consumption-based model described above would surely fail to generate anything like the volatility and option price dynamics observed in the data. A particularly powerful empirical feature of the data is the so-called variance premium, which is the difference between the "risk neutral" expected conditional variance of the stock market index and the actual

expected variance under the physical probability measure. The CBOE's VIX index essentially provides direct readings on the risk-neutral variance; see Bollerslev, Gibson and Zhou (2008) and Carr and Wu (2008) for more details. Not only does the VIX show considerable time variation, Bollerslev, Tauchen and Zhou (2009) show that the variance premium is a good predictor of stock returns. Other stylized facts about the risk neutral conditional distribution of returns include time-varying (but generally negative) skewness, and time-varying fat tails (see, for instance, Figlewski (2009)).

To generate these features of the risk-neutral distribution, structural models must endogenously generate time-varying non-Gaussianity in returns. However, most existing models would fail to do so, as the technology for fundamentals is too close to normality, and the models therefore generate near-Gaussian asset return dynamics.

We set out to integrate the two literatures by proposing a simple, tractable consumption based asset pricing model, where preferences are as in Campbell and Cochrane (1999), but the consumption technology is non-linear, following what we call a "Bad Environment – Good Environment" framework, "BEGE" for short. We essentially assume that the consumption growth process receives two types of shocks, both drawn from potentially fat-tailed, skewed distributions. While one shock has positive skewness, the other shock has negative skewness. Because the relative importance of these shocks varies through time, there are "good times" where the good distribution dominates, and "bad times" where the bad distribution dominates. An implication of the framework is that even during bad times, large good shocks can occur persistently, and vice versa. Such behavior has been very apparent in stock return dynamics during the 2007-2009 crisis. Economically, the BEGE model creates a riskier consumption growth environment, which, in equilibrium, leads to a large equity premium and substantial precautionary savings demands, keeping risk free rates low. Because the riskiness varies through time, the model generates intricate return dynamics, and realistic variance risk premiums. Crucially, we demonstrate that fundamentals indeed exhibit the kind of non-linearities that are generated by the BEGE framework.

The BEGE framework is reminiscent of regime –switching models, where a Markov variable generates switches between two normally distributed regimes. In principle, such mixture models can also generate time-varying skewness and kurtosis. The impact of such models in consumption based asset pricing was explored by Whitelaw (2000), Kandel and Stambaugh (1990), Bonomo and Garcia (1994), Epstein and Zin (2001) and Cecchetti, Lam and Mark (1990). Regime switching models have much of the same economic appeal as the model we propose, but unfortunately, they are fairly intractable in an equilibrium pricing context. In contrast, we use the gamma distribution for our shocks resulting in an affine term structure and quasi-closed form expressions for equity prices and the variance premium. This greatly increases the appeal of the framework as we can obtain useful intuition on what drives asset prices, and can easily estimate the

structural parameters. We formally test the performance of a simple version of our modeling framework with respect to a large number of empirical features of asset returns and fundamentals.

The remainder of the article is organized as follows. Section 2 introduces the model. We present simple solutions for the risk free rate, price dividend ratios and the variance premium. Section 3 introduces the data and documents that there are indeed time-varying non-linearities in the consumption growth process. We also set out the estimation strategy. Section 4 discusses our parameter estimates and the fit of the model. Apart from most salient asset price features, the model also fits the variance premium and other stylized facts about option prices. In addition, we show that the BEGE framework fits the consumption data better than the present versions of some alternative models that also introduce nonlinearities in fundamentals to fit asset pricing puzzles. These models include the volatility-in-volatility model of Bollerslev, Tauchen and Zhou (2009), the jumps and long-run risk model of Drechsler and Yaron (2009) and the time-varying disaster models of Nakamura, Steinsson, Barro and Ursua (2009), and Gabaix (2010). The final section offers some concluding remarks.

2. The Bad Environment-Good Environment (BEGE) Model

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

2.1. Fundamentals

Our model for consumption is given by the following equation:

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

where $\Delta c_t = \ln(C_t) - \ln(C_{t-1})$ is the logarithmic change in consumption. $\overline{g}$ is the unconditional mean rate of consumption growth, x_t is the deviation of the conditional growth rate from $\overline{g}$, so that the conditional mean of consumption growth is $E_t[\Delta c_{t+1}] = \overline{g} + x_t$. The final two terms reflect non-Gaussian innovations. The parameters σ_{cp} and σ_{cn} are both positive. The shocks, $\omega_{p,t+1}$ and $\omega_{n,t+1}$, are zero-mean innovations with the following distributions,

$$\begin{aligned} \omega_{p,t+1} &= ge_{t+1} - p_t \\ \omega_{n,t+1} &= be_{t+1} - n_t \end{aligned} \tag{2}$$

Above, ge_{t+1} represents the "good environment" variable and be_{t+1} represents the "bad environment" variable. Both follow gamma distributions. Specifically, $ge_{t+1} \sim \Gamma(p_t, 1)$ where $\Gamma(p_t, 1)$ represents a gamma

distribution with shape parameter, p_t, and size parameter equal to 1. Analogously, $be_{t+1} \sim \Gamma(n_t, 1)$. The shape parameters, p_t and n_t vary through time according to a stochastic process to be introduced shortly. These processes thus govern the conditional higher-order moments of Δc_t. Specifically, p_t governs the width of the positive tail, and n_t governs the width of the negative tail. Because the mean of the gamma distribution is equal to its shape parameter (when the size parameter is 1), the terms, $-p_t$ and $-n_t$ in Equation (2) ensure that the shocks each have conditional mean 0.

Similar to Bansal and Yaron (2004), we assume an AR(1) process for x_t:

$$x_t = \rho_x x_{t-1} + \sigma_{xp}\omega_{p,t} + \sigma_{xn}\omega_{n,t}$$

where ρ_x is the autocorrelation of x_t, and σ_{xp} and σ_{xn} determine the exposure of x_t to the consumption shocks.

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

$$\begin{aligned} MGF_m(\Delta c_{t+1}) &\equiv E_t[\exp(m\Delta c_{t+1})] \\ &= \exp\begin{pmatrix} m(\overline{g} + x_t) \\ -p_t(m\sigma_{cp} + \ln(1 - m\sigma_{cp})) \\ -n_t(-m\sigma_{cn} + \ln(1 + m\sigma_{cn})) \end{pmatrix} \end{aligned} \tag{3}$$

This follows directly from the MGF of the gamma distribution and the fact that $\omega_{p,t+1}$ and $\omega_{n,t+1}$ are independent.[1] Next, we solve for the first few conditional centered moments of Δc_{t+1} by evaluating subsequent derivatives of the MGF at $m = 0$, which provides uncentered moments, and then translating to their centered counterparts in the usual way. This yields:

$$\begin{aligned} E_t\left[(\Delta c_{t+1} - (\overline{g} + x_t))^2\right] &= \sigma_{cp}^2 p_t + \sigma_{cn}^2 n_t \equiv vc_t \\ E_t\left[(\Delta c_{t+1} - (\overline{g} + x_t))^3\right] &= 2\sigma_{cp}^3 p_t - 2\sigma_{cn}^3 n_t \equiv sc_t \\ E_t\left[(\Delta c_{t+1} - (\overline{g} + x_t))^4\right] - 3vc_t^2 &= 6\sigma_{cp}^4 p_t + 6\sigma_{cn}^4 n_t \equiv kc_t \end{aligned} \tag{4}$$

The top line of Equation (4) shows that both p_t and n_t contribute positively to the conditional variance of consumption, defined as vc_t. They differ, however, in their implications for the conditional skewness of consumption. As can be seen in the expression for the centered third moment, sc_t, skewness, which is defined as $sc_t/vc_t^{3/2}$, is positive when p_t is relatively large, and negative when n_t is large. This is the essence

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

of the BEGE model: the bad environment refers to an environment in which the $\omega_{n,t}$ shocks dominate; in the good environment the $\omega_{p,t}$ shocks dominate. Of course, in both environments shocks are zero on average, but there is a higher probability of large positive shocks in a "good environment" and vice versa. Whether good or bad shocks dominate depends on the relative values of p_t and n_t, and the sensitivity of consumption growth to both shocks. Finally, the third line of the equation is the excess centered fourth moment, kc_t. The conditional excess kurtosis of consumption growth is given by kc_t/vc_t^2. Both p_t and n_t contribute positively to this moment, though in different proportions than they do for vc_t. Note that there is a linear dependence among higher moments of Δc_t, all of which are linear in p_t and n_t.[2]

Figure 1 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 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 red 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 blue 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 makes tangible the role of p_t as the "good environment" variable and n_t as "the bad environment" variable.

We now turn to the assumed dynamics for p_t and n_t. We model the p_t factor as following a simple, autoregressive process with "square-root-like" volatility dynamics,

$$p_t = \overline{p} + \rho_p \left(p_{t-1} - \overline{p}\right) + \sigma_{pp}\omega_{p,t} \tag{5}$$

where $\overline{p}$ is the unconditional mean of p_t, ρ_p is its autocorrelation coefficient, and σ_{pp} governs the conditional volatility of the process. Specifically, the conditional volatility of p_{t+1} is $\sigma_{pp}\sqrt{p_t}$ since the variance of $\omega_{p,t+1}$ is p_t. With fine enough time increments, this ensures that 0 is a reflecting boundary for the process. We model n_t symmetrically,

$$n_t = \overline{n} + \rho_n \left(n_{t-1} - \overline{n}\right) + \sigma_{nn}\omega_{n,t}. \tag{6}$$

Note that the conditional covariances between Δc_{t+1} and p_{t+1} and n_{t+1} are, respectively,

$$\begin{aligned} COV_t\left[\Delta c_{t+1}, p_{t+1}\right] &= \sigma_{cp}\sigma_{pp}p_t \\ COV_t\left[\Delta c_{t+1}, n_{t+1}\right] &= -\sigma_{cn}\sigma_{nn}n_t \end{aligned} \tag{7}$$

[2] While we have represented the BEGE distribution as a combination of two independent shocks for illustrative purposes, it can, of course, also be represented as a univariate distribution with a density function that depends on four parameters, p_t, n_t, σ_{cp}, and σ_{cn}. A closed-form (albeit messy) analytic solution for the BEGE density function is also available.

so that we have hard-wired a positive conditional correlation between Δc_{t+1} and p_{t+1}, and a negative conditional covariance between Δc_{t+1} and n_{t+1}. This assumes that positive shocks to consumption tend to increase the variability of "good" shocks while negative consumption shocks are associated with a greater negative tail. Despite this assumption, the conditional covariance of Δc_{t+1} and its own conditional variance, vc_t, is:

$$COV_t\left[\Delta c_{t+1}, vc_{t+1}\right] = \sigma_{cp}^3 \sigma_{pp} p_t - \sigma_{cn}^3 \sigma_{nn} n_t, \tag{8}$$

which can take on either sign and, indeed, generally varies through time.

2.2. Preferences

Consider a complete markets economy as in Lucas (1978), but modify the preferences of the representative agent to have the form:

$$E_0\left[\sum_{t=0}^{\infty} \beta^t \frac{\left(C_t - H_t\right)^{1-\gamma} - 1}{1-\gamma}\right], \tag{9}$$

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

One motivation for an external habit stock is the "keeping up with the Joneses" framework of Abel (1990, 1999). There, individual investors evaluate their own consumption relative to a benchmark representing past or current aggregate consumption, H_t. In Campbell and Cochrane (1999), H_t is an exogenously modelled subsistence or habit level. Hence, the local coefficient of relative risk aversion equals $\gamma \frac{C_t}{C_t - H_t}$, where $\left(\frac{C_t - H_t}{C_t}\right)$ is defined as the surplus ratio[3]. As the surplus ratio goes to zero, the consumer's risk aversion goes to infinity. In our model, we define the inverse of the surplus ratio, Q_t, so that. $\gamma \cdot Q_t$ $(Q_t > 1)$ represents stochastic risk aversion. As Q_t changes over time, the representative investor's risk tolerance changes.

The 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:

$$\begin{aligned} M_{t+1} &= \beta \frac{\left(C_{t+1}/C_t\right)^{-\gamma}}{\left(Q_{t+1}/Q_t\right)^{-\gamma}} \\ &= \beta \exp\left[-\gamma \Delta c_{t+1} + \gamma\left(q_{t+1} - q_t\right)\right], \end{aligned} \tag{10}$$

where $q_t = \ln(Q_t)$.

This model may better explain the return predictability evidence than the standard model with power utility because it can generate counter-cyclical expected returns and prices of risk. We specify the process

[3] Of course, this is not actual risk aversion defined over wealth, which depends on the value function. The Appendix to Campbell and Cochrane (1995) examines the relation between "local" curvature and actual risk aversion, which depends on the sensitivity of consumption to wealth. In their model, actual risk aversion is simply a scalar multiple of local curvature. In the present article, we only refer to the local curvature concept, and slightly abuse terminology in calling it "risk aversion."

for $q_t \equiv \ln(Q_t)$ directly as follows:

$$q_{t+1} = \mu_q + \rho_q q_t + \sigma_{qp}\omega_{p,t+1} + \sigma_{qn}\omega_{n,t+1} \tag{11}$$

where μ_q, ρ_q and σ_{qp} and σ_{qn} are scalar parameters. As in CC, the risk aversion process is persistent, governed by the parameter ρ_q, and heteroskedastic, governed by time-variation in p_t and n_t. We also follow CC in having the innovation in q_t entirely spanned by the consumption shocks, but there are two such shocks in our framework and these shocks are heteroskedastic.[4] The conditional covariance between risk aversion and consumption is given by:

$$COV_t\left[\Delta c_{t+1}, q_{t+1}\right] = \left(\sigma_{cp}\sigma_{qp}\right)p_t - \left(\sigma_{cn}\sigma_{qn}\right)n_t. \tag{12}$$

The external habit interpretation of the model requires this covariance to be negative: positive consumption shocks decrease risk aversion. In CC, this correlation is a non-linear process increasing in q_t. Our modeling here is different and more flexible. We would expect σ_{qp} to be negative and σ_{qn} to be positive. When that occurs, shocks that increase the relative importance of "good environment" shocks ($\omega_{p,t}$) decrease risk aversion, and shocks that increase the relative importance of "bad environment" shocks" ($\omega_{n,t}$) increase risk aversion. Moreover, the conditional covariance between consumption growth and risk aversion is then always negative. We will not, however, impose this restriction in the estimation stage.

2.3. Asset prices

In this subsection, we present solutions for asset prices in the BEGE framework.

2.3.1. The risk free short rate

To solve for the real risk free short rate, rrf_t, we use the usual no-arbitrage condition,

$$\exp\left(rrf_t\right) = E_t\left[\exp\left(m_{t+1}\right)\right].^{-1} \tag{13}$$

To simplify this expectation, it will be convenient to define the quantities,

$$\begin{aligned} a_p &= \gamma\left(\sigma_{qp} - \sigma_{cp}\right) \\ a_n &= \gamma\left(\sigma_{qn} + \sigma_{cn}\right) \end{aligned} \tag{14}$$

These quantities measure the impact of the two sources of uncertainty on the pricing kernel, as can be seen in the equation,

$$m_{t+1} - E_t\left[m_{t+1}\right] = a_p\omega_{p,t+1} + a_n\omega_{n,t+1} \tag{15}$$

[4] In this sense, our modeling differs from Bekaert, Engstrom and Grenadier (2005) and Bekaert, Engstrom and Xing (2009) who let q_t depend on a shock not spanned by fundamental shocks.

For ease of interpretation, we focus on the case where $a_p < 0$ and $a_n > 0$. This corresponds to a situation where positive $\omega_{p,t+1}$ shocks decrease marginal utility (good news) while positive $\omega_{n,t+1}$ shocks increase marginal utility (bad news). Using Lemma 1 in the appendix, the real short rate can be expressed as,

$$rrf_t = \begin{pmatrix} -\ln\beta + \gamma\left(\overline{g} + x_t\right) + \gamma\left(1-\rho_q\right)\left(q_t - \overline{q}\right) \\ + f\left(a_p\right)p_t + f\left(a_n\right)n_t \end{pmatrix} \tag{16}$$

where the function, $f(x)$, is defined by

$$f(x) = x + \ln(1-x) \tag{17}$$

The first line in the solution for rrf_t has the usual consumption and utility smoothing effects: to the extent that marginal utility is expected to be lower in the future (that is, when $(\overline{g} + x_t) > 0$ and/or, $q_t > \overline{q}$), investors desire to borrow to smooth marginal utility, and so risk free rates must rise. The bottom line captures precautionary savings effects, that is, the desire of investors to save more in uncertain times. Notice that because the function $f(x)$ is always negative[5], the precautionary savings effects are also always negative. A third-order Taylor expansion of the log function helps with the interpretation of rrf_t:

$$rrf_t \approx \begin{pmatrix} -\ln\beta + \gamma\left(\overline{g} + x_t\right) + \gamma\left(1-\rho_q\right)\left(q_t - \overline{q}\right) \\ + \left(-\frac{1}{2}a_p^2 - \frac{1}{3}a_p^3\right)p_t \\ + \left(-\frac{1}{2}a_n^2 - \frac{1}{3}a_n^3\right)n_t \end{pmatrix} \tag{18}$$

The first precautionary savings terms,$-\frac{1}{2}a_p^2 p_t$ and $-\frac{1}{2}a_n^2 n_t$ capture the usual precautionary savings effects: higher volatility generally leads to increased savings demand, depressing interest rates. The cubic terms represent a novel feature of the BEGE model. Consider again the case where $a_p < 0$ and $a_n > 0$. Under this assumption the term, $-\frac{1}{3}a_p^3 p_t > 0$, mitigates the precautionary savings effect to the extent that the good-environment variable, p_t, is large. This makes perfect economic sense. When good environment shocks dominate, the probability of large positive shocks is relatively large, and the probability of large negative shocks is small, decreasing precautionary demand. Conversely, the $-\frac{1}{3}a_n^3 n_t < 0$ term indicates that precautionary savings demands are exacerbated with 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. Through this mechanism, our model may generate the kind of extremely low but also very volatile risk free rates witnessed in the 2007-2009 crisis period.

2.3.2. Equity valuation

Following Campbell and Cochrane (1999), we assume that dividends equal consumption and solve for equity prices as a claim to the consumption stream. In any present value model, under a no-bubble

[5] We also require $x < 1$, a weak technical condition that is always met in our estimations.

transversality condition, the equity price-dividend ratio (the inverse of the dividend yield) is represented by the conditional expectation,

$$\frac{P_t}{D_t} = E_t\left[\sum_{i=1}^{\infty}\exp\left(\sum_{j=1}^{i}\left(m_{t+j}+\Delta d_{t+j}\right)\right)\right] \tag{19}$$

where $\frac{P_t}{D_t}$ is the equity price-dividend ratio and Δd_t represents logarithmic dividend growth. This conditional expectation can also be solved in our framework as an exponential-affine function of the state vector, as is summarized in the following proposition, which is proved in the appendix.

Proposition1

For the economy described by Equations (1) through (11), the price-dividend ratio of equity is given by

$$\frac{P_t}{D_t} = \sum_{i=1}^{\infty}\exp\left(\widetilde{A}_i + \widetilde{B}_i p_t + \widetilde{C}_i n_t + \widetilde{D}_i q_t + \widetilde{E}_i x_t\right) \tag{20}$$

where the initial values of the parameter sequences are given by

$$\begin{aligned}
\widetilde{A}_1 &= \ln\beta + (1-\gamma)\,\overline{g} + \gamma\left(1-\rho_q\right)\overline{q}\\
\widetilde{B}_1 &= -f\left(a_p + \sigma_{cp}\right)\\
\widetilde{C}_1 &= -f\left(a_n - \sigma_{cn}\right)\\
\widetilde{D}_1 &= -\gamma\left(1-\rho_q\right)\\
\widetilde{E}_1 &= 1-\gamma
\end{aligned}$$

where the functions providing the coefficients for $n \geq 2$ are represented by

$$\begin{aligned}
\widetilde{A}_i &= \widetilde{A}_1 + \widetilde{A}_{i-1} + \widetilde{B}_{i-1}\mu_p + \widetilde{C}_{i-1}\mu_c + \widetilde{D}_{i-1}\mu_q\\
\widetilde{B}_i &\equiv \widetilde{B}_{i-1}\rho_p - f\left(a_p + \sigma_{cp} + \widetilde{B}_{i-1}\sigma_{pp} + \widetilde{D}_{i-1}\sigma_{qp} + \widetilde{E}_{i-1}\sigma_{xp}\right)\\
\widetilde{C}_i &\equiv \widetilde{C}_{i-1}\rho_n - f\left(a_n - \sigma_{cn} + \widetilde{C}_{i-1}\sigma_{nn} + \widetilde{D}_{i-1}\sigma_{qn} + \widetilde{E}_{i-1}\sigma_{xn}\right)\\
\widetilde{D_i} &\equiv \widetilde{D}_{i-1}\rho_q - \gamma\left(1-\rho_q\right)\\
\widetilde{E_i} &\equiv \widetilde{E}_{i-1}\rho_x - \gamma
\end{aligned}$$

First, note that $\widetilde{B}_1$ and $\widetilde{C}_1$ are always positive because the function $-f(x)$ is always positive. Moreover, one can easily check that $\widetilde{B}_i$ and $\widetilde{C}_i$ are positive for all i as well. In other words, positive shocks to n_t and p_t drive up the price-dividend ratio. This is because p_t and n_t increase the volatility of the pricing kernel, inducing precautionary savings demands, which increases the current price of future cash flows, all else equal. As we will see below, however, increases in p_t and n_t also raise the equity premium, which serves to depress equity prices relative to safe assets.[6]

[6] There is a large literature examining the effects of uncertainty on equity prices. The folklore wisdom is that increased economic uncertainty ought to depress stock prices because it raises the equity premium (see Poterba and Summers (1986) and Wu (2001)). However, such a conclusion is by no means general. Pastor and Veronesi (2006) stress that uncertainty about cash flows should increase stock values (as it makes the distribution of future cash flows positively skewed), whereas Abel (1988) 's Lucas –tree model can generate either effect, depending on the coefficient of relative risk aversion. In Barsky (1989) and Bekaert, Engstrom, and Xing (2009), similar to this paper, the term structure effects of increased uncertainty cause equity prices to (potentially) rise.

Next, $\widetilde{D}_i$ term captures the effect of the risk aversion variable, q_t, which affects equity price-dividend ratios through utility smoothing channels; increases in q_t tend to depress equity prices as investors' desire to save diminishes Note that there is no marginal pricing difference in the effect of q_t on a riskless versus risky coupon stream. This is true by construction in this model because the preference variable, q_t, affects neither the conditional mean nor volatility of cash flow growth, nor the conditional covariance between the cash flow stream and the pricing kernel at any horizon. We purposefully excluded such relationships because, economically, it does not seem reasonable for investor preferences to affect productivity.

Finally, the effect of x_t on equity is represented by $\widetilde{E_i}$ Assuming $\gamma > 1$, $\widetilde{E_i}$ is negative for all i, so that increases in x_t lower equity prices. While an increase in x_t raises expected dividend growth one-for-one, suggesting higher equity prices, this effect is more than offset because higher x_t simultaneously increases the expected growth in marginal utility (and thus interest rates), and by a larger factor, γ.

2.3.3. Approximations to the exact equity solution

While the above solution for the equity price-dividend ratio is exact, it is a non-linear function of the state vector. To simplify our subsequent calculations, it is useful to calculate a log-linear approximation to the price-dividend ratio. It is shown in the appendix that the logarithmic dividend-price ratio, dp_t, is approximately,

$$dp_t \approx d_0 + d_1' Y_t \tag{21}$$

where $Y_t = [p_t, n_t, \Delta c_t, q_t, x_t]'$ is the state vector and the coefficients d_0, d_1, etc. are functions of the deep model parameters with explicit formulae provided in the appendix. In light of the discussion above, we expect the following signs for dependence of dp_t on Y_t (for $\gamma > 1$):

d_1^p	d_1^n	$d_1^{\Delta c}$	d_1^q	d_1^x
$(-)$	$(-)$	(0)	$(+)$	$(+)$

For a more tractable linear expression for logarithmic returns, ret_t, we first note that ret_t can be expressed as

$$ret_{t+1} = dp_t + \Delta d_{t+1} + \ln\left(1 + \frac{P_{t+1}}{D_{t+1}}\right).$$

Using a second linearization of the final term, we can approximate equity returns as

$$ret_{t+1} \approx r_0 + r_1' Y_{t+1} + r_2' Y_t \tag{22}$$

Because the dependencies of $\ln\left(1+\frac{P_{t+1}}{D_{t+1}}\right)$ on the state vector have the same sign as those of P_{t+1}/D_{t+1}, it follows that returns load onto the contemporaneous shocks to elements of Y_{t+1} with the following signs:

r_1^p	r_1^n	$r_1^{\Delta c}$	r_1^q	r_1^x
$(+)$	$(+)$	$(+)$	$(-)$	$(-)$

2.3.4. The distribution of equity returns

We now examine the implications of the BEGE model for the conditional distribution of equity returns. We examine the physical and risk-neutral distributions separately.

The equity risk premium. As is well-known, the standard no-arbitrage condition,

$1 = E_t\left[\exp\left(m_{t+1} + ret_{t+1}\right)\right]$, leads to the following expression for the (gross) equity risk premium,

$$\frac{E_t\left[\exp\left(ret_{t+1}\right)\right]}{\exp\left(rrf_t\right)} = 1 - COV_t\left[\exp\left(m_{t+1}\right), \exp\left(ret_{t+1}\right)\right]$$

Under the BEGE model, using the linear approximation for ret_t and Lemmas 1 and 2 in the appendix, this expression simplifies to

$$\frac{E_t\left[\exp\left(ret_{t+1}\right)\right]}{\exp\left(rrf_t\right)} = \exp\left(\kappa^p p_t + \kappa^n n_t\right).$$

The equity premium only depends on the two factors that affect the moments of the pricing kernel and its covariance with returns. Whether the premium increases or decreases in p_t and n_t depends on the signs of κ^p and κ^n respectively, which, in turn depend on the deep model parameters. To gain some intuition, let's first look at κ^p. Define

$$\overline{r}^p = r_1^p\sigma_{pp} + r_1^c\sigma_{cp} + r_1^q\sigma_{qp} + r_1^x\sigma_{xp}$$

which measures the exposure of returns to p_t shocks. Then,

$$\kappa^p = \frac{\frac{\overline{r}^p}{1-a_p} - 1}{\overline{r}^p - 1}$$

Under our maintained assumption that $a_p < 0$ (that is, that positive shocks to p_t lower marginal utility), we can derive two useful facts. First, $\kappa^p < 0$ only if $1 < \overline{r}^p < 1 - a_p$. Because a_p is a relatively small number this is unlikely to happen. Second, κ^p is strictly increasing in $\overline{r}^p$. Hence, any term that increases $\overline{r}^p$ increases the dependence of the equity premium on p_t and its unconditional value. Given the expected signs of the r_1 coefficients derived above and the fact that we expect σ_{cp} to be positive and σ_{qp} to be negative, the equity premium is increasing in the equity return's exposures to shocks in p_t (variance risk), consumption growth,

and q_t (risk aversion). However, because positive shocks to x_t contribute to negative returns ($r_1^x < 0$, a term structure effect), a positive loading of x_t onto p_t shocks would decrease the equity premium. This makes sense, as the x_t shocks have an opposite effect on marginal utility as do p_t shocks. In sum, it seems likely for κ^p to be positive and for the equity premium to be increasing in p_t.

A similar expression is available for κ^n. Define,

$$\overline{r}^n = r_1^n \sigma_{nn} - r_1^c \sigma_{cn} + r_1^q \sigma_{qn} + r_1^x \sigma_{xn}$$

Then,

$$\kappa^n = \frac{\frac{\overline{r}^n}{1-a_n} - 1}{\overline{r}^n - 1}$$

Maintaining our assumption that $a_n > 0$ (positive n_t shocks raise marginal utility), we can now derive that $\kappa^n < 0$ only when $1 - a_n < \overline{r}^n < 1$, again a condition unlikely to be satisfied. Now, however, the expression for κ^n is *decreasing* in $\overline{r}^n$. Therefore any term that decreases $\overline{r}^n$ increases the equity premium. For example, the positive exposure of returns to consumption growth ($r_1^c > 0$) contributes positively to the equity risk premium through the n_t channel as well as $\sigma_{cn} > 0$. In contrast, because n_t shocks raise marginal utility, the positive dependence of returns on n_t shocks ($r_1^n > 0$) provides a hedge, and lowers the equity risk premium, all else equal. The expected negative exposure of returns to risk aversion ($r_1^q < 0$), together with the presumed positive exposure of risk aversion to n_t ($\sigma_{qn} > 0$) implies a positive contribution to the equity risk premium.

Higher-order physical moments. The appendix shows how to calculate the (physical) moment generating function for any affine function of the state vector. Armed with that, it is possible to calculate any moment of interest. These calculations are straightforward and similar to those for computing the conditional moments of consumption growth, as shown in Section 2.1. We begin by calculating the physical measure of conditional equity return volatility, $pvar_t$. For more compact notation in this subsection, we will continue to use use $\overline{r}^p$ and $\overline{r}^n$, which featured in the above discussion of the equity risk premium. Using the approximation in Equation (22) and Lemma 1 yields:

$$pvar_t = \left(\overline{r}^p\right)^2 p_t + \left(\overline{r}^n\right)^2 n_t \tag{23}$$

Not surprisingly, both p_t and n_t contribute to return variance in a positive, linear fashion. Similar calculations show that the conditional (centered) third moment and excess fourth moment, denoted psk_t and pku_t

respectively, can be expressed as:

$$\begin{aligned} psk_t &= 2\left(\overline{r}^p\right)^3 p_t - 2\left(\overline{r}^n\right)^3 n_t \\ pku_t &= 6\left(\overline{r}^p\right)^4 p_t + 6\left(\overline{r}^n\right)^4 n_t \end{aligned} \tag{24}$$

The BEGE model is therefore clearly able to generate time-varying skewness which can change sign over time as well as time-varying kurtosis. It is worth highlighting that because there are only two state variables driving these (and all higher) moments, there is a linear dependence among the moments' dynamics, which may be counterfactual.

Higher-order risk-neutral moments. Many stylized facts about the risk-neutral distributions of returns have emerged in the literature, see Figlewski (2009) for a good survey. We focus our analysis of the BEGE system on the following empirical regularities:

1. The risk-neutral conditional variance of returns usually exceeds the physical variance of returns; the difference is called the variance premium.[7]

2. The variance premium covaries positively with the equity risk premium. (See Bollerslev, Gibson and Zhou (2009), for instance.)

3. The risk-neutral distribution of equity returns is negatively skewed and fat tailed.[8].

To facilitate the calculation of the risk-neutral distribution of returns in the BEGE framework, let us first define the risk-neutral expectation of any variable, $E_t^Q\left[\exp\left(w_{t+1}\right)\right]$,as

$$E_t^Q\left[\exp\left(w_{t+1}\right)\right] = E_t\left[\exp\left(m_{t+1}+w_{t+1}\right)\right]\left(E_t\left[\exp\left(m_{t+1}\right)\right]\right)^{-1} \tag{25}$$

Based on this definition, Lemma 2 of the appendix shows how to calculate the risk-neutral moment generating function for the BEGE system, which renders the calculation of any risk-neutral moment straightforward, if tedious. For instance, the risk-neutral variance measure, $qvar_t$, simplifies to:

$$qvar_t = \left(\frac{\overline{r}^p}{1-a_p}\right)^2 p_t + \left(\frac{\overline{r}^n}{1-a_n}\right)^2 n_t \tag{26}$$

This expression is intuitive when compared with the solution for $pvar_t$, adding a simple denominator term to the parameters multiplying p_t and n_t in Equation (23). Consider first the denominator term

[7] In the options literature, researchers often reserve the term "variance premium" for the negative of what we call the variance premium, which is also the expected payoff to long position in a variance swap (see e.g. Carr and Wu (2008)), and may term our variable, a volatility spread (see e.g. Bakshi and Madan, 2006).

[8] This is consistent with the older options pricing literature that focused on implied volatility "smirks" and "smiles," using the Black-Scholes option pricing model to back out implied volatilities at various strike prices. See, for instance, Bakshi, Kapadia and Madan (2003).

multiplying p_t. Maintaining our assumption that $a_p < 0$, the denominator is strictly greater than 1. This implies that p_t, the good environment variable, serves to reduce risk neutral variance relative to its physical measure counterpart. On the other hand, as long as $a_n > 0$[9] (which is consistent with positive n_t shocks raising marginal utility), n_t will generally increases the risk-neutral variance relative to its physical measure counterpart. This is intuitive and suggests that the BEGE system is potentially capable of matching stylized fact 1: the so-called variance premium, $qvar_t - pvar_t$ (henceforth denoted $vprem_t$) is positive. The Appendix shows that the presence of the non-linear BEGE shocks are essential to generate a positive variance premium; with only Gaussian shocks, the variance premium is zero. This is reminiscent of a result in the jump model of Drechsler and Yaron (2009). Moreover, if, as expected, increases in n_t raise the equity risk premium, then the variance premium may covary positively with the equity risk premium, consistent with stylized fact 2. Finally, if the variance premium is indeed increasing in n_t, then the BEGE framework may exhibit the property that the variance premium is higher when the physical return distribution is more leptokutotic and/or more left-skewed, a feature emphasized by Bakshi and Madan (2006) as being consistent with a broad range of preference specifications and also having strong empirical support.

We now turn to higher risk-neutral moments. Simple calculations using Lemma 2 show that the risk neutral conditional (centered) third moment and excess fourth moment, qsk_t and qku_t respectively, can be expressed as:

$$\begin{aligned} qsk_t &= 2\left(\frac{\overline{r}^p}{1-a_p}\right)^3 p_t - 2\left(\frac{\overline{r}^n}{1-a_n}\right)^3 n_t \\ qku_t &= 6\left(\frac{\overline{r}^p}{1-a_p}\right)^4 p_t + 6\left(\frac{\overline{r}^n}{1-a_n}\right)^4 n_t \end{aligned} \tag{27}$$

Clearly, qsk_t will be negative when n_t is large and qku_t will be high to the extent that p_t or n_t are large. These effects make the BEGE system potentially consistent with stylized fact 3.

3. Empirical Implementation

In this section, we introduce the data used in the study and present reduced-form evidence for the kind of variation in consumption growth implied by our model in Section 1. We then outline the estimation strategy for the asset pricing model.

3.1. Asset Price Data

The asset pricing data sample is by necessity relatively short, spanning from January 1990 through December 2009, since it uses option prices. We estimate the real short rate, rrf_t, as the 30-day nominal T-

[9] We also need $a_n < 2$, a technical condition which is always met in our estimations.

bill yield provided by the Federal Reserve less expected quarter-ahead inflation (at a monthly rate) measured from the Blue Chip survey. In doing so, we implicitly assume that the inflation risk premium is zero at the monthly horizon and that the term structure of expected inflation is flat at horizons less than one quarter. For equity prices, we use the logarithmic dividend yield, dp_t, for the S&P 500, calculated as trailing 12-month dividends (divided by 12) divided by the month-end price. The real return to equity, ret_t, is the logarithmic change in the month-end level of the S&P 500 plus the monthly dividend yield defined above minus PCE inflation over the month. We calculate the risk-neutral conditional second, third and fourth moments of equity returns, $qvar_t$, $qskw_t$, and $qkur_t$ respectively, using the method of Bakshi, Kapadia and Madan (2003). This involves calculating the prices of portfolios of options designed to have payoffs that are determined by particular higher order moments of returns. We obtained a panel of option prices across the moneyness spectrum for the S&P 500 index from 1996-2009 from OptionMetrics, and from 1990-1995 from DeltaNeutral. We used the option contracts that have maturity closest to one month, and filtered out illiquid options according to the rules described in Figlewski (2009). Finally, we calculate the physical probability measure of equity return conditional variance, $pvar_t$, in two steps. We begin with the monthly realized variance, $rvar_t$, calculated as squared 5-minute capital appreciation returns over the month. Then we project $rvar_t$ onto one-month lags of the variables: $rvar_t$, and $qvar_t$.[10] The fitted values from this regression are used to measure $pvar_t$. This procedure is quite close to that used by Drechsler and Yaron (2009) and others.

Table 1 reports simple univariate sample statistics for these data. The standard errors reported in parentheses below the statistics are the standard deviations of 10,000 replications of a VAR bootstrap. Specifically, we estimate a first-order VAR on the data, from which we block-bootstrap the residuals using 12 months per block. We then use the VAR parameters to generate bootstrapped asset price data of the same length as our sample, for which we calculate the univariate sample statistics.

The annualized average risk free real short rate in our sample is 0.0009×12 or about 1.1 percent. Its volatility, at an annual rate, is $0.0012 \times \sqrt{12}$ or about 1.4 percent. Log dividend yields are quite variable and highly persistent. The equity premium is $0.0069 - 0.0009 = 0.0060$ or about 7 percent at an annual rate. The volatility of returns is about 15 percent annualized. All these statistics are similar to what other researchers have documented for this sample period.

The conditional variance of equity returns under the physical measure, $pvar_t$, has an unconditional mean of 0.0022, or an annualized volatility of about 16 percent. The gap between the risk-neutral and physical variance, the variance "premium," $vprem_t$, has an unconditional mean of 0.0014, implying a risk-neutral

[10] This regression suggests that $pvar_t$ loads heavily (and roughly equally) onto both lagged $rvar_t$ and $qvar_t$. We very strongly reject the hypothesis that there is no dependence on $qvar_t$.

annualized volatility of about 21 percent. Figure 2 plots the physical and risk-neutral conditional *volatilities* for our sample. As expected, the estimated risk neutral volatility is always higher than the physical volatility. Returning to Table 1, the risk neutral centered third moment qsk_t, has a negative mean, consistent with conditional skewness of the risk neutral distribution being negative on average, and the excess centered fourth moment of the risk-neutral distribution, qku_t, is greater than 0, suggesting positive excess kurtosis on average. These features of the risk-neutral distribution of return are consistent with those documented in other papers such as Bakshi and Madan (2006) and Figlewski (2009).

3.2. Consumption Moments and Dynamics

Our asset price data sample covers a relatively mild period for consumption growth. Monthly consumption growth volatility for the decade ending in 2009, even though it includes the financial crisis, was lower than for any other decade since the start of the data series the 1950's. Meanwhile, the upheaval in stock prices in 2008 and 2009 is more reminiscent of stock returns during the Great Depression. Of course, as emphasized by Barro (2006), it is certainly possible that market participants thought that Great Depression consumption dynamics were likely to return as well. Thus, to understand asset price dynamics in the recent period, it is likely necessary to take a longer-term view of possible consumption outcomes.

Unfortunately, monthly data on real consumption expenditures on nondurables and services do not extend back to the Depression era, but annual data is available back to 1929. Because we are trying to match features of the tails of the conditional consumption growth distribution we focus on long-term annual consumption data.

The first two columns of Table 2 report the features of consumption data we attempt to match. The first column reports sample statistics and the second reports block-bootstrapped standard errors, where the block length used was 5 years. Note that these standard errors are generally considerably larger than standard asymptotic standard errors. The top panel reports simple univariate statistics. For this sample, the mean rate of real consumption growth is about 3 percent per year and the sample standard deviation is about 2 percent. Contrary to the assumption of i.i.d. Gaussian dynamics that characterizes much of the asset pricing literature, the conditional distribution of annual consumption data for the U.S. exhibits quite rich dynamics. First, the autocorrelation coefficient is about 0.5. Second, the sample unconditional skewness and excess kurtosis of consumption growth are -1.8 and 6.4, respectively. While the kurtosis statistic is less than two standard deviations away from its value under Gaussianity, a standard Kolmogorov-Smirnov test of Gaussianity for the consumption sample (not reported) rejects at any conventional significance level. Finally, the probability of 2 and 4 standard deviation declines (or "crashes") are about 5.0 and 1.3 percent respectively, whereas the probabilities of crashes of these magnitudes under Gaussianity are 2.28 and 0.0031

percent respectively. Note that a 4 standard deviation decline occurred once in our 80 year sample.

In the middle panel, we report statistics describing the shift in the distribution of consumption growth following a "bad" realization in the prior year. Specifically, we report the change in the one-year ahead 10th, 25th, 50th, 75th and 90th percentiles for consumption growth, in units of unconditional standard deviation. The threshold for the "bad" outcome in the prior year is defined as a realization that is at or below the 15th percentile of the unconditional distribution. For example, a number of -2.00 for a particular quantile would mean that this quantile is 2 standard deviations below that quantile value for the unconditional distribution. The pattern is striking. Following a negative realization, the left tail of the distribution blows out. Specifically, the 10th and 25th percentiles of the conditional distribution are more than 2 standard deviations lower following a negative realization. On the other hand, the remaining conditional quantiles are not significantly different from their unconditional counterparts. The result is a conditional distribution that is much more sharply negatively skewed. This persistence in probabilities of extreme outcomes is exactly the kind of dynamics the BEGE model is designed to capture.

The bottom panel reports the conditional distribution following a one standard deviation positive shock. In contrast to the response to a negative shock, no clear shift in the skewness of the distribution is evident, although the distribution does shift up a bit.

These effects are plotted in Figure 3. The blue squares represent the unconditional distribution. The red down-triangles represent the distribution in the years following a "bad" consumption growth realization in the prior year. The green up-triangles represent the distribution following a "good" consumption growth realization in the prior year, defined as a realization exceeding the 85th percentile of the unconditional distribution.

Figure 4 presents a similar analysis for the shift in the *subjective* distribution of the growth rate of fundamentals, as measured using data from the Survey of Professional Forecasters (SPF). Participants in the survey are asked to complete histograms for the distribution of annual GDP growth outcomes for the current and next calendar year. The appendix describes how we use the histogram responses to construct the aggregate subjective distribution of one year-ahead GDP growth for each quarterly survey. The survey sample is from 1981Q3 through 2009Q4. The blue squares plot the unconditional average percentiles of the subjective distribution. The red down-triangles present the average subjective distribution conditional on the most recently published GDP growth rate having been "bad" (defined as reported four quarter GDP growth being less than the 15th percentile of the distribution of actual GDP growth outcomes over the survey sample). The green up-triangles represent the subjective distribution conditional on the most recent GDP release having been "good" (exceeding the 85th percentile of the unconditional distribution of actual four-quarter GDP growth). To ascertain that the most recent reading of four-quarter GDP growth was available

to survey respondents in real time, we use the Philadelphia Fed's real-time vintage data set. The figure shows that, even though the survey has been conducted during a period of mostly benign growth outcomes, a similar skewed downward shift is apparent following relatively adverse outcomes.

3.3. Structural Model Estimation

We use classical minimum distance (CMD) for estimation of the BEGE model, which relies on the matching of sample statistics.[11] All the sample statistics we attempt to fit using the BEGE model are collected into a vector, $\widehat{p}$, with estimated covariance matrix $\widehat{V}$. For $\widehat{p}$, we use all the sample statistics reported in Tables 1 and 2. Specifically, we ask the model to match all the long-term features of consumption growth reported in Table 2. To that set of 17 statistics, we add 9 unconditional sample statistics of asset prices: the means, volatilities and autocorrelations of the real short rate, rrf_t, the dividend yield, dp_t, real equity returns, ret_t. Finally, we seek to match 8 additional statistics about the higher-order return moments. These include the mean and volatility of: the conditional variance of returns under the physical measure, $pvar_t$, the variance premium, $vprem_t$, and the conditional risk-neutral third and fourth moments of returns, qsk_t and qku_t. In all, we ask the model to match 34 reduced-form statistics. By any measure, this represents an extremely challenging set of moments for a relatively parsimonious structural model. In fact, the model has only two stochastic shocks!

To estimate $\widehat{V}$, we assume a block diagonal structure. Let the consumption statistics, denoted $\widehat{p}_c$, be ordered as the first block in $\widehat{p}$ with the asset price statistics, $\widehat{p}_a$, second. For the upper left block of $\widehat{V}$, $\widehat{V}_{cc}$, we estimate the full covariance matrix for the BEGE consumption parameter estimates in Table 1 using the bootstrap method described above. We also estimate the full covariance matrix of the asset price statistics, $\widehat{V}_{aa}$, using the bootstrap procedure described earlier. We assume that the sampling errors for the consumption statistics, which use data back to 1929, are orthogonal to those of all the asset price statistics, which are available only from 1990 forward. That is, $\widehat{V}_{ac} = 0$.

We denote the true structural parameters by the vector, θ_0. The 18 parameters to be estimated are,

$$\theta = \left[\overline{g}, \sigma_{cp}, \sigma_{cn}, \overline{p}, \rho_p, \sigma_{pp}, \overline{n}, \rho_n, \sigma_{nn}, \overline{q}, \rho_q, \sigma_{qp}, \sigma_{qn}, \rho_x, \sigma_{xp}, \sigma_{xn}, \ln(\beta), \gamma\right]' \tag{28}$$

Under the null hypothesis that our model is true,

$$p_0 = h(\theta_0) \tag{29}$$

where $h(\theta)$ is a vector-valued function that maps the structural parameters into the reduced-form statistics. For the consumption statistics, we use a simulation of 10000 observations of the model, but the asset price

[11] See Wooldridge (2002), pg. 445-446 for a good textbook exposition on CMD.

statistics are all available in closed form[12]. To estimate the structural parameters, $\widehat{\theta}$, we minimize an objective function of the form,

$$\min_{\theta\in\Theta}\left\{\widehat{p}-h\left(\theta\right)\right\}'\widehat{W}\left\{\widehat{p}-h\left(\theta\right)\right\} \tag{30}$$

where $\widehat{W}$ is a symmetric, positive semi-definite, data-based weighting matrix. Efficient CMD suggests $\widehat{V}^{-1}$ for the weighting matrix, which we do employ for the block of consumption statistics. That is, $\widehat{W}_{cc}=\widehat{V}_{cc}^{-1}$. However, we use a diagonal weighting matrix for the asset price statistics, $\widehat{W}_{aa}=diag\left(\widehat{V}_{aa}^{-1}\right)$. We do this because some of the asset price statistics are nearly linearly dependent, rendering $\widehat{V}_{aa}$ nearly singular. Standard asymptotic arguments (relegated to the appendix) lead to a Gaussian limiting distribution of $\widehat{\theta}$, even under our nonstandard weighting matrix.

4. Results

In this section, we report on the estimation of the structural model parameters and then explore the model's implications for consumption dynamics and a variety of asset pricing phenomena.

4.1. Model estimation results

We fix two of the 18 parameters ex-ante. First, because the scale of the latent factor q_t is not well identified using our set of reduced-form statistics to be matched, we fix $\overline{q}=1$. Note that this does not restrict the level of risk aversion in the economy because γ is freely estimated. Second, we also fix $\ln\left(\beta\right)=-0.0003$ to aid in identification. This parameter is also only weakly identified using our estimation strategy, and fixing it does not seem to materially impact our ability to fit the moments of interest. Table 3 reports the remaining parameter estimates.

We first examine the dynamics of the conditional distribution of consumption growth. The mean of p_t is estimated at around 12.1, rendering shocks to ω_{pt} fairly close to being normally distributed when p_t is at its unconditional mean. In contrast, n_t has a very low mean of about 0.6, suggesting a strongly nonlinear distribution of $\omega_{n,t}$ shocks on average.[13] The distribution of consumption growth that emerges is one that is close to Gaussian over much of the range of Δc_t, but with a longer negative tail, suggesting occasional sharp declines in consumption. To illustrate this, Figure 5 shows the density of demeaned consumption growth under various configurations for p_t and n_t. To facilitate the visibility of the tails of the distribution, the base-10 logarithms of the densities are plotted. The top left panel shows that when n_t and p_t are at their

[12] The linear mappings between asset prices and the state vector are given in equations (16), (21), (22), (23), (26), and (27). Unconditional asset price statistics are then simple transformations of the unconditional moments of the state vector.

[13] For a $\Gamma\left(12,1\right)$ random variable, skewness is $2/\sqrt{12}=.6$ and excess kurtosis is $6/12=0.5$. For a $\Gamma\left(0.60,1\right)$ random variable, skewness is 2.6 and excess kurtosis is 10.

median values, the distribution of consumption growth does indeed have fatter tails than a corresponding Gaussian density with the same variance. Moreover, the left tail of the distribution is much fatter than the right tail relative to normality. The top right panel shows the density of consumption growth when p_t is at its 95th percentile value. At this configuration, even though the variance of consumption growth is high, its distribution is actually closer to the normal distribution. This is because the gamma distribution approaches the normal distribution for large values of the shape parameter. Nevertheless, it is clear that elevating p_t raises the right tail much more than the left tail, so that p_t is indeed a "good environment" state variable. The bottom left panel shows that when n_t is at its 95th percentile value, the distribution of consumption growth is still highly non-Gaussian, and the left tail is notably thicker compared to the upper right panel, justifying n_t 's role as a "bad environment" state variable. Finally, when both n_t and p_t take on their 95th percentile values (which happens very infrequently since they are independent), the distribution of consumption growth is again closer to normality due to the very high level of p_t and its large contribution to the overall variance of consumption growth. Note that n_t and p_t are both very persistent processes. Both consumption and expected consumption growth (x_t) are more sensitive to n_t than to p_t, but of course p_t has higher variance.

The estimated parameters relating to the properties of risk preferences are reported at the bottom. First, q_t, is found to be highly persistent, and significantly exposed to the n_t shock, with a coefficient of 0.0479 and a standard error of 0.0045. This implies that a positive shock to n_t, while lowering consumption growth, also raises risk aversion. This is quite consistent with the notion of habit persistence-based risk aversion like that in Campbell and Cochrane (1999). However, we do not find any significant exposure of q_t to p_t shocks. Finally, we find that γ is estimated to be 3.07. While this appears a reasonable number, recall that (local) risk aversion equals $\gamma \exp(q_t)$. Risk aversion is generally fairly mild, but with a long positive tail. The mean, median, and standard deviation of its distribution are 8.8, 7.8, and 2.8 respectively.

The final row of Table 1 reports over-identification tests for (1) the full set of statistics being fit, (2) the block of consumption statistics, and (3) the block of asset price statistics. The test for the full set of moments marginally fails to reject at the 5 percent level. Among the two subsets of moments, the asset price test statistic clearly rejects, while the test for the consumption moments does not. Overall, the model fit with respect to most of the salient features of consumption and asset price behavior is quite impressive, as we demonstrate next.

4.2. Fit with asset prices

In Table 1, we report the model-implied values for the fitted asset price statistics in square brackets for comparison to the previously discussed sample statistics. For the real short rate, the dividend yield and

equity returns, the means, volatilities and autocorrelations are all comfortably within two standard errors of the sample counterparts. Hence, the model fits the standard moments that are the focus of articles such as Bansal and Yaron (2004) and Campbell and Cochrane (1999).

Next, we report the means and volatilities of the conditional physical variance of equity returns, the variance premium, and the conditional risk-neutral third and fourth moments of returns. Although the mean of the variance premium is too low by about 40 percent (with the miss representing just over two standard errors of the sample statistic), the conditional third and fourth risk neutral moments are hit near-perfectly[14]. In addition, the model generates substantial volatility of the third and fourth moments, but still somewhat lower than is observed in the data.

In summary, the fit of the model with respect to a very wide class of asset pricing statistics is quite good. Drechsler and Yaron (2009), who introduce jumps in expected consumption growth and the volatility of consumption growth in the context of a long-run risk model, also fit asset price data very well, with the exception of the variability of the price-dividend ratio, which they under-estimate. They also slightly underestimate the mean of the variance premium (but less than we do) and over-estimate its variance.

4.3. The conditional distribution of consumption growth

The key property of the BEGE model relative to Campbell and Cochrane (1999) is that it permits substantial non-linearities in the consumption growth process. With sufficient non-linearities, it may be not surprising that the model can also fit option price dynamics in addition to standard asset pricing dynamics. It is therefore important to verify that the model's (conditional) consumption growth distribution fits the data well. We report the properties of annual consumption growth under BEGE in Table 2. In terms of the unconditional statistics, the implied moments are all comfortably within the two standard error band of the data moments, and often within one standard error. This is also true for the crash probabilities. If anything, the BEGE consumption model is slightly less non-normal than the data. In terms of the conditional distribution, the BEGE model is within two standard errors of the 10 empirical quantiles. In particular, the BEGE model captures the asymmetric increase in variance and skewness after a negative shock.

We also investigate how competing models fare with respect to the properties of consumption growth. We looked at four articles introducing non-linear features into the data generating process of consumption growth. Obviously, because these researchers do not estimate their models to be consistent with the statistics we introduce, we anticipate a less than perfect fit. Bollerslev, Tauchen and Zhou (2009) introduce stochastic volatility into the stochastic volatility process of consumption growth. However, they do not use consumption

[14]The level of our variance premium may be slightly over-estimated as we have used cash returns to measure the physical variance. Drechsler and Yaron (2009) claim this induces a slight upward bias in physical variance computations, due to serial correlation in the data and therefore use futures returns to measure realized variances.

data to calibrate the model, taking many parameters from the extant literature such as Bansal and Yaron's (2004) work, so that the consumption growth implications are, not surprisingly, not realistic. For example, because the stochastic volatility process is not constrained to be positive, at their parameter values, the conditional volatility of consumption growth is negative in more than half of all draws. Drechsler and Yaron (2009) do calibrate their model to fit properties of annual consumption growth. We take their parameter values and report the statistics implied by their model in Table 2. While they match the autocorrelation and standard deviation of consumption growth, the unconditional properties of consumption growth are too Gaussian relative to the data. In terms of the conditional distribution, the data moments are estimated too imprecisely to reject their model, but it clearly completely misses the pattern observed in the data. The model generates a variance preserving drop in conditional mean with no additional negative skewness after a bad consumption shock, and a symmetric increase in the conditional mean after a positive shock. This is perhaps not surprising as the DY model has jumps in the conditional growth rates (and in conditional variances), while consumption growth itself is conditionally Gaussian. To fit the pattern in the consumption data, the DY model could be amended to have jumps in the consumption growth process itself, but this may have weaker pricing implications than the jumps they include now.

Another important class of models considers the possibility of low-probability consumption crashes or "disasters." Barro's (2006) original work focused on matching the equity premium, and did not consider the pricing of options. His model in fact would not endogenously generate time-variation in risk premiums, let alone a variance premium. A recent article by Gabaix (2010) suggests that a Barro-type model can fit a large number of asset pricing puzzles, including features of option prices. However, the main mechanism to generate time-variation in risk premiums in Gabaix's work is a stochastic "recovery rate" process; consumption growth simply is a two-state process, which cannot match any of the consumption dynamics shown in Table 2. Nakamura, Steinsson, Barro, Ursua, (NSBU 2010) does not focus on options prices at all but presents a very intricate consumption growth process in this class. When a disaster hits, it potentially causes a permanent loss in output, at the same time as a temporary disaster. The model can generate disasters that unfold over time, and may also involve a rapid recovery after a disaster hits. The model features three shocks and two "disaster shocks," one being normally distributed, one featuring a skewed distribution (in one specification actually a gamma distribution). Ex-ante, this model may generate "BEGE-like" dynamics. NSBU estimate the model using a panel of countries assuming the disaster parameters to be the same across countries, but allowing other parameters to differ across countries. We take the estimated parameters for the US and reproduce the implied consumption dynamics in Table 2. The NSBU model fits the data considerably better than the DY model. The NSBU model generates consumption growth that is more volatile and leptokurtic than the data, but is realistic with respect to skewness and crash probabilities. The

additional volatility and leptokurtosis is likely an artifact of the use of disasters happening in other countries that were more severe than those occurring in the US. The model also fails to fit the conditional patterns we unearthed in the data. Essentially, the NSBU model generates a large increase in variance after a bad consumption shock, without generating additional skewness, and produces a similar albeit somewhat smaller increase in variance after a positive shock. It therefore also entirely misses the asymmetric pattern observed in the data. Nevertheless, it is possible that when the model is asked to fit these patterns, it may well succeed.[15]

One remaining potential problem is that our model may fit the annual consumption data well, but imply grossly unrealistic monthly consumption growth data, perhaps because non-normalities may weaken after temporal aggregation. To verify this, we filtered the state variables from the asset price data and monthly consumption data for the 1990-2009 sample period (see the appendix for a full description of the procedure). It is indeed the case that the relative importance of n_t increases dramatically during the recent crisis period. Figure 6 graphs the resulting first four conditional moments of monthly logarithmic consumption growth. The conditional mean of consumption growth shows very little time-variation indicating that x_t indeed plays a rather minor role in our model. Consumption growth volatility was very low in the mid-nineties, rising to between 2 and 2.5% for the early 2000s and then drops to below 1.5%. In 2008, consumption volatility rises to a peak of over 3%. Its skewness is always negative, varying between about 0 and -2. While the skewness is lowest in 2008, it was also quite low during the mid-90s. Excess kurtosis varies between about 9 and slightly less than 2. Remarkably, kurtosis was not highest during the recent crisis, but rather in the mid 90s. Overall, these conditional moments are plausible and by no means extreme.

It turns out that this reality check has bite. We also estimated a model imposing the Campbell -Cochrane (1999) assumption that shocks between consumption growth and q_t are perfectly correlated. In particular, we specify the q_t shock as:

$$q_t - E_t\left[q_{t+1}\right] = \phi_q\left(\Delta c_{t+1} - E_t \Delta c_{t+1}\right).$$

We estimate ϕ_q to be -3.13 with a standard error of 0.71. While the model is now rejected at the 5% level, the fit with most asset price data and the annual consumption data is not much worse than that of the more flexible model. However, the filtered monthly consumption growth statistics are much more extreme than what we just reported for the more flexible model and do not represent a plausible representation of monthly consumption growth distributions.

[15] Julliard and Ghosh (2010) formulate a variety of criticisms on the original Barro (2006) specification, one being that the disasters needed to explain the equity premium puzzle are simply not empirically plausible. Backus, Chernov and Martin (2009) show that index option data imply a more modest distribution of economic disasters than what is typically used in rare disaster models. We avoid both critiques by jointly fitting consumption and asset price data.

4.4. The dynamics of asset prices

Recall that there are only two fundamental shocks that drive all the state variables in the BEGE system, $\omega_{p,t}$ and $\omega_{n,t}$. Under our points estimates, $\omega_{p,t}$ and $\omega_{n,t}$ contribute significantly and roughly equally to consumption growth variation. For all the asset prices however, the story is much different. The impact of $\omega_{p,t}$, a positive p_t shock on asset prices, is very small due to concavity of preferences over consumption and also because q_t is not very sensitive to $\omega_{p,t}$. In contrast, the representative investor in this economy experiences substantial increases in marginal utility after positive $\omega_{n,t}$ shocks through lower consumption and higher risk aversion. Consequently, most of the variation in asset prices is accounted for by $\omega_{n,t}$ shocks. Figure 7 shows impulse responses from the two shocks to q_t, x_t, and selected asset prices. Positive $\omega_{n,t}$ shocks have a sharp negative impact on short rates, but increase the dividend yield. That the short rate and the dividend yield respond differently to $\omega_{n,t}$ shocks is reminiscent of a flight-to-safety event in which interest rates decrease sharply and at the same time equity prices decrease. The mechanism is on display in Figure 7. When $\omega_{n,t}$ increases, x_t decreases, but n_t and q_t increase. For the short rate, the x_t and n_t effects dominate and sharply decrease it. While this effect by itself drives up equity prices, the simultaneous increase in q_t dominates for equity prices, so that equity prices fall and the dividend yield moves up. Figure 7 shows that virtually all variation in higher-order asset price moments is also due to the $\omega_{n,t}$ shocks.

These dynamic effects also imply that the model endogenously generates asymmetric volatility (see Wu (2001)): the conditional covariance between returns and the conditional variance of returns is state dependent but mostly negative. The main mechanism is the positive dependence of both the conditional variance and the dividend yield (the inverse of the price dividend ratio) on $\omega_{n,t}$ that was just illustrated, and the fact that this shock dominates the variance of both variables.

4.5. Endogenous equity return predictability

Much of the asset pricing literature focuses on equity return predictability. Nevertheless, the return predictability evidence is rather weak. In Table 4, we present some univariate statistics for regressions of excess equity returns on the short rate, the dividend yield and the variance premium. None of the variables significantly predict future stock returns. The dividend yield is in fact the strongest predictor and the only one with a t-ratio larger than 1. Bollerslev, Tauchen and Zhou (2009) report that the variance premium is a highly significant predictor of equity returns. However, their main measure of the variance premium simply uses $rvar_t$ as the measure of conditional variance, $pvar_t$. In contrast, we use a projection of $rvar_t$ onto lagged variables to identify $pvar_t$. The last column of Table 4 shows that the variance premium measured as in Bollerslev, Tauchen and Zhou (2009) indeed significantly predicts equity returns for our sample.

Our structural model also generates a modest amount of return predictability. The maximum available r-squared statistic for the BEGE model (not reported in the table) is about 1 percent. We report the model-implied projection coefficients in brackets above the sample coefficients. For the dividend yield and variance premium, the model generates positive coefficients with magnitudes somewhat larger than their counterparts in the data, while the model generates a negative coefficient on the short rate. Economically, the equity premium depends predominantly and positively on n_t (κ^n is much larger than κ^p). Therefore, these results are easy to understand given the dynamic patterns we illustrated in the previous sub-section, with $\omega_{n,t}$ positively related to the dividend yield and the variance premium and negatively to short interest rates. The lack of a significant negative relationship between the short rate and future returns in the data appears to be a bit anomalous in this sample, as other researchers have generally found a significant negative relationship (see, for instance, Ang and Bekaert 2007). This could be due to our use of the real short rate which strips out inflation compensation.

The conditional Sharpe ratio for equity, the ratio of the conditional expected excess return to the conditional volatility, does vary substantially through time under the BEGE model. Figure 8 plots the Sharpe ratio as a function of n_t and p_t. The Sharpe ratio takes on a typical value of around 30 percent at an annual rate), and is not very sensitive to p_t. However, the conditional Sharpe ratio is very sensitive to shocks to n_t and can become as high as 50% when n_t exceeds its mean by about 3 standard deviations. Because this happens infrequently and in relatively bad times, the BEGE model's implications for the Sharpe ratio are potentially consistent with recent evidence on the counter-cyclical nature and rare occurrence of return predictability (see Henkel, Martin and Nardari (2009)).

In the context of our model, it is also of interest to examine the cyclical properties of the variance premium. In Table 5, we report the data correlation of the physical conditional variance and the variance premium with a business cycle indicator, namely expected four quarter consumption growth taken from the Survey of Professional Forecasters. We use this measure as we have a natural counterpart of expected consumption growth in our model, namely x_t.[16] The growth forecast has a -0.63 negative correlation with a dummy variable that takes a value of one for NBER-defined recessions. Both the conditional variance and the variance premium are, perhaps not surprisingly, strongly countercyclical, featuring a correlation of about -0.30 with expected consumption growth. Our model also generates countercyclical variances and variance premiums, with the correlation between x_t and the physical variance and variance premium being -0.79.

[16] The exact concept corresponding to the measure in the data would actually depend on other state variables because it represents the expectation of annual exponentiated consumption growth.

5. Conclusion

We have presented a new framework to model economic shocks. In our BEGE framework, there are two types of shocks: good environment shocks, which are positively skewed, and bad environment shocks, which are negatively skewed. Using this simple device and the convenience of gamma distributions, we can generate non-linear dynamics in a very tractable fashion. Most of the related work such as Drechsler and Yaron (2009), Bollerslev, Tauchen and Zhou (2009), and Bollerslev, Sizova and Tauchen (2009), uses Epstein-Zin preferences and other forms of nonlinearities in fundamentals. Instead, we append the BEGE technology to the well-known Campbell–Cochrane (1999) consumption-based asset pricing model. We demonstrate that the model fits the data very well, and fits features of the data that the Campbell-Cochrane model cannot fit, such as the conditional variance dynamics of equity returns, the variance premium, and other features of the risk-neutral distribution of returns which have received much recent attention. In addition, our model seems to fit consumption data better than the competing models.

Of course, many realistic features are missing from the particular model explored in this paper. The recent crisis reinforces the potential importance of Knightian uncertainty (see Drechsler (2009) and Epstein and Schneider (2007) for recent efforts) parameter uncertainty (Weitzman, 2007), and learning (see Veronesi (1999) and Shaliastovich, 2009) for understanding the joint dynamics of asset returns and fundamentals. Nevertheless, we feel that the technology introduced here can be very helpful to make headway in formulating models that break the curse of Gaussianity in a tractable fashion.

Two simple extensions may be worth exploring. First, the time-varying mean in consumption growth could be more directly linked to expectations about the future state of the economy. The current crisis again shows that anticipation of future bad economic conditions has marked implications on asset prices, yet, in our Campbell-Cochrane specification, the conditional mean of consumption growth plays a minor role and the fundamental drivers to the recovery are shocks to consumption growth.[17] Second, to keep the model as simple as possible, we priced a claim to consumption, not dividends. However, Longstaff and Piazzesi (2007) argue that corporate earnings are much more volatile than consumption growth and also more sensitive to economic conditions. They introduce jumps in the dividend process not present in the consumption growth process to generate more intricate stock return dynamics. Introducing more intricate cash flow dynamics into our model can be easily accomplished in a tractable fashion. This will also reduce the relatively high correlation between equity returns and consumption growth that our model currently implies.

[17] That said, recent work by Beeler and Campbell (2008) shows that a Campbell-Cochrane specification may be more consistent with the joint dynamics of stock prices and consumption growth than a "long-run risk" model.

6. Appendix

6.1. The General Model

We can write our model in general terms as follows

$$Y_{t+1} = \mu + AY_t + \Sigma_H \varepsilon_{t+1} + \Sigma_F \omega_{t+1} \tag{31}$$

Where Y_t $(n \times 1)$ is the state vector, μ $(n \times 1)$ is the associated mean parameter vector, A $(n \times n)$ is a transition parameter matrix, Σ_H $(n \times q)$ is the conditional volatility matrix for normally-distributed shocks, $\varepsilon_{t+1}(q \times 1)$, and Σ_F $(n \times p)$ is the conditional volatility matrix for the gamma-distributed shocks, ω_{t+1} $(p \times 1)$. The specific distributional assumptions for the shocks are

$$\begin{aligned} \varepsilon_{t+1}^i &\sim N(0,1),\ i = 1, ..., q \\ \omega_{t+1}^i &\sim \Gamma\left(k_t^i, 1\right) - k_t^i,\ k_t = \Phi Y_t,\ i = 1, ..., p \end{aligned} \tag{32}$$

and all the shocks are independent. The additive term in the ω_{t+1}^i definition, $-k_t^i$, sets the mean of the shock to zero. The parameter matrix $\Phi(p \times n)$ is comprised of only zeros and ones and selects which elements of Y_t determine the "shape" parameter of each ω_{t+1}^i shock.[18] For our main model, $Y_t = [p_t, n_t, \Delta c_t, q_t, x_t]'$, and the system matrices are:

$$\begin{aligned} \mu &= \left[\left(1-\rho_p\right)\overline{p}, \left(1-\rho_n\right)\overline{n}, \overline{g}, \left(1-\rho_q\right)\overline{q}, 0\right]' \\ A &= diag\left(\left[\rho_p, \rho_n, 0, \rho_q, \rho_x\right]\right), \Sigma_H = 0 \\ \Sigma_F &= \begin{bmatrix} \sigma_{pp} & 0 \\ 0 & \sigma_{nn} \\ \sigma_{cp} & -\sigma_{cn} \\ \sigma_{qp} & \sigma_{qn} \\ \sigma_{qx} & \sigma_{qn} \end{bmatrix}, \Phi = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{33}$$

The moment generating function of Y_{t+1} is given by Lemma 1.

Lemma 1 *For the random variable Y_t in Equation (31) the conditional expectation of an exponential-affine function of the state vector, $E_t\left[\exp\left(v'Y_{t+1}\right)\right]$,where v is a vector of constants (that is, the moment generating function under the physical probability measure), is given by*

$$\begin{aligned} E_t\left[\exp\left(v'Y_{t+1}\right)\right] &= \exp\left(v'\mu + v'AY_t\right) E_t\left[\exp\left(v'\Sigma_H\varepsilon_{t+1}\right)\right] E_t\left[\exp\left(v'\Sigma_F\omega_{t+1}\right)\right] \\ &= \exp\left(v'\mu + v'AY_t + \frac{1}{2}v'\Sigma_H\Sigma_H'v - \left(v'\Sigma_F + \ln\left(\mathbf{1} - v'\Sigma_F\right)\right)\Phi Y_t\right) \end{aligned}$$

where $\mathbf{1}$ is a $(1 \times p)$ vector of ones. The physical measures of the expectation and variance of $v'Y_{t+1}$ are defined, respectively, as

$$\begin{aligned} &\frac{d}{ds}\left[E_t\left[\exp\left(sv'Y_{t+1}\right)\right]\right]_{s=0} \\ &\frac{d^2}{ds^2}\left[E_t\left[\exp\left(sv'Y_{t+1}\right)\right]\right]_{s=0} \end{aligned} \tag{34}$$

They are given by:

$$\begin{aligned} E_t\left[v'Y_{t+1}\right] &= v'\mu + v'AY_t \\ V_t\left[v'Y_{t+1}\right] &= v'\Sigma_H\Sigma_H'v + \left(v'\Sigma_F\right)^{\bullet 2}\Phi Y_t \end{aligned}$$

[18] For a $\Gamma(k,1)$ distribution, the mean equals k, the variance equals k, the skewness is $2/\sqrt{k}$, and the kurtosis is $6/k$. The moment generating function is: $MGF_m = E\left[\exp\left(m\Gamma(k,1)\right)\right] = \exp\left(-k\ln(1-m)\right)$. The MGF is undefined for $m > 1$.

where $a^{\bullet 2}$ denotes the element-by-element exponentiation. For the third and fourth centered moments, straightforward calculations yield,

$$\begin{aligned} E_t\left[\left(v'Y_{t+1}\right)^3 - E_t\left[\left(v'Y_{t+1}\right)\right]^3\right] &= 2\left(v'\Sigma_F\right)^{\bullet 3}\Phi Y_t \\ E_t\left[\left(v'Y_{t+1}\right)^4 - E_t\left[\left(v'Y_{t+1}\right)\right]^4\right] - 3V_t\left[v'Y_{t+1}\right]^2 &= 6\left(v'\Sigma_F\right)^{\bullet 4}\Phi Y_t \end{aligned}$$

Lemma 2 *For the random variable Y_t in Equation (31), and a real pricing kernel, m_t, that is affine in current and lagged values of Y_t:*

$$m_t = m_0 + m_1'Y_t + m_2'Y_{t-1}, \tag{35}$$

the conditional risk-neutral expectation of an exponential-affine function of the state vector is defined as

$$E_t^Q\left[\exp\left(v'Y_{t+1}\right)\right] \equiv E_t\left[\exp\left(m_{t+1}\right)\right]^{-1} E_t\left[\exp\left(m_{t+1} + v'Y_{t+1}\right)\right] \tag{36}$$

and is given, using Lemma 1, by

$$\begin{aligned} & E_t^Q\left[\exp\left(v'Y_{t+1}\right)\right] \\ = \quad & \exp\left(\begin{array}{c} v'\mu + v'AY_t + \frac{1}{2}v'\Sigma_H\Sigma_H'v + m_1'\Sigma_H\Sigma_H'v \\ -\left(v'\Sigma_F + \ln\left(\mathbf{1} - \bullet\frac{v'\Sigma_F}{1-m_1'\Sigma_F}\right)\right)\Phi Y_t \end{array}\right) \end{aligned} \tag{37}$$

where $\bullet\frac{a}{b}$ denotes element-by-element division. Moreover, $E_t^Q\left[\exp\left(sv'Y_{t+1}\right)\right]$ is the risk-neutral moment generating function for $v'Y_{t+1}$. The risk neutral first and second moments of $v'Y_{t+1}$ can be found, respectively, by evaluating

$$\begin{aligned} & \frac{d}{ds}\left[E_t^Q\left[\exp\left(sv'Y_{t+1}\right)\right]\right]_{s=0} \\ & \frac{d^2}{ds^2}\left[E_t^Q\left[\exp\left(sv'Y_{t+1}\right)\right]\right]_{s=0} \end{aligned} \tag{38}$$

Upon evaluation, these reduce to:

$$\begin{aligned} E_t^Q\left[v'Y_{t+1}\right] &= v'\mu + v'AY_t + m_1'\Sigma_H\Sigma_H'v + \left(-v'\Sigma_F + \bullet\frac{v'\Sigma_F}{1-m_1'\Sigma_F}\right)\Phi Y_t \\ V_t^Q\left[v'Y_{t+1}\right] &= v'\Sigma_H\Sigma_H'v + \left(\bullet\frac{v'\Sigma_F}{1-m_1'\Sigma_F}\right)^{\bullet 2}\Phi Y_t \end{aligned}$$

Comparing the expressions for $V_t\left[v'Y_{t+1}\right]$ and $V_t^Q\left[v'Y_{t+1}\right]$, it is apparent that any variable that is an affine function of Y_{t+1} (such as returns) will have a zero variance risk premium (that is, its risk-neutral and physical conditional volatilities are equal) if the innovations to Y_{t+1} are limited to Gaussian shocks (that is, if $\Sigma_F = 0$). For the third and fourth centered moments, straightforward calculations yield,

$$\begin{aligned} E_t^Q\left[\left(v'Y_{t+1}\right)^3 - E_t^Q\left[\left(v'Y_{t+1}\right)\right]^3\right] &= 2\left(\bullet\frac{v'\Sigma_F}{1-m_1'\Sigma_F}\right)^{\bullet 3}\Phi Y_t \\ E_t^Q\left[\left(v'Y_{t+1}\right)^4 - E_t^Q\left[\left(v'Y_{t+1}\right)\right]^4\right] - 3V_t^Q\left[v'Y_{t+1}\right]^2 &= 6\left(\bullet\frac{v'\Sigma_F}{1-m_1'\Sigma_F}\right)^{\bullet 4}\Phi Y_t \end{aligned}$$

The risk-neutral moment generating function is thus very useful as direct computations of the risk neutral moments are quite involved. For example, even with Gaussian shocks, the first-order risk neutral expectation requires the use of Stein's lemma.

6.2. Risk Free Rate

It is well known that the gross real risk free rate is given by the inverse of the conditional expectation of the real pricing kernel. Then,

$$\exp(rrf_t) = (E_t[\exp(m_{t+1})])^{-1} \tag{39}$$

Applying Lemma 1, it is immediate that

$$r_t^f = -a_1^0 - a_1' Y_t \tag{40}$$

where a_0 and a_1' are given by Lemma 1 and shown explicitly in Section 2.

6.3. Equity Valuation

Let e_c be the selection vector such that $\Delta c_t = e_c' Y_t$. Maintaining our assumption that consumption equals dividends, the price dividend ratio is given by

$$\frac{P_t}{D_t} = E_t \sum_{n=1}^{\infty} \exp\left(\sum_{j=1}^{n} m_{t+j} + \Delta c_{t+j}\right) \equiv \sum_{n=1}^{\infty} q_{n,t}^0 \tag{41}$$

where $q_{n,t}^0 \equiv E_t \exp\left(\sum_{j=1}^{n} m_{t+j} + e_c' Y_{t+j}\right)$ are scalars. We will prove that $q_{n,t}^0 = \exp\left(b_n^0 + b_n' Y_t\right)$ where b_n^0 (scalar) and b_n' (k-vectors) are defined below. The proof is accomplished by induction. Consider $q_{1,t}^0$:

$$q_{1,t}^0 = E_t(\exp(m_{t+j} + e_c' Y_{t+j})) \tag{42}$$

Using Lemma 1,

$$q_{1,t}^0 = \exp\left(b_1^0 + b_1' Y_t\right) \tag{43}$$

where the exact expressions for b_1^0 and b_1' are given in Lemma 1. Next, suppose that $q_{n-1,t}^0 = \exp\left(b_{n-1}^0 + b_{n-1}' Y_t\right)$. Then rearrange $q_{n,t}^0$ as follows.

$$\begin{aligned}
q_{n,t}^0 &= E_t \exp\left(\sum_{j=1}^{n} m_{t+j} + e_c' Y_{t+j}\right) \\
&= E_t E_{t+1}\left\{\exp(m_{t+1} + e_c' Y_{t+1}) \exp\left(\sum_{j=1}^{n-1} m_{t+j+1} + e_c' Y_{t+j+1}\right)\right\} \\
&= E_t\left\{\exp\left(m_{t+1}^r + e_c' Y_{t+1}\right) E_{t+1} \exp\left(\sum_{j=1}^{n-1} m_{t+j+1} + e_c' Y_{t+j+1}\right)\right\} \\
&= E_t\left\{\exp(m_{t+1} + e_c' Y_{t+1})\, q_{n-1,t+1}^0\right\} \\
&= E_t\left\{\exp(m_{t+1} + e_c' Y_{t+1}) \exp\left(b_{n-1}^0 + b_{n-1}' Y_{t+1}\right)\right\} \\
&= E_t\left\{\exp\left(b_{n-1}^0 + m_{t+1} + (e_c + b_{n-1})' Y_{t+1}\right)\right\} \\
&= \exp\left(b_n^0 + b_n' Y_t\right)
\end{aligned} \tag{44}$$

where b_n^0 and b_n' are easily calculated using Lemma 1. Upon substitution, the recursions are those given in Section 2.

6.4. Log Linear Approximation of Equity Prices

In the estimation, we use a linear approximation to the logarithm of the price-dividend ratio. From the previous subsection, we see that the price dividend ratio is given by

$$\frac{P_t}{D_t} = \sum_{i=1}^{\infty} \exp\left(b_i^0 + b_i' Y_t\right) \tag{45}$$

where $Y_t = [p_t, n_t, \Delta c_t, q_t, x]$, $b_i^0 = \widetilde{A}_i$ and $b_i = \left[\widetilde{B}_i, \widetilde{C}_i, \widetilde{D}_i, 0, \widetilde{E}_i\right]$ with the coefficient sequences given in the text. We seek to approximate the log price-dividend ratio, pd_t, using a first order Taylor approximation of Y_t about $\overline{Y}$, the unconditional mean of Y_t. Let

$$\overline{q}_i^0 = \exp\left(b_i^0 + b_i'\overline{Y}\right) \tag{46}$$

and note that

$$\frac{\partial}{\partial Y_t}\left(\sum_{i=1}^{\infty} q_{i,t}^0\right) = \sum_{i=1}^{\infty}\frac{\partial}{\partial Y_t} q_{i,t}^0 = \sum_{i=1}^{\infty} q_{i,t}^0 \cdot b_i' \tag{47}$$

Approximating,

$$\begin{aligned} pd_t &\simeq \ln\left(\sum_{i=1}^{\infty}\overline{q}_i^0\right) + \frac{1}{\sum_{i=1}^{\infty}\overline{q}_i^0}\left(\sum_{i=1}^{\infty}\overline{q}_i^0 \cdot b_i'\right)\left(Y_t - \overline{Y}\right) \\ &= d_0 + d_1' Y_t \end{aligned} \tag{48}$$

where d_0 and d_1 are implicitly defined. Similarly,

$$\begin{aligned} gpd_t \equiv \ln\left(1 + \frac{P_t}{D_t}\right) &\simeq \ln\left(1 + \sum_{i=1}^{\infty}\overline{q}_i^0\right) + \frac{1}{1+\sum_{i=1}^{\infty}\overline{q}_i^0}\left(\sum_{n=1}^{\infty}\overline{q}_i^0 \cdot b_i'\right)\left(Y_t - \overline{Y}\right) \\ &= h_0 + h_1' Y_t \end{aligned} \tag{49}$$

where h_0 and h' are implicitly defined. Using pd_t and gpd_t we can span log equity returns. Using the definition of equity returns,

$$\begin{aligned} ret_{t+1} &= -pd_t + \Delta c_{t+1} + gpd_{t+1} \\ &\sim (h_0 - d_0) + \left(e_c' + h_1'\right) Y_{t+1} - d_1' Y_t \\ &= r_0 + r_1' Y_{t+1} + r_2' Y_t \end{aligned} \tag{50}$$

where r_0, r_1 and r_2 are implicitly defined

6.5. CMD Estimation Asymptotics

First note that the first order condition for our CMD optimization, described in Equation (30), is

$$\widehat{H}'\widehat{W}\left\{\widehat{p} - h\left(\widehat{\theta}\right)\right\} = 0. \tag{51}$$

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

$$h\left(\widehat{\theta}\right) = h(\theta_0) + H_0\left(\widehat{\theta} - \theta_0\right). \tag{52}$$

where $H_0 = \nabla_\theta h(\theta_0)$ is the gradient of $h(\theta)$ at the true parameter values. Combining Equations (51) and (52), we have,

$$\widehat{H}'\widehat{W}H_0\left(\widehat{\theta} - \theta_0\right) = \widehat{H}'\widehat{W}\left(\widehat{p} - h(\theta_0)\right) \tag{53}$$

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

$$\left(\widehat{\theta} - \theta_0\right) \sim N\left(0, \widehat{V_\theta}\right) \tag{54}$$

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, $\widehat{p}$.

6.5.1. Overidentification Test

Under efficient CMD, a simple overidentification test is available,

$$\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}\widehat{V}^{-1}\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}\sim\chi^2_{ns-np} \tag{55}$$

where ns and np are the size of $\widehat{p}$ and $\widehat{\theta}$ respectively. Under an alternative weighting matrix such as ours, a similar test statistic is available, but its distribution is different. To establish the distribution of

$$\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}\widehat{W}\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}=Obj \tag{56}$$

for $\widehat{W}\neq\widehat{V}^{-1}$, we follow Jagannathan and Wang (1996, JW henceforth). From the previous subsection, we obtain:

$$\begin{aligned}\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\} &= \widehat{p}-\left(h\left(\theta_0\right)+H_0\left(\widehat{\theta}-\theta_0\right)\right) &(57)\\ &= \left(I-H_0\left(H_0'\widehat{W}H_0\right)^{-1}H_0'\widehat{W}\right)\left(\widehat{p}-h\left(\theta_0\right)\right) &(58)\end{aligned}$$

Substitution into the objective function and rearrangement yields,

$$\begin{aligned}Obj &= \left(\widehat{p}-h\left(\theta_0\right)\right)'\left(\widehat{W}-\widehat{W}H_0\left(H_0'\widehat{W}H_0\right)^{-1}H_0'\widehat{W}\right)\left(\widehat{p}-h\left(\theta_0\right)\right) &(59)\\ &= Z'\left(\widehat{W}-\widehat{W}H_0\left(H_0'\widehat{W}H_0\right)^{-1}H_0'\widehat{W}\right)Z &(60)\end{aligned}$$

where Z is an ns dimensional random vector which is asymptotically normally distributed with zero mean and covariance matrix $\widehat{V}$. Defining $Z=\widehat{V}^{1/2}z$ where $\widehat{V}^{1/2}$ is the lower triangular Cholesky decomposition of $\widehat{V}$ and $z\sim N\left(0,I\right)$, we obtain,

$$Obj=z'Az \tag{61}$$

where $A=\widehat{V}^{1/2\prime}\widehat{W}^{1/2}\left(I-\widehat{W}^{1/2}\widehat{H}\left(\widehat{H}'\widehat{W}\widehat{H}\right)^{-1}\widehat{H}'\widehat{W}^{1/2}\right)\widehat{W}^{1/2}\widehat{V}^{1/2}$. JW show that A has $(ns-np)$ positive eigenvalues. Moreover, the quadratic form, $z'Az$, is easily simulated to derive critical values for Obj. Similar methods are used to calculate the critical values for the tests of fit for subsets of the matched statistics, reported in Table 3.

6.5.2. Filtering

To filter the state variables from observable endogenous variables conditional on the estimated model parameters reported in Table 3, we employ quasi-maximum likelihood and the Kalman filter. To do so, we replace the BEGE data generating process for Δc_t, p_t, n_t, q_t and x_t described in Section 2 with a conditionally Gaussian alternative model that preserves the dynamics of the conditional mean and covariance of the state variables, but ignores higher-order moments.

$$\begin{aligned}\Delta c_t &= \overline{g}+x_t+\sigma_{cp}\varepsilon_{p,t}-\sigma_{cn}\varepsilon_{n,t}\\ p_t &= \overline{p}+\rho_p\left(p_{t-1}-\overline{p}\right)+\sigma_{pp}\varepsilon_{p,t}\\ n_t &= \overline{n}+\rho_n\left(n_{t-1}-\overline{n}\right)+\sigma_{nn}\varepsilon_{n,t}\\ q_t &= \mu_q+\rho_q q_t+\sigma_{qp}\varepsilon_{p,t+1}+\sigma_{qn}\varepsilon_{n,t+1}\\ x_t &= \rho_x x_{t-1}+\sigma_{xp}\varepsilon_{p,t}+\sigma_{xn}\varepsilon_{n,t} &(62)\end{aligned}$$

where the shocks to the system are now Gaussian:

$$\begin{aligned}\varepsilon_{p,t} &\sim N\left(0,p_{t-1}\right) &(63)\\ \varepsilon_{n,t} &\sim N\left(0,n_{t-1}\right)\end{aligned}$$

As emphasized in the text, such a model would not be able to generate a positive variance premium or a non-Gaussian equity return risk neutral distribution. However, we are not using the above equations to estimate the BEGE model parameters or the dependence of the endogenous variables on the state variables. Rather, we simply preserve the estimated linear relationships between the endogenous variables and the state variables from the full BEGE model

$$Z_t = \widehat{\theta} + \widehat{\Theta} Y_t + u_t \tag{64}$$

where $Z_t = [\Delta c_t, rrf_t, dy_t, pvar_t, vprem_t, qsk_t, qkt_t]'$ and $Y_t = [p_t, n_t, \Delta c_t, q_t, x_t]'$ and $\widehat{\theta}$ and $\widehat{\Theta}$ are the loadings of Z_t on Y_t at the parameter estimates reported in Table 3. Because of the stochastic singularity of the system, we append Gaussian measurement error, u_t, to the measurement equations, (64),

$$u_t \sim N(0, R)$$

where R is a diagonal matrix. To estimate R and produce filtered values of Y_t, we maximize the likelihood of the entire system embodied by equations (62) and (64) using the standard Kalman filter. Because the Gaussian data generating process differs from the nonlinear BEGE process, the resulting filtered values of Y_t and the estimates of R are only approximate. However, it should be emphasized that when provided with correctly specified processes for the conditional means and covariances of the state and endogenous variables, the Kalman filter is the optimal linear filter in terms of root mean squared error.

Data for this estimation are described in the text with the exception of the monthly consumption growth variable, Δc_t. For this variable, we use real monthly consumption expenditures on nondurables and services as reported by the BEA over our sample period, January 1990 through December 2009.

6.6. Survey data

To create a measure of the subjective distribution of year-ahead real GDP growth, we utilize survey data available from the Survey of Professional Forecasters (SPF). Since 1981Q3, the SPF has asked respondents to fill in probabilities for histograms over real GDP growth outcomes for the coming year.[19] More specifically, respondents are asked to fill in the probability that real GDP growth will within specific ranges (e.g. 0 to 1 percent). Generally, several "bins" are presented in the surveys, but the bin boundaries have not been constant over time. To summarize the subjective distribution in a consistent manner across surveys, we use a procedure similar to that proposed by Giordani and Soderlind (2002). Specifically, for each survey date, we begin with the mean (across respondents) probabilities assigned to each bin, the "subjective bin probabilities" denoted $sprob_{i,t}$ where the subscript i indexes the $nbin$ bins and t denotes the date of the survey. We then fit a distribution, $f_{mix}(\cdot)$ to the subjective bin probabilities. For this purpose, we choose a flexible distribution, a Gaussian mixture:

$$f_{mix,t}(x) = \alpha_t \phi(m_{1,t}, s_{1,t}) + (1 - \alpha_t)\phi(m_{2,t}, s_{2,t})$$

where $\phi(m, s)$ is the Gaussian distribution with mean, m, and standard deviation s, and α_t, $m_{1,t}$, $s_{1,t}$, $m_{1,t}$, and $s_{1,t}$ are parameters to be estimated. Denote the probability assigned to each bin by $f_{mix,t}(x)$ as $tprob_{i,t}$. To estimate the parameters, we minimize the Pearson χ^2 test statistic measuring the distance between the subjective bin probabilities and the probabilities implied by $f_{mix,t}(x)$:

$$\{\widehat{\alpha}_t, \widehat{m}_{1,t}, \widehat{s}_{1,t}, \widehat{m}_{2,t}, \widehat{s}_{2,t}\} = \arg\min \left[\sum_{i=1}^{nbin} \frac{(sprob_{i,t} - tprob_{i,t})^2}{tprob_{i,t}} \right]$$

This identifies all the parameters of the mixture distribution for each time period, and the quantiles of the conditional distribution are then easily computed (simulated). In Figure 4, the squares plot the unconditional mean (over time) of several quantiles of the distribution, and the triangles plot the mean values of the

[19] Actually, the SPF asks for separate histograms for the current and following calendar years. To avoid seasonality and to roughly maintain a 1-year-ahead forecast horizon, we use a weighted average of the probabilities in the current and next calendar year. For first quarter surveys, we assign the full weight to the current year forecast. For second quarter surveys, we assign three-quarters weight to the current calendar year and one-quarter to the next calendar year, and so on.

quantiles conditional on the most recently published four-quarter GDP growth estimate from the BEA (see Figure 4 notes).

References

Abel, A., 1988, Stock Prices under Time-Varying Dividend Risk: An Exact Solution in an Infinite-Horizon General Equilibrium Model, Journal of Monetary Economics, 22, 375-393.

Abel, A., 1990, Asset Prices under Habit Formation and Catching up with the Joneses, American Economic Review, 80, 38-42.

Abel, A., 1999, Risk premia and term premia in general equilibrium, Journal of Monetary Economics, 43, 3-33.

Ang, A., Bekaert, G., 2007, Stock Return Predictability: Is it There?, Review of Financial Studies, 20(3), 651-707

Backus, D., M. Chernov and I. Martin, 2009, Disasters Implied by Equity Index Options, working paper.

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

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

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

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

Barsky, R., 1989, Why Don't the Prices of Stocks and Bonds Move Together?, American Economic Review, 79, 1132-45.

Beeler, J. and J. Y. Campbell, 2009, the Long-Run Risks Model: An Empirical Assessment.

Bekaert, G., Engstrom, E., Grenadier, S., 2005, Stock and Bond Returns with Moody Investors, working paper.

Bekaert, G., Engstrom, E., Xing, Y., 2009, Risk, Uncertainty, and Asset Prices, Journal of Financial Economics, 91, 59-82.

Bekaert, G., Liu, J., 2004, Conditioning Information and Variance Bounds on Pricing Kernels, Review of Financial Studies 17, 2, 2004, 339-378.

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

Bollerslev, Gibson, M., and H. Zhou, 2009, Estimation of Volatility Risk Premia and Investor Risk Aversion from Option-Implied and Realized Volatilities, Journal of Econometrics, forthcoming.

Bollerslev, T., Sizova, S., and G. Tauchen, 2009, Volatility in Equilibrium: Asymmetries and Dynamic Dependencies, working paper.

Bonomo, M., Garcia, R., 1994, Can a Well-Fitted Equilibrium Asset-Pricing Model Produce Mean Reversion?, Journal of Applied Econometrics, 9, 19-29.

Broadie, M., Chernov, M., Johannes, M., 2007, Model Specification and Risk Premia: Evidence from Futures Options, Journal of Finance, 62, 1453-1490.

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.

Campbell, J.Y., Thompson, S.B., 2008, Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?, Review of Financial Studies, 21, 1509-1531.

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

Cecchetti, S. G., Lam, P., Mark, N. C., 1990, Mean Reversion in Equilibrium Asset Prices, American Economic Review, 80, 398-418.

Chernov, M., Gallant, A.R., Ghysels, E., Tauchen, G., 2003, Alternative Methods for stock price dynamics, Journal of Econometrics, 116, 225-257.

Chernov, M., Ghysels, E., 2000, A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation, Journal of Financial Economics, 56, 407-458.

Drechsler, I. 2009, Uncertainty, Time-Varying Fear, and Asset Prices, working paper.

Drechsler, I., and Yaron, A., 2008, What's Vol Got to Do With It, The Wharton School.

Epstein, L.G., Schneider, M., 2007, Ambiguity, Information Quality and Asset Pricing, Journal of Finance, forthcoming.

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

Epstein, L. G., Zin, S. E., 2001,The independence axiom and asset returns, Journal of Empirical Finance, 8, 537-572.

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., 2008, Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance, NBER working paper.

Giordani, P. and P. Soderlind, Inflation Forecast Uncertainty, European Economic Review, 47, 1037-1059.

Gourio, F., 2010, Disaster Risk and Business Cycles, working paper.

Goyal, A., and Welch, I., 2008, A Comprehensive Look at the Empirical Performance of Equity Premium Prediction, Review of Financial Studies 21(4) 1455-1508.

Jagannathan, R. and Z. Wang, 1996, The conditional CAPM and the Cross-Section of Expected Returns, Journal of Finance, LI(1), 3-53.

Julliard, C. and A. Ghosh, 2010, Can Rare Events Explain the Equity Premium Puzzle, working paper.

Henkel, S., J. Martin and F. Nardari, 2008, Time-Varying Short-Horizon Return Predictability, working paper.

Kandel, S., Stambaugh, R., 1990, Expectations and Volatility of Consumption and Asset Returns, Review of Financial Studies, 3, 207-232.

Longstaff. F.A., and M. Piazzesi, 2004, Corporate Earnings and the Equity Premium, Journal of Financial Economics, 74, 401-421.

Lucas, R. E. Jr., 1978, Asset prices in an exchange economy, Econometrica 46, 1426-1446.

Nakamura, E., R. Barro, J. Steinsson, J. Ursua, 2010, "Crises and Recoveries in an Empirical Model of Consumption Disasters, working paper.

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

Pan, J., 2002, The jump-risk premia implicit in options: evidence from an integrated tie-series study, Journal of Financial Economics, 63(1), 3-50.

Pastor, L., Veronesi, P., 2006, Was There a NASDAQ Bubble in the Late 1990s?, Journal of Financial Economics, 81,61-100.

Poterba, M., Summers, L., 1986, The Persistence of Volatility and Stock Market Fluctuations, American Economic Review, 76, 1142-1151.

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

Shaliastovich, I, 2009, Learning, confidence and Option Prices, working paper.

Veronesi, P., 1999, Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model, Review of Financial Studies, 12, 975-1007.

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

Whitelaw, R., 2000, Stock Market Risk and Return: An Equilibrium Approach, Review of Financial Studies, 13, 3, 521-547.

Weitzman, M.L., 2007, Subjective Expectations and Asset-Return Puzzles, American Economic Review, 97, 1102-1130.

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

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

Table 1: Asset Price Sample Statistics and Model Fit

	$mean$	vol	$ac\,(1)$
rrf_t	[0.00038]	[0.0013]	[0.9674]
	0.00092	0.0012	0.9745
	(0.00052)	(0.00023)	(0.0258)
dp_t	[−6.2379]	[0.2825]	[0.9874]
	−6.3845	0.3369	0.9901
	(0.1642)	(0.0579)	(0.0149)
ret_t	[0.0050]	[0.0486]	[−0.0051]
	0.0069	0.0435	0.0859
	(0.0019)	(0.0036)	(0.1027)
$pvar_t$	[0.0023]	[0.0028]	
	0.0022	0.0030	
	(0.00047)	(0.0010)	
$vprem_t$	[0.0010]	[0.00117]	
	0.0014	0.00099	
	(0.00014)	(0.00017)	
qsk_t	[−0.00049]	[0.00057]	
	−0.00041	0.00085	
	(0.00013)	(0.00031)	
qkt_t	[0.000106]	[0.00012]	
	0.000097	0.00033	
	(0.000044)	(0.00013)	

This table reports on the ability of the structural model and parameter estimates shown in Table 3 to match the reduced-form statistics used in the CMD estimation. The model-implied statistics are shown in square brackets. The sample statistics are reported below with bootstrapped standard errors in parentheses. Data are monthly from January 1990 through December 2009. All variables are expressed at a monthly rate. The variables include the real short rate, rrf_t, the logarithmic dividend yield, dp_t, equity returns, ret_t, the conditional variance of returns under the physical measure, $pvar_t$ the variance premium, $qvar_t - pvar_t$, and the risk-neutral conditional third and fourth moments of month-ahead returns, qsk_t and qkt_t, respectively.

Table 2: Consumption growth statistics and model fit

	sample	*std.err*	*BEGE*	*DY*	*NSBU*
Unconditional moments					
mean	0.0320	(0.0024)	0.0348	0.0209	0.0200
std	0.0226	(0.0037)	0.0174	0.0257	0.0504
$ac\,(1)$	0.4773	(0.1553)	0.3285	0.5068	−0.0758
skew	−1.8280	(0.8002)	−0.9028	−0.0594	−1.5381
kurt	6.4287	(3.4952)	3.9967	0.7257	15.6917
Unconditional crash probabilities					
$2\sigma.$	0.0500	(0.0159)	0.0313	0.0250	0.0287
$4\sigma.$	0.0125	(0.0098)	0.0048	0.0005	0.0112
Distribution shift after negative realization					
Δq_{10}	−2.3333	(1.0266)	−1.6387	−0.7381	−0.8014
Δq_{25}	−1.7948*	(0.9366)	−0.6438	−0.7433	0.2575
Δq_{50}	−0.6684	(0.6345)	−0.2943	−0.7860	0.3422
Δq_{75}	−0.2846	(0.4592)	−0.2851	−0.8342	0.3877
Δq_{90}	−0.1887	(0.4517)	−0.2031	−0.8796	0.6021
Distribution shift after positive realization					
Δq_{10}	0.3221	(0.5714)	0.0707	0.8957	−0.3774
Δq_{25}	0.2029	(0.4115)	0.1924	0.8301	−0.3197
Δq_{50}	0.4127	(0.3371)	0.4522	0.7616	−0.2919
Δq_{75}	0.4111	(0.3191)	0.5237	0.6944	−0.1932
Δq_{90}	0.2656	(0.4256)	0.3849	0.6790	0.2810

For all rows, the first column reports unconditional univariate statistics for annual real consumption expenditures on nondurables and services from the US NIPA accounts between 1929 and 2009. The units are log growth of annual aggregate consumption. The second column reports bootstrapped standard errors for the US data. To construct these, the consumption growth series is sampled (with replacement and in 2 year blocks) to create bootstrapped samples. The standard deviation over 10,000 bootstrapped samples is reported. Estimates in the column "BEGE" reports simulated unconditional statistics implied by the BEGE model under the point estimates in Table 3. Because the BEGE model is specified at the monthly frequency in units of log growth, the simulated monthly data are time aggregated, in levels, over non-overlapping 12 month periods. The column labeled "DY" performs the same exercise for the model of Drechsler and Yaron (2010), which is also simulated at the monthly frequency and aggregated annually. The column labeled "NSBU" reports simulated statistics from the model estimated by Nakamura et al (2010) for the US. That model is specified and simulated at the annual frequency, so no time-aggregation is performed.

The top panel reports moments of the unconditional distribution of annual consumption growth. The two rows in the second panel report the unconditional probability of 2 and 4 standard deviation crashes under each model, where the standard deviation is that reported for each model in the top panel. The third panel reports the distribution of consumption growth for year $(t+1)$ conditional on a "bad" realization of consumption growth in year (t), with a "bad realization" defined as a reading of Δc_t that is less than the 15th percentile of the unconditional distribution. The fourth panel repeats the analysis of the third panel, but conditioning on the occurrence of a positive shock, defined as a Δc_t reading greater than the 85th percentile of the unconditional distribution.

$*$ Although its implied t-ratio is slightly below the standard 5 percent critical value, this statistic is significantly different from zero in that the bootstrapped 95 percent confidence interval is comprised of only negative values.

Table 3: BEGE Model Parameter Estimates and Overidentification Tests

Technology parameters								
p_t	$\overline{p}$	12.1244	ρ_p	0.9728	σ_{pp}	0.4832		
		(0.3220)		(0.0002)		(0.4034)		
n_t	$\overline{n}$	0.6342	ρ_n	0.9810	σ_{nn}	0.1818		
		(0.0656)		(0.0028)		(0.0008)		
Δc_t	$\overline{g}$	0.0031			σ_{cp}	0.0012	σ_{cn}	0.0052
		(0.0003)				(0.0004)		(0.0015)
x_t			ρ_x	0.9467	$\sigma_{xp}^{\times 10^3}$	-0.0016	$\sigma_{xn}^{\times 10^3}$	-0.0140
				(0.2684)		(0.0303)		(0.0735)
Preference parameters								
q_t	$\overline{q}$	1.0000	ρ_q	0.9841	σ_{qp}	0.0005	σ_{qn}	0.0479
		$(fixed)$		(0.0023)		(0.0029)		(0.0045)
m_t	$\ln(\beta)$	-0.0003	γ	3.0690				
		$(fixed)$		(0.3969)				
Overidentification test p-values								
	Overall			Consumption			Asset Prices	
	(0.0557)			(0.8463)			(0.0122)	

The model being estimated is summarized by the equations

$$\begin{aligned}
\Delta c_{t+1} &= \overline{g} + x_t + \sigma_{cp}\omega_{p,t+1} - \sigma_{cn}\omega_{n,t+1} \\
p_t &= \overline{p} + \rho_p (p_t - \overline{p}) + \sigma_{pp}\omega_{p,t} \\
n_t &= \overline{n} + \rho_n (n_t - \overline{n}) + \sigma_{nn}\omega_{n,t} \\
q_t &= \overline{q} + \rho_q (q_{t-1} - \overline{q}) + \sigma_{qp}\omega_{p,t+1} + \sigma_{qn}\omega_{n,t+1} \\
x_t &= \rho_x x_{t-1} + \sigma_{xp}\omega_{p,t} + \sigma_{xn}\omega_{n,t} \\
m_{t+1} &= \ln\beta - \gamma\Delta c_{t+1} + \gamma\Delta q_{t+1}
\end{aligned}$$

where all these equations are described in detail in Section 2. The model is estimated by CMD in which parameters are chosen to match the consumption statistics in Table 2 and the set of reduced-form asset price statistics in Table 1. Asymptotic standard errors are reported in parentheses. In the bottom panel, overidentification test p-values are presented for the full set of moments, the subset of moments based on consumption (the Table 2 statistics) and the subset of moments based on asset prices (those reported in Table 1). The overidentification test distributions and p-values are determined by Monte Carlo methods as described in the appendix.

Table 4: Equity Return Predictability

	rrf_t	dp_t	$qvar_t - pvar_t$	$qvar_t - rvar_t$
$model - implied\ coef$	[−1.9704]	[0.0160]	[3.9572]	–
$sample\ coef$	1.8455	0.0121	1.5878	3.4982
$sample\ s.e.$	(2.2409)	(0.0082)	(2.7954)	(1.2821)
$sample\ R^2$	0.0028	0.0090	0.0013	0.0296

This table presents the univariate predictability of one month ahead excess equity returns with respect to instruments listed in columns. All regressions are of the form

$$ret_{t+1} - rrf_t = \alpha + \beta \cdot inst_t + u_{t+1}$$

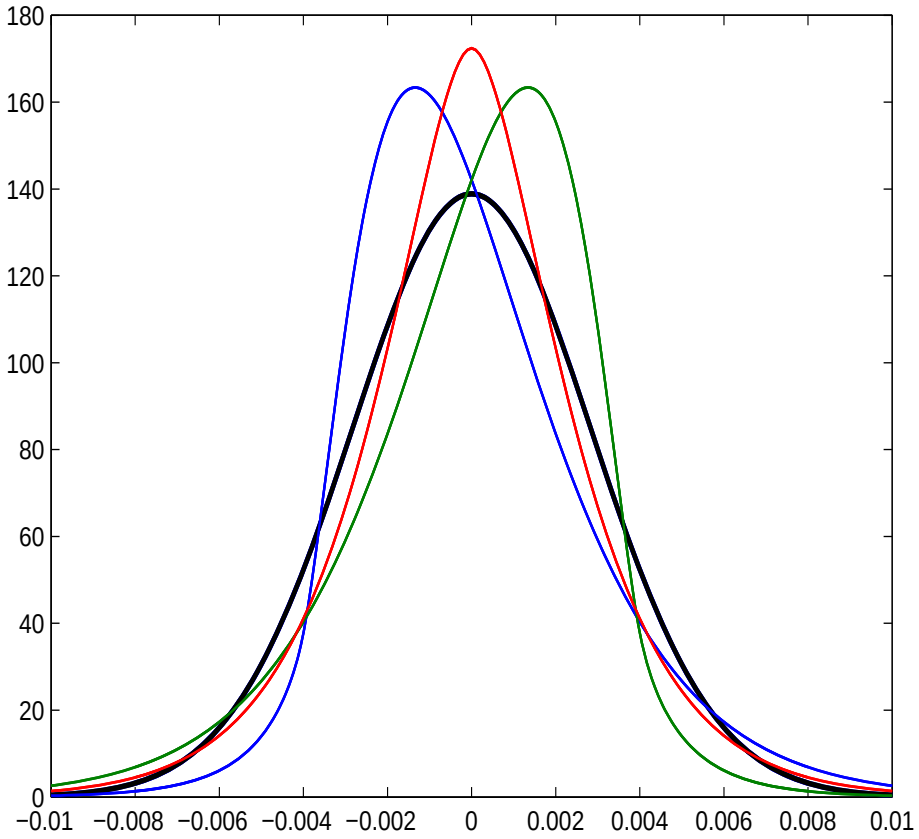
where $inst_t$ refers to the instrument listed at the top of each column. The sample is monthly from January 1990 through December 2009. The coefficients implied by the model and parameter estimates presented in Table 3 are listed first in square brackets. The corresponding coefficient in the data sample, along with its OLS standard error (in parentheses) and the associated R^2 statistic are listed below.

Table 5: Cyclicality of the Variance Premium

	sample	model-implied
	expcons	x
pvar.	−0.3587	[−0.7868]
	(0.0570)	
$qvar - pvar$	−0.2873	[−0.7868]
	(0.0599)	

This table reports the correlation of the conditional variance of equity returns, $pvar_t$, and the variance premium, $qvar_t - pvar_t$ with a measure of the state of the business cycle, expected four-quarter real U.S. NIPA consumption growth taken from the Survey of Professional Forecasters ("*expcons*"). Quarterly survey data is interpolated to the monthly frequency. The standard errors, reported in parentheses, are based on an asymptotic normal distribution of $0.5 * \ln((1+\rho)/(1-\rho))$, where ρ is the correlation, with a variance equal to $1/(N-3)$, where N is the number of observations. The final column reports the correlations of $pvar_t$, and $qvar_t - pvar_t$ with x_t, the state variable governing conditional consumption growth under the BEGE model at the point estimates in Table 3..

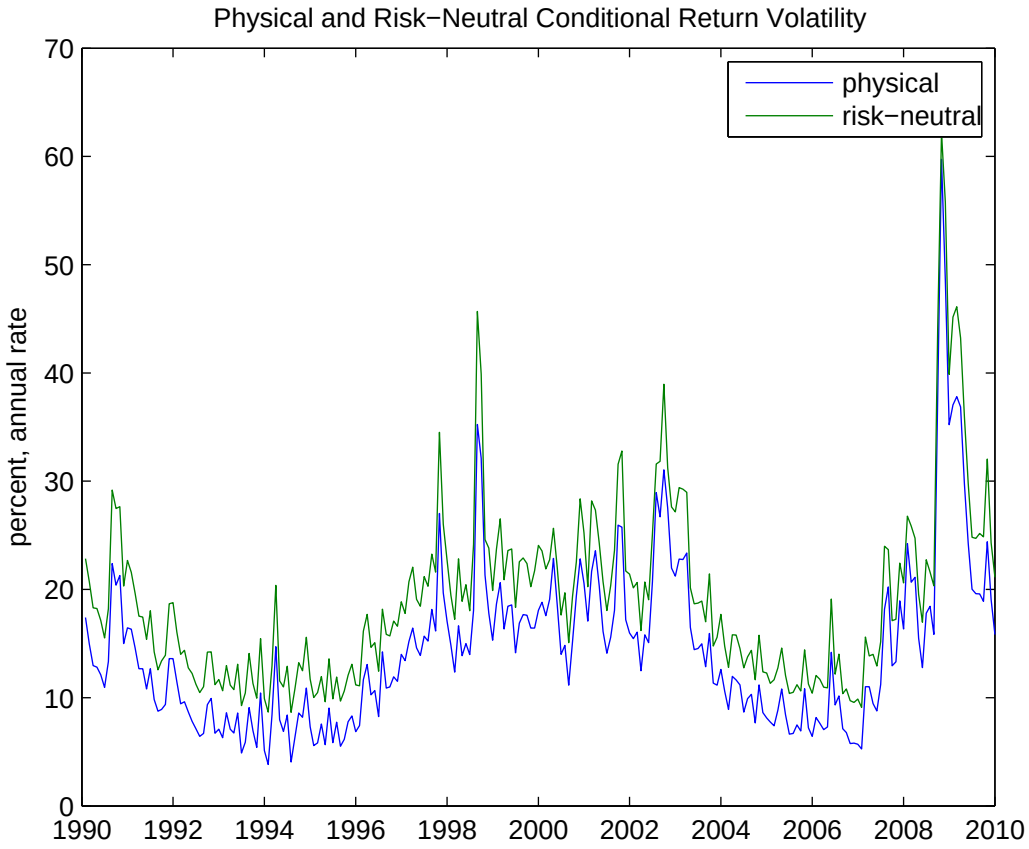
Figure 1: Examples of the BEGE distribution



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

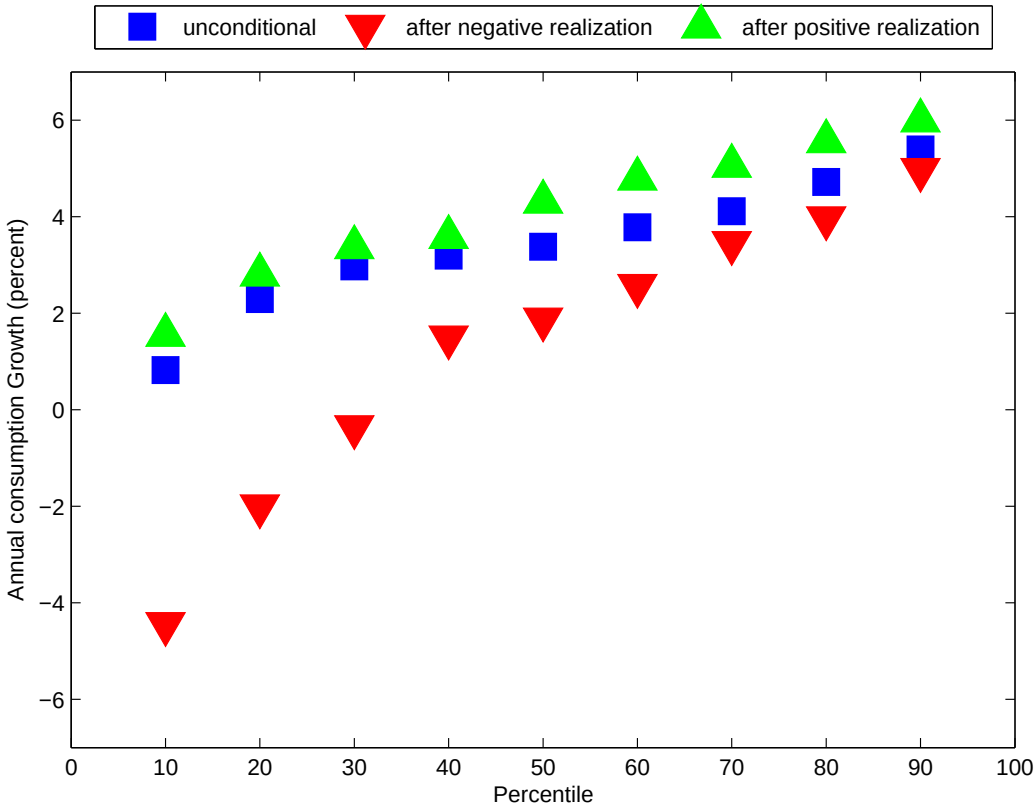
	p_t	n_t	σ_{cp}	σ_{cn}
black	40	40	0.0003	0.0003
red	2	2	0.0014	0.0014
green	.4	3	0.0016	0.0016
blue	3	.4	0.0016	0.0016

Figure 2: Physical and Risk-Neutral Conditional Equity Return Volatility



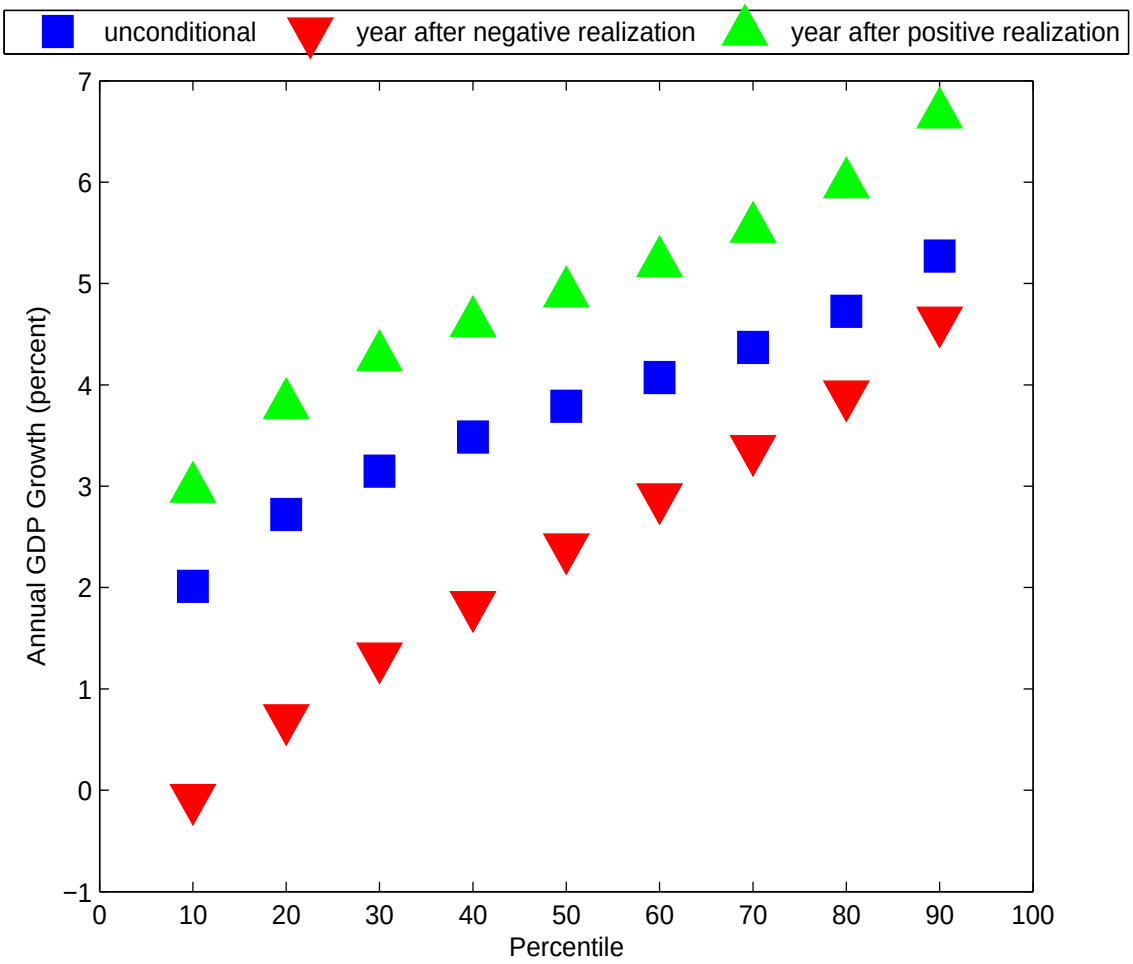
The Figure presents the time series of the conditional volatility of equity returns under the physical and risk-neutral measures. We calculate the risk-neutral conditional second moments of equity returns , $qvar_t$, using the method of Bakshi, Kapadia and Madan (2003). This involves calculating the prices of portfolios of options designed to have payoffs that are determined by particular higher order moments of returns. We obtained a panel of option prices across the moneyness spectrum for the S&P 500 index from 1996-2009 from OptionMetrics, and from 1990-1995 from Delta Neutral. We used the option contracts that have maturity closest to one month, and filtered out illiquid options according to the rules described in Figlewski (2009). We calculated the physical probability measure of equity return conditional variance, $pvar_t$, in two steps. We begin with the monthly realized variance, $rvar_t$, calculated as squared 5-minute capital appreciation returns over the month. Then we project $rvar_t$ onto one-month lags of the variables: $rvar_t$, and $qvar_t$. The fitted values from this regression are used to measure $pvar_t$.

Figure 3: Historical Distribution of Annual Consumption Growth



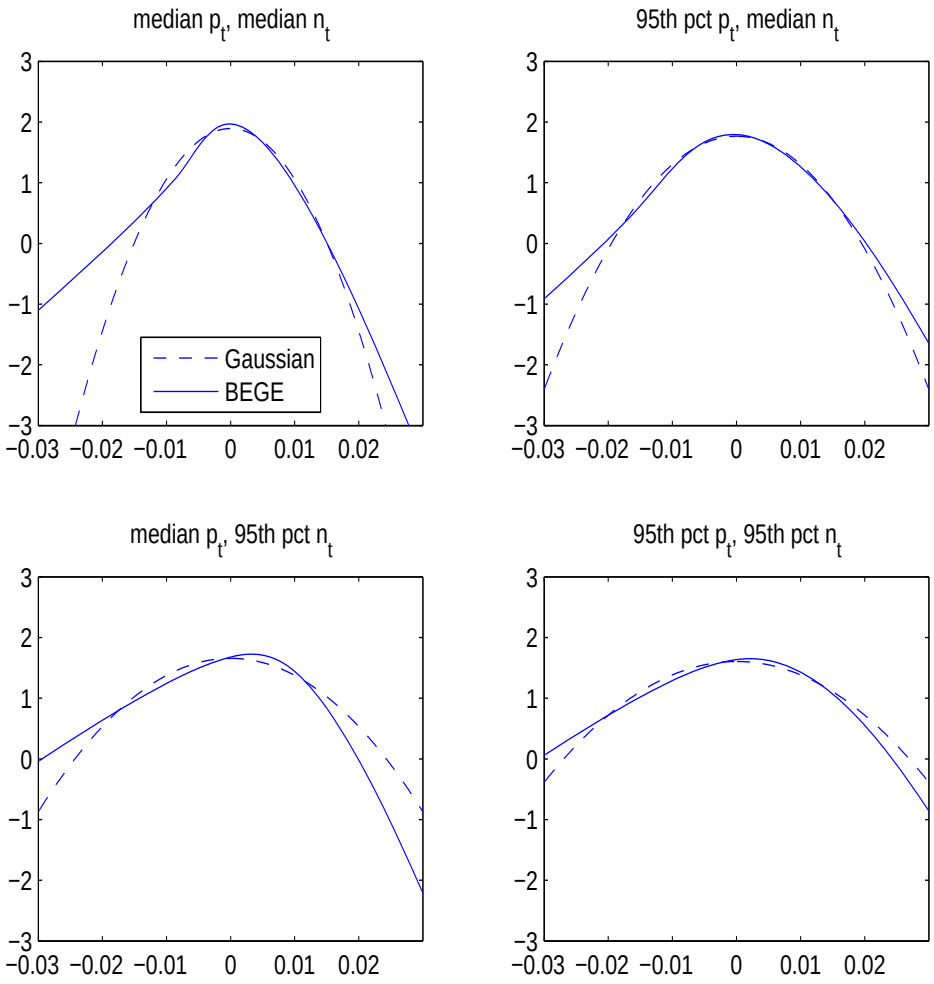
The figure presents the percentiles of the distribution of the log growth rate in annual real consumption expenditures on nondurables and services reported in the U.S. NIPA accounts. The blue squares represent the unconditional distribution. The red down-triangles represent the distribution in the years following a "bad" consumption growth realization in the prior year. A bad realization is defined as a consumption growth reading of less than the 15th percentile of the unconditional distribution (about 2 percent). The green up-triangles represent the distribution following a "good" consumption growth realization in the prior year, defined as a reading exceeding the 85th percentile of the unconditional distribution (about 4 percent).

Figure 4: Survey-Based Distribution of Annual GDP Growth



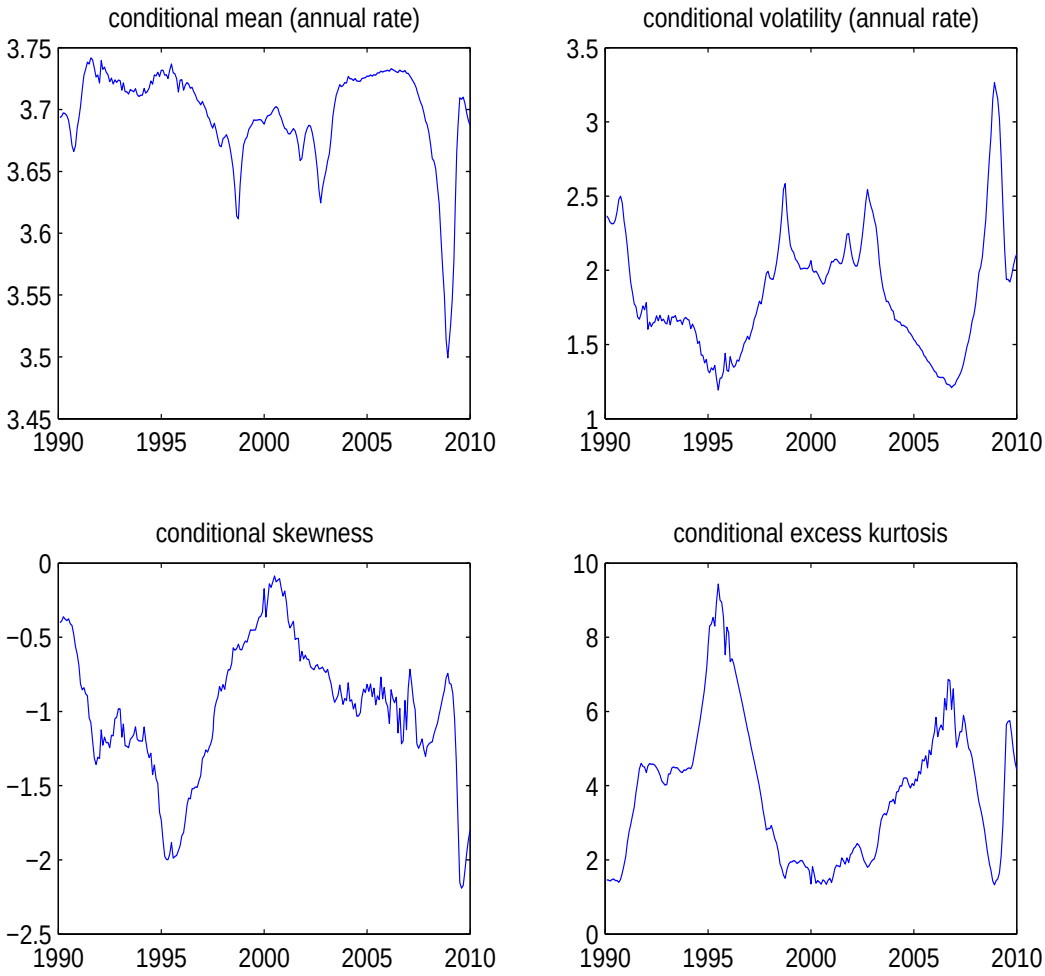
This figure presents the percentiles of the subjective conditional distribution of annual real GDP growth measured using survey data from the Survey of Professional Forecasters. Participants in the survey are asked to complete histograms for the distribution of annual GDP growth outcomes for the current and next calendar year. The appendix describes how we use the histogram responses to construct the aggregate subjective distribution of year-ahead GDP growth for each quarterly survey. The survey sample is from 1981Q3 through 2009Q4. The blue squares plot the unconditional average percentiles of the subjective distribution. The red down-triangles present the average subjective distribution conditional on the most recently released GDP growth figures from the BEA's quarterly release of having been "bad" (defined as reported four quarter GDP growth being less than the 15th percentile of the distribution of actual GDP growth outcomes). The green up-triangles represent the subjective distribution conditional on the most recent GDP release having been "good" (exceeding the 85th percentile of the unconditional distribution of actual four-quarter GDP growth). To ascertain that the most recent reading of four-quarter GDP growth was available to survey respondents in real time, we use the Philadelphia Fed's real-time vintage data set.

Figure 5: Log Density of Δc_t



This figure plots the log (base 10) density of (demeaned) monthly consumption growth under the BEGE model estimates presented in Table 3. Each panel presents the log density at a different configuration of p_t and n_t with each either at its model-implied median value, or its 95th percentile value. The quantiles of p_t and n_t are determined by simulation. Also plotted are Gaussian log densities with the same mean and variance as the BEGE density for each configuration of p_t and n_t.

Figure 6: Filtered Conditional Consumption Moments

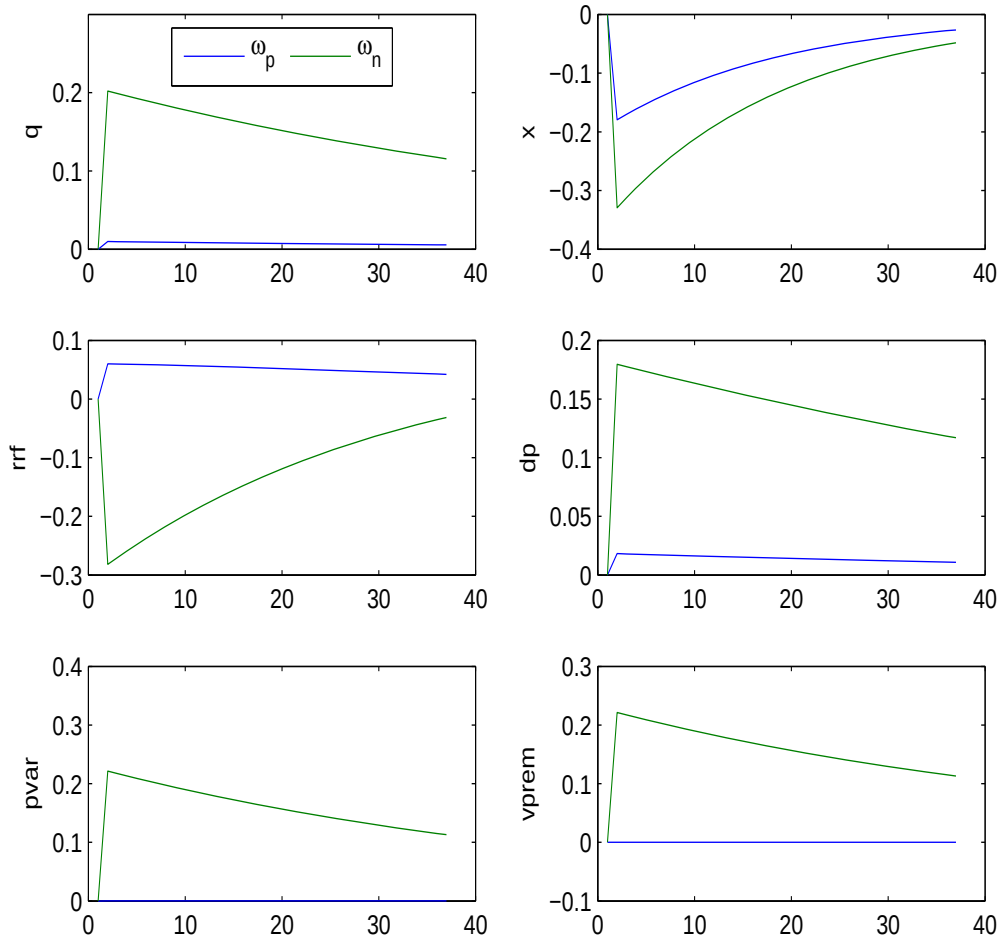


This figure reports filtered monthly time series for the conditional mean, volatility, skewness and kurtosis of month-ahead consumption growth. using the model and point estimates reported in Table 3. The blue lines show the smoothed (conditional on the full sample) estimates of the moments, which are functions of the filtered estimates of p_t,n_t,and $x_t$shown in Figure 9:

$$\begin{aligned} E_t\left[\Delta c_{t+1}\right] &= x_t + \overline{g} \\ \sigma_t\left[\Delta c_{t+1}\right] &= 1\left(\sigma_{cp}^2 p_t + \sigma_{cn}^2 n_t\right)^{1/2} \\ skew_t\left[\Delta c_{t+1}\right] &= 2\left(\sigma_{cp}^3 p_t - \sigma_{cn}^3 n_t\right)/\sigma_t^{3/2}\left[\Delta c_{t+1}\right] \\ xkurt_t\left[\Delta c_{t+1}\right] &= 6\left(\sigma_{cp}^4 p_t + \sigma_{cn}^4 n_t\right)/\sigma_t^{4/2}\left[\Delta c_{t+1}\right] \end{aligned}$$

A standard linear Gaussian Kalman filter is used in conjunction with an approximate Gaussian version of the structural model. The filtering procedure is described in detail in the appendix.

Figure 7: Impulse Responses of Asset Prices to Shocks

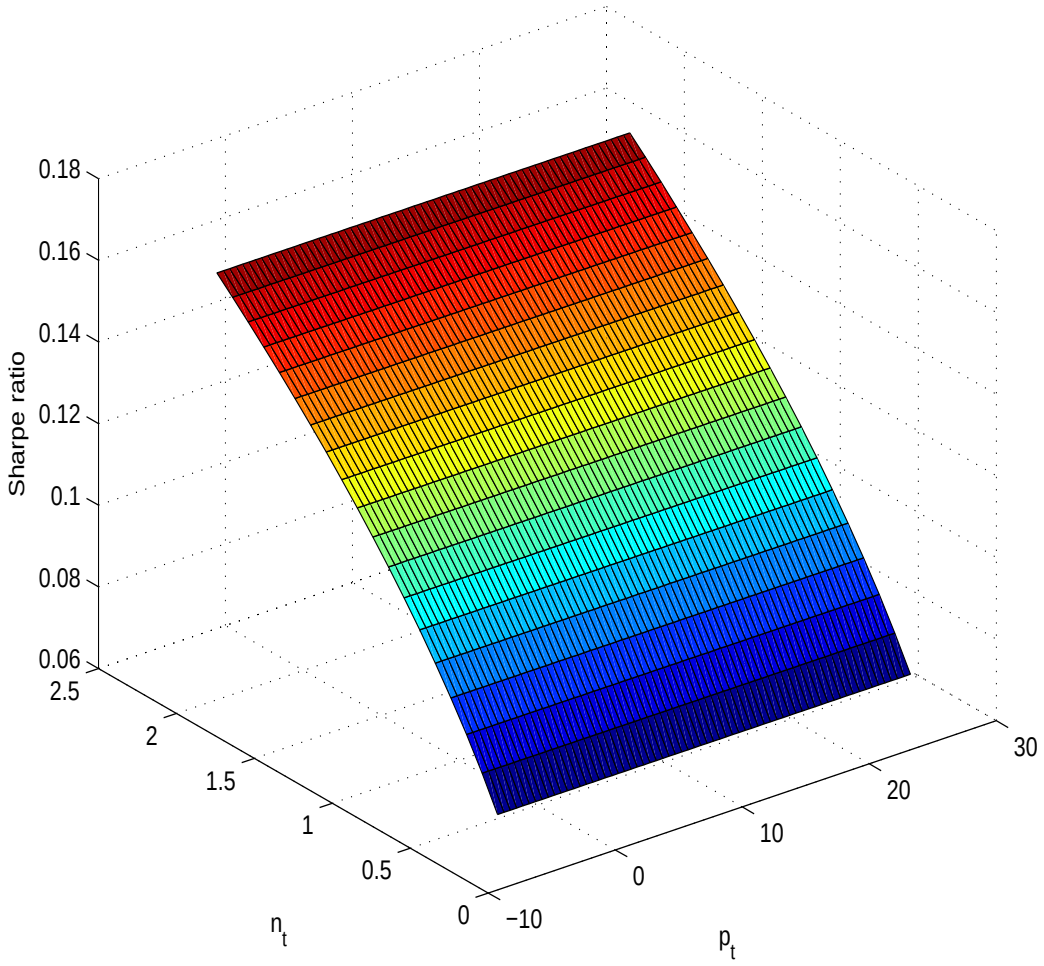


This figure shows the impulse response of the state variables q_t, x_t, and the endogenous variables, rrf_t, dp_t, $pvar_t$, and $vprem_t$ to the shocks, $\omega_{p,t}$ and $\omega_{n,t}$. For all variables, the units on the vertical axis are unconditional standard deviations. In all panels, the shocks occur at month 1 and the horizontal axis runs from 0 months (prior to the shock) through 36 months. Impulse responses to 90th percentile values of $\omega_{p,t}$ and $\omega_{n,t}$ are reported. The response of each endogenous variable in j periods, iz_{t+j}, is given by

$$iz_{t+j} = h_z \begin{bmatrix} \rho_p & 0 & 0 & 0 \\ 0 & \rho_n & 0 & 0 \\ 0 & 0 & \rho_q & 0 \\ 0 & 0 & 0 & \rho_x \end{bmatrix}^{j-1} \begin{bmatrix} \sigma_{pp} & 0 \\ 0 & \sigma_{nn} \\ \sigma_{qp} & \sigma_{qn} \\ \sigma_{xp} & \sigma_{xn} \end{bmatrix} \begin{bmatrix} \omega_{p,t+1} \\ \omega_{n,t+1} \end{bmatrix}$$

where h_z is the loading of the variable on $Y_t = [p_t, n_t, q_t, x_t]$.

Figure 8: State Dependence of the Equity Conditional Sharpe Ratio



This figure reports the monthly Sharpe ratio for one-month ahead equity returns under the structural model and point estimates in Table 3 calculated as:

$$Sharpe\ ratio = \frac{E_t\left[ret_{t+1} - rrf_t\right]}{VAR_t\left[ret_{t+1}\right]^{1/2}}$$