Asset Return Dynamics under Bad Environment-Good Environment Fundamentals

Geert Bekaert*
Columbia University and NBER

Eric Engstrom†
Federal Reserve Board of Governors

This Draft: June 2009

JEL Classifications
G12, G15, E44

Keyphrases
Equity premium, Volatility Premium, Countercyclical risk aversion, Economic Uncertainty, Dividend yield

Abstract:

We present a consumption-based asset pricing model that not only fits the standard salient asset return features including the equity premium, low risk free rate and realistic volatility of equity returns and yields, but also generates a realistic variance premium and option prices. The model borrows its preference structure largely from Campbell and Cochrane (1999) but introduces a new technology for consumption growth. The model incorporates a "bad environment-good environment" framework that generates realistic time-varying volatility, skewness and kurtosis in fundamentals while still permitting closed-form solutions for asset prices.

*Columbia Business School, 802 Uris Hall, 3022 Broadway, New York New York, 10027; ph: (212)-854-9156; fx: (212)-662-8474; gb241@columbia.edu.

†Board of Governors of the Federal Reserve System, Washington, DC, 20551, ph: (202)-452-3044; fx: (202)-728-5887; eric.engstrom@frb.gov.

The authors especially thank Stephen Figlewski for providing time series data on the risk neutral density of returns. The views expressed in this article do not necessarily represent those of the Federal Reserve System or its 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 cannot be time variation in risk premiums unless the probability of the "crash" is assumed to vary through time (see Gabaix (2009), and Wachter (2008)).

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 contract essentially provides direct readings on the risk-neutral variance; see 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 include time-varying (but generally negative) skewness, fat tails, and a strong negative correlation between return realizations and risk-neutral volatility (see, for instance, Figlewski (2009)).

To generate these features of the risk-neutral distribution in the reduced-form literature, structural models must endogenously generate time-varying skewness in returns. However, most existing structural 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 generates 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.

The framework is also 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). We feel that 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. Of course, the model can only be deemed successful if fundamentals indeed exhibit non-linearities in the data, which, through an acceptable preference structure, lead to realistic asset pricing dynamics. 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 we use and documents that there are indeed time-varying non-linearities in the consumption growth process. Much of what we do here confirms results in the literature, with some additions regarding the conditional skewness of consumption growth. Section 4 sets out the estimation strategy. Section 5 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. Section 6 discusses some robustness checks and extensions of the BEGE framework. The final section offers some concluding remarks, and compares our findings to contemporaneous articles by Bollerslev, Tauchen and Zhou (2009) and Drechsler and Yaron (2008), that have similar goals but a very different framework. We also provide further motivation for the BEGE fundamental dynamics using survey-based measures of the conditional distribution of economic growth.

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:

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

where $\Delta c_t = \ln(C_t) - \ln(C_{t-1})$ is the logarithmic change in consumption, $\overline{g}$ is the mean rate of consumption growth, which we assume is constant, and the parameters σ_{cp} and σ_{cn} are both positive. The shocks, $\omega_{p,t+1}$ are $\omega_{n,t+1}$ zero-mean innovations with the following distributions,

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

where $\Gamma(p_t, 1)$ represents the gamma distribution with shape parameter, p_t, and size parameter equal to 1. The shape parameters, p_t and n_t will be modeled as time-varying (positive) latent factors, the data generating process for which will be introduced shortly. These factors 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. To understand what this 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,

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

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.[3] 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:

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

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 $sc_t/vc_t^{3/2}$, will be positive when p_t is relatively large, and negative when

[3] 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, $E[\exp(m(x_1 - x_2))] = E[\exp(mx_1)]/E[\exp(mx_2)]$.

n_t is large. This is the essence of the BEGE model: the bad environment refers to an environment in which the $\omega_{n,t}$ 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 p_t and n_t. 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.

While we have represented the BEGE distribution as a mixture 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 (but very messy) analytic solution for the BEGE density function is also available (upon request from the authors). 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 latent factor p_t as following a simple, autoregressive process with square-root volatility dynamics,

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

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,

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

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

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

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. However, this assumption could be easily relaxed within our general framework. Moreover, the conditional covariance of Δc_{t+1} and its own conditional variance, vc_t is:

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

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

2.2 Preferences

We now describe the preferences of the representative agent in our model. Consider a complete markets economy as in Lucas (1978), but modify the preferences of the representative agent to have the form:

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

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 framework of Abel (1990, 1999) who specifies preferences where H_t represents past or current aggregate consumption, which a small individual investor takes as given, but she then evaluates her own utility relative to that benchmark.[4] That is, utility has a "keeping up with the Joneses" feature. In Campbell and Cochrane (1999), H_t is taken as an exogenously modelled subsistence or habit level. Hence, the local coefficient of relative risk aversion equals $\gamma \cdot \frac{C_t}{C_t-H_t}$, where $\left(\frac{C_t-H_t}{C_t}\right)$ is defined as the surplus ratio[5]. As the surplus ratio goes to zero, the consumer's risk aversion goes to infinity. In our model, we view the inverse of the

[4] For empirical analyses of habit formation models, where habit depends on past consumption, see Heaton (1995) and Bekaert (1996).

[5] 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."

surplus ratio as a preference shock, which we denote by Q_t. Thus, $Q_t = \frac{C_t}{C_t - H_t}$. Risk aversion is now characterized by $\gamma \cdot Q_t$, and $Q_t > 1$. 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 of 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 predictability evidence than the standard model with power utility because it can generate counter-cyclical expected returns and prices of risk. The unobserved process for $q_t \equiv \ln(Q_t)$ 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_q and ϕ_p and ϕ_n are parameters. Here, we have allowed the innovation in q_t to be spanned by the consumption innovations, $\sigma_{cp}\omega_{p,t+1}$ and $\sigma_{cn}\omega_{n,t+1}$. 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 CC in q_t entirely spanned by the consumption shocks, but there are two such shocks in our framework and these shocks are heteroskedastic.[6] 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 was a non-linear process that was increasing in q_t. Our modeling here is different and a bit 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

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

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 term structure

We first solve for the real risk free short rate,. rrf_t, in our framework and then the price of a real consol. The latter will be useful for comparison with equity prices.

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

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

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

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

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

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

For ease of interpretation, we will 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,

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

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 $\overline{g} > 0$ and/or,

$q_t > \overline{q}$), investors desire to borrow to smooth marginal utility, and so risk free rates must rise. The bottom two lines capture precautionary savings effects, that is, the desire of investors to save more in uncertain times. Notice that because the function $f(x) = x + \ln(1-x)$ is always negative, the precautionary savings effects are also always negative. A third-order Taylor expansion of the log function helps with the interpretation of rrf_t:

$$rrf_t \approx \begin{pmatrix} -\ln\beta + \gamma\overline{g} + \gamma\left(1-\rho_q\right)(q_t - \overline{q}) \\ +\left(-\frac{1}{2}a_p^2 - \frac{1}{3}a_p^3\right)p_t \\ +\left(-\frac{1}{2}a_n^2 - \frac{1}{3}a_n^3\right)n_t \end{pmatrix} \tag{17}$$

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. In this way, our model may generate the kind of extremely low but also very volatile risk free rates witnessed in the 2007-2009 crisis period.

The price of a risk free real consol We now extend the characterization of the real term structure to a risk-free real consol, that is as asset that pays a real coupon, normalized to 1, each period. Under standard no-arbitrage arguments, the price of the consol, PC_t, must obey:

$$PC_t = E_t\left[\sum_{i=1}^{\infty} \exp\left(\sum_{j=1}^{i} m_{t+j}\right)\right] \tag{18}$$

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.

Proposition 1 *For the economy described by Equations (1) through (11), the price of a risk free* *real consol paying one unit of the consumption good is given by*

$$PC_t = \sum_{i=1}^{\infty} \exp\left(A_i + B_i p_t + C_i n_t + D_i q_t\right) \tag{19}$$

where the initial values of the parameter sequence are given by

$$\begin{aligned}
A_1 &= \ln\beta - \gamma\overline{g} + \gamma\left(1-\rho_q\right)\overline{q} \\
B_1 &= -a_p - \ln\left(1-a_p\right) \\
C_1 &= -a_n - \ln\left(1-a_n\right) \\
D_1 &= -\gamma\left(1-\rho_{qq}\right)
\end{aligned}$$

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

$$\begin{aligned}
A_i &= A_{i-1} + B_{i-1}\mu_p + C_{i-1}\mu_n + D_{i-1}\mu_q \\
B_i &\equiv \left(-a_p + B_{i-1}\left(\rho_p - \sigma_{pp}\right) - D_{i-1}\sigma_{qp}\right) - \ln\left(1 - a_p - B_{i-1}\sigma_{pp} - D_{i-1}\sigma_{qp}\right) \\
C_i &\equiv \left(-a_n + C_{i-1}\left(\rho_n - \sigma_{nn}\right) - D_{i-1}\sigma_{qn}\right) - \ln\left(1 - a_n - C_{i-1}\sigma_{nn} - D_{i-1}\sigma_{qn}\right) \\
D_i &\equiv D_1 + D_{i-1}\rho_{qq}
\end{aligned}$$

(Proof is available in separate appendix).

The most useful expressions above for gaining intuition about consol pricing are those for B_1 and C_1. First, note that B_1 and C_1 are always positive because the function $f(x) = -x - \ln(1-x)$ is always positive. Moreover, one can easily show that B_i and C_i are positive for all i as well. Hence, increases in n_t and p_t always increase real consol prices, another implication of the precautionary savings channel. Finally, the D_n term captures the effect of the risk aversion variable, q_t, which affects bond prices through utility smoothing channels; therefore increases in q_t tend to depress consol prices.

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

$$\frac{P_t}{D_t} = E_t\left[\sum_{i=1}^{\infty} \exp\left(\sum_{j=1}^{i}\left(m_{t+j} + \Delta d_{t+j}\right)\right)\right] \tag{20}$$

where $\frac{P_t}{D_t}$ is the equity price-dividend ratio and Δd_t represents 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.

Proposition 2 *For the economy described by Equations (1) through (11), the price-dividend ratio of equity is given by*

$$\frac{P_t}{D_t} = \sum_{i=1}^{\infty} \exp\left(\widetilde{A}_i + \widetilde{B}_i p_t + \widetilde{C}_i n_t + \widetilde{D}_i q_t\right) \tag{21}$$

where the initial values of the parameter sequence are given by

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

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

$$\begin{aligned}
\widetilde{A}_i &= \widetilde{A}_1 + \widetilde{A}_{i-1} + \widetilde{B}_{i-1}\mu_p + \widetilde{C}_{i-1}\mu_c + \widetilde{D}_{i-1}\mu_q \\
\widetilde{B}_i &\equiv \left(-a_p - \sigma_{cp} + \widetilde{B}_{i-1}\left(\rho_p - \sigma_{pp}\right) - \widetilde{D}_{i-1}\sigma_{qp}\right) - \ln\left(1 - a_p - \sigma_{cp} - \widetilde{B}_{i-1}\sigma_{pp} - \widetilde{D}_{i-1}\sigma_{qp}\right) \\
\widetilde{C}_i &\equiv \left(-a_n + \sigma_{cn} + \widetilde{C}_{i-1}\left(\rho_n - \sigma_{nn}\right) - \widetilde{D}_{i-1}\sigma_{qn}\right) - \ln\left(1 - a_n + \sigma_{cp} - \widetilde{C}_{i-1}\sigma_{nn} - \widetilde{D}_{i-1}\sigma_{qn}\right) \\
\widetilde{D}_i &\equiv \widetilde{D}_1 + \widetilde{D}_{i-1}\rho_{qq}
\end{aligned}$$

(Proof is available in separate appendix).

First, note that there is no marginal pricing difference in the effect of q_t on riskless versus risky coupon streams: the expression for $\widetilde{D}_n$ is the same as D_n. 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. The implication is that increases in q_t always depress equity prices. Second, the $\widetilde{B}_1$ and $\widetilde{C}_1$ terms do differ from their consol counterparts. However, the pricing functions are still such that these coefficients are always positive. In other words, shocks to n_t and p_t that drive up the variability of cash flows, always increase the price-dividend ratio. 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. Let us maintain the assumption that $a_p < 0$ and $a_n > 0$. Because $-a_p - \sigma_{cp}$ is less positive than $-a_p$, an increase in p_t raises equity prices less than it raises real consol prices because the equity cash flow is risky. Similarly, because $-a_n + \sigma_{cn}$ is less negative than $-a_n$, equity prices rise by less than real consol prices when n_t increases. However, the risky cash flow and pure term structure effects offset one another. Again, this is only under our maintained assumption of the signs of a_p and a_n, which are in turn consistent with a counter-cyclical risk aversion process. That equity prices are so closely tied to consol prices is a quite strong restriction. Nevertheless, it is an artifact of our desire to follow the simple structure in CC, setting consumption equal to dividends and excluding time-varying cash flow expectations effects in equity pricing. We consider a simple extension in the final section that relaxes these assumptions.

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,

$$dp_t \approx d_0 + d_1' Y_t \tag{22}$$

where $Y_t = [p_t, n_t, \Delta c_t, q_t]'$ 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. Further, we can approximate logarithmic equity returns as

$$r_{t+1} \approx r_0 + r_1' Y_{t+1} + r_2' Y_t \tag{23}$$

with these results also described in detail in the appendix.

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.

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, $pvar_t$. Importantly, this computation uses the approximation in Equation (23). That approximation and Lemma 1, yield:

$$pvar_t = \left(\sigma_{pp} r_p + \sigma_{cp} r_c + \sigma_{qp} r_q\right)^2 p_t + \left(\sigma_{nn} r_n - \sigma_{cn} r_c + \sigma_{qn} r_q\right)^2 n_t \tag{24}$$

where r_p is the loading of returns onto on p_t, in Equation (23), etc. Unsurprisingly, 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:

$$\begin{aligned} psk_t &= 2\left(\sigma_{pp} r_p + \sigma_{cp} r_c + \sigma_{qp} r_q\right)^3 p_t - 2\left(\sigma_{nn} r_n + \sigma_{cn} r_c + \sigma_{qn} r_q\right)^3 n_t \\ pku_t &= 6\left(\sigma_{pp} r_p + \sigma_{cp} r_c + \sigma_{qp} r_q\right)^4 p_t + 6\left(\sigma_{nn} r_n + \sigma_{cn} r_c + \sigma_{qn} r_q\right)^4 n_t \end{aligned} \tag{25}$$

The BEGE model is therefore clearly able to generate time-varying skewness which can change sign over time as well as kurtosis which varies in magnitude. 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. Of course, the BEGE system can always be augmented with additional state variables to break this dependence.

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.

2. The wedge between risk-neutral and physical variance covaries positively with the equity risk premium.

3. Negative shocks to returns are associated with contemporaneous increases in risk-neutral variance that tend to persist.

4. The risk-neutral distribution is negatively skewed and fat tailed.[7].

[7] This is consistent with the older options pricing literature that focused on implied volatility smirks and smiles found when using the Black-Scholes option pricing model to back out implied volatilities at various strike prices.

We now examine the risk-neutral distribution of returns under the BEGE framework to see whether the framework is likely to be capable of matching the stylized facts. To facilitate the calculation of the risk-neutral distribution of returns, let us first define the risk-neutral expectation of any variable, $E_t^Q\left[\exp\left(x_{t+1}\right)\right]$,as

$$E_t^Q\left[\exp\left(x_{t+1}\right)\right] = E_t\left[\exp\left(m_{t+1} + x_{t+1}\right)\right]\left(E_t\left[\exp\left(m_{t+1}\right)\right]\right)^{-1} \tag{26}$$

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:

$$qvar_t = \left(\frac{\sigma_{pp}r_p + \sigma_{cp}r_c + \sigma_{qp}r_q}{1-a_p}\right)^2 p_t + \left(\frac{\sigma_{nn}r_n - \sigma_{cn}r_c + \sigma_{qn}r_q}{1-a_n}\right)^2 n_t \tag{27}$$

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 (24). Consider first the denominator in term multiplying p_t. Maintaining our assumption that $a_p < 0$ (that is, that positive p_t shocks lower marginal utility) 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 $a_n > 0$[8] (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, $qvar_t - pvar_t$, which we henceforth denote $vprem_t$ is simply the difference between Equations (27) and (24), and can potentially be positive. Moreover, if, as expected, increases in n_t tend to increase the equity risk premium, then the variance premium may covary positively with the equity risk premium, consistent with stylized fact 2. If n_t is persistent, then negative return shocks may coincide with higher risk-neutral variance that persists for several periods, consistent with stylized fact 3.

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

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

respectively, can be expressed as:

$$
\begin{aligned}
qsk_t &= 2\left(\frac{\sigma_{pp}r_p + \sigma_{cp}r_c + \sigma_{qp}r_q}{1-a_p}\right)^3 p_t - 2\left(\frac{\sigma_{nn}r_n - \sigma_{cn}r_c + \sigma_{qn}r_q}{1-a_n}\right)^3 n_t \\
qku_t &= 6\left(\frac{\sigma_{pp}r_p + \sigma_{cp}r_c + \sigma_{qp}r_q}{1-a_p}\right)^4 p_t + 6\left(\frac{\sigma_{pp}r_p + \sigma_{cp}r_c + \sigma_{qp}r_q}{1-a_n}\right)^4 n_t
\end{aligned}
\tag{28}
$$

By examining these expressions, we see that qsk_t will be negative when n_t is large and 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 4 .

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.

3.1 Data

The main data we use are monthly and span the period from January 1990 through March 2009. For consumption growth, Δc_t, we use real personal consumption expenditures (PCE) on nondurables and services from the Bureau of Economic Analysis (BEA). To calculate an inflation-adjusted series, we first sum the two nominal consumption series, calculate the nominal growth rate, and then deflate using the overall PCE deflator from the BEA. We estimate the real short rate, rrf_t, as the 30-day nominal T-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 month-end price. The equity return, ret_t, we use 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 use the realized and risk-neutral expected variance data provided on Hao Zhou's website, and updated through March 2009. We measure risk-neutral equity conditional variance, $qvar_t$, following Bollerslev, Tauchen and Zhou (2009) as the month-end value of the VIX, squared. We calculate the physical probability measure of equity return conditional variance, $pvar_t$, in two steps. We begin with 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: $rvar_t$, rrf_t, dp_t, and $qvar_t$.[9] 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.

Panel A of Table 1 reports some simple statistics for the monthly sample. Note that the average real return on equity for this sample is only 0.0037 per month, or about 4.4 percent per year. Given that the real short rate averaged about 1.2 percent per year, the realized average excess return on equity for the sample is only about 3.2 percent per year. The usual stylized facts are present: a low risk free rate with low volatility, a volatile dividend yield and volatile equity returns. In addition, we note the properties of the variance premium, which has a significantly positive mean. Also note that unconditional higher-order moments of consumption suggest little departure from normality: Sample skewness and kurtosis are -0.1 and 3.7 respectively, with only the latter significantly different from its value under normality. Nevertheless, when we examine the data more carefully for nonlinearities in the consumption process in the next subsection, significant time-varying departures from normality do emerge.

3.2 Empirical evidence for non-linearities in fundamentals

While the evidence of time-variation in consumption growth volatility is abundant (see Bekaert, Engstrom, Xing (2009) for a survey), there exists considerably less empirical work on higher-order moments of consumption growth. The regime switching models in Whitelaw (2000) and Bekaert and Liu (2004) do imply that US consumption exhibits time-varying skewness. For our main monthly dataset, we measure conditional higher-order consumption moments in a reduced-form fashion using asset prices as instruments. Specifically, we estimate the following system of equations:

$$
\begin{aligned}
\Delta c_{t+1} &= \overline{g} + u_t^1 \\
(\Delta c_{t+1} - \overline{g})^2 &= m_2 + x_t' \beta_2 + u_t^2 \\
(\Delta c_{t+1} - \overline{g})^3 &= m_3 + x_t' \beta_3 + u_t^3
\end{aligned} \tag{29}
$$

On the left-hand side of the bottom two equations are realized, demeaned consumption growth raised to the second and third powers. We maintain the assumption of a constant conditional mean. On the right-hand side are simple linear specifications using a vector of instruments, x_t. For

[9] This regression suggests that $cvar_t$ loads heavily onto both lagged $rvar_t$ and $qvar_t$. We cannot reject the joint hypothesis that the loadings on lagged rrf_t and dp_t are zero, but we very strongly reject the hypothesis that there is no dependence on lagged $qvar_t$.

the monthly dataset, x_t is comprised of the real short rate, rrf_t, the dividend yield, dp_t, the physical and risk-neutral equity return variance measures, $pvar_t$ and $qvar_t$, and exponentially-weighted (with parameter 0.1) moving averages of squared and cubed demeaned consumption growth. In column 1 of Table 2, the top row reports the p-value for the joint significance of β_2 and the second row reports the joint significance for β_3. We strongly reject the null hypothesis that the conditional variance and centered third moment are constant, as p-values for the joint significance of β_2 and β_3 are substantially below 0.01.

Recall that we denote $E_t\left(\Delta c_{t+1}-\overline{g}\right)^2$ by vc_t and $E_t^3\left(\Delta c_{t+1}-\overline{g}\right)^3$ by sc_t. Columns 2 through 4 of Table 2 report some univariate statistics for vc_t and sc_t, revealing significant variability and autocorrelation in both. These conditional moments also correlate in the expected manner with asset prices. The dividend yield, the physical conditional variance of returns, and the risk-neutral conditional variance of returns all vary strongly and positively with the conditional variance of fundamentals, and negatively with the conditional third moment. The signs of correlations with the real short rate follow the opposite pattern. Hence, when consumption shocks are negatively skewed, equity prices, the VIX and the conditional variance of equity returns are relatively high and real short rates are low.

Of course, our short sample period is not well suited to detect strong non-linearities in consumption growth. For example, relaxing the restriction of a constant conditional mean weakens the evidence in Table 2 for time-varying skewness. We nevertheless believe that the evidence for these non-linearities is strong. In Section 6, we consider a longer sample using data going back to the Great Depression, to estimate consumption moments. In the conclusion, we show how such non-linearities are more apparent in survey data reflecting expectations of economic conditions. If anything, the estimation conducted here will underestimate the importance of consumption growth non-linearities.

4 Structural Model Estimation

In this section, we outline our estimation strategy for the structural model. We use classical minimum distance (CMD) for estimation, which relies on the matching of sample statistics.[10]

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

4.1 Reduced form statistics to be matched

We begin by calculating a vector of sample statistics, $\widehat{p}$, with estimated covariance matrix $\widehat{V}$ to be matched by the structural model. For $\widehat{p}$, we use all the statistics reported in Table 1 and Panel A of Table 2. In doing so, we ask the model to match the conditional means, volatilities and autocorrelations of consumption growth, Δc_t, the real short rate, rrf_t, the dividend yield, dp_t, real equity returns, ret_t, the conditional variance of returns under the physical and risk-neutral measures, $pvar_t$ and $qvar_t$ respectively, and the conditional second and third centered moments of consumption growth, vc_t and sc_t respectively.[11] Further, we require that the model match the unconditional sample skewness and kurtosis of consumption growth. We also seek to fit the unconditional correlation between changes in $pvar_t$ and the variance premium, $vprem_t \equiv qvar_t$-$pvar_t$. We find that this statistic is useful in helping to identify the correlation between risk aversion, q_t, and the p_t and n_t processes more precisely. In all, we ask the model to match 26 reduced-form statistics. By any measure, this represents an extremely challenging set of moments for a relatively parsimonious structural model. We use a heteroskedasticity and autocorrelation consistent (HAC) estimator for $\widehat{V}$ employing the Newey-West (1987) methodology with 20 Newey-West lags. The sample statistics are related to the population statistics, p_0, by

$$\sqrt{T}\left(\widehat{p}-p_0\right) \sim N\left(0, \widehat{V}\right). \tag{30}$$

4.2 Objective function and distribution of structural parameters

Under the model to be estimated, the sample statistics of the endogenous variables are nonlinear functions of the deep model parameters. The mapping is described in the appendix. We denote the true structural parameters by the vector, θ_0. The structural parameters to be estimated are,

$$\theta = \left[\overline{g}, \sigma_{cp}, \sigma_{cn}, \overline{p}, \rho_p, \sigma_{pp}, \overline{n}, \rho_n, \sigma_{nn}, \overline{q}, \rho_q, \sigma_{qp}, \sigma_{qn}, \ln\left(\beta\right), \gamma\right]' \tag{31}$$

Under the null hypothesis that our model is true,

$$p_0 = h\left(\theta_0\right) \tag{32}$$

[11] We do not attempt to match the consumption growth autocorrelation, which our model implicitly fixes at 0.

where $h(\theta)$ is a vector-valued function that maps the structural parameters into the reduced-form statistics. To form estimates of the structural parameters, $\widehat{\theta}$, we minimize an objective function of the form,

$$\min_{\theta \in \Theta} \{\widehat{p} - h(\theta)\}' \widehat{W}^{-1} \{\widehat{p} - h(\theta)\} \tag{33}$$

where $\widehat{W}^{-1}$ is a symmetric, positive semi-definite, data-based weighting matrix. Efficient CMD suggests $\widehat{V}^{-1}$ for the weighting matrix, but we instead use a diagonal weighting matrix , $\widehat{W} = diag\left(\widehat{V}\right)^{-1}$. We do this because because vc_t and sc_t are very nearly exact linear combinations of the other variables,[12] rendering $\widehat{V}$ nearly singular.

Standard CMD arguments lead to the asymptotic distribution of $\widehat{\theta}$ and a test of the overidentifying restrictions (see appendix).

5 Results

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

5.1 Model estimation results

We only estimate 13 of the 15 parameters listed above in θ because we fix two parameters ex-ante. First, because the scale of the latent factor q_t is not well identified using our set of reduced-form parameters, 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 $\ln(\beta) = -0.0003$ 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 on the remaining parameters' estimates. Of the three state variable process, n_t and q_t are highly persistent, whereas p_t's autocorrelation coefficient is only 0.6. Of particular interest are the parameters σ_{qp} and σ_{qn} which govern the correlation between consumption shocks and risk aversion. As expected, positive "good environment" consumption shocks reduce risk aversion, but positive "bad environment" shocks lead to higher risk aversion. Both coefficients are significantly different from zero.

Note that the test of the over-identifying restrictions rejects at the 1 percent level, but the model

[12] Because vc_t and sc_t are spanned in part by lagged (exponentially-weighted) moving averages of squared and cubed consumption growth in addition to the other instruments, there is no exact dependence with the other variabels used in estimation. However, in practice the regression places very low weights on these variables, so that vc_t and sc_t are almost perfecetly linearly dependent on the other variables.

does have an overall satisfactory fit with the moments used in the estimation. To make this more concrete, Table 4 compares some basic moments for a number of critical variables in the model with the data. The model moments are in square brackets above the data moment; the number in parentheses is a data-based standard error.

Let's first focus on the fitted consumption growth statistics. The fit is nearly perfect. Not only do we fit the mean and volatility exactly, we also nearly perfectly fit the near-zero skewness and mild kurtosis of consumption growth. Of course, the autocorrelation of consumption growth in the model is by definition zero, whereas the monthly data show slight negative autocorrelation. In Panel B, we also look at the conditional variance and centered third moment of consumption growth, vc_t and sc_t respectively, and the model fits the first three moments of vc_t near perfectly, but has trouble matching the volatility of sc_t.

For the real short rate, the dividend yield and equity returns, we also match the first three moments, producing moments comfortably within one standard error of the data moment. Hence, the model fits the standard moments that are the focus of articles such as Bansal and Yaron (2004) and Campbell and Cochrane (1999). However, the model generates a correlation between equity returns and consumption growth of 0.7, while that moment in the data is only 0.2, estimated with a standard error of 0.1. While the model-implied correlation is thus too high, it is lower than the correlation implied by some other popular consumption-based models (for instance, Campbell and Cochrane (1999)). If we add this statistic to the set being matched during estimation, we find that we can lower this correlation somewhat without dramatically worsening the fit elsewhere. Moreover, the model extension we propose in Section 6 can easily break the strong correlation by introducing a dividend process that is not perfectly correlated with consumption.

Finally, we report some characteristics of the conditional variance of equity returns and the variance premium. While the model generates a good fit for the mean of the physical volatility of returns and the variance premium, the volatility of the physical volatility of returns is somewhat too low. In section 6, we show how this miss owes to the mild consumption data we have used in the study. To preview those results: when we taker a longer view of consumption growth dynamics, not surprisingly we find stronger nonlinearities in consumption. If we then allow the model to "see" the stronger consumption dynamics, the estimation procedure can then match all the moments of $pvar_t$ and $vprem_t$ almost perfectly.

5.2 The conditional distribution of consumption growth

We now examine the dynamics of the conditional distribution of consumption growth in more detail. The mean of p_t is estimated at around 26. At this value, shocks to ω_{pt} are fairly close to being normally distributed. In contrast, n_t has a very low mean of about 0.06, suggesting a strongly nonlinear distribution of $\omega_{n,t}$ shocks on average.[13] However, the mean contribution of the $\omega_{p,t}$ shocks to the consumption growth variance is $\left(\sigma_{cp}^2\overline{p}\right)$ is an order of magnitude larger then the contribution of $\omega_{n,t}$ shocks, $\left(\sigma_{np}^2\overline{n}\right)$. 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 2 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 logarithms of the densities are plotted. The top left panel shows that when n_t and p_t are at their 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 (holding the variance constant). 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 moderately 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. In summary, at the point estimates presented in Table 2, p_t basically serves to govern the overall variance of the distribution of consumption growth and the thickness of the positive tail, while n_t determines the size of the negative tail with less of an impact on overall consumption growth variance.

[13] For a $\Gamma(26,1)$ random variable, skewness is $2/\sqrt{26} \sim .4$ and excess kurtosis is $6/26 \sim 0.2$. For a $\Gamma(0.06,1)$ random variable, skewness is about 8 and excess kurtosis is about 33.

5.3 The dynamics of asset prices

Table 5 reports the dependence of the various endogenous variables on the state vector. We first focus on Panel A, which reports the factor loadings on the three state variables (p_t, n_t and q_t). Not surprisingly, positive shocks to p_t and n_t lower real interest rates through precautionary savings effects, while a positive shock to q_t increases the interest rate through a consumption smoothing effect. These effects are also present with the same sign for the dividend yield. This parity arises because our model lacks interesting equity cash flow dynamics—the main effects of the state variables for all long-lived assets work through the term structure. Positive shocks to n_t increase the equity premium, $eqprem_t$, with the effect of the other variables being negligible. The conditional variance of equity returns is increasing in all three state variables, but the variance premium only loads positively on n_t. It is shocks to n_t that should cause a positive correlation between the equity risk premium and the variance premium.

Figure 3 plots impulse response functions of some of the asset prices to $\omega_{p,t}$ and $\omega_{n,t}$. Recall that q_t is spanned by the two fundamental consumption shocks. Hence, a positive $\omega_{p,t}$ shock not only increases p_t but also decreases q_t. Consequently, the effect of $\omega_{p,t}$ on interest rates is negative. For $\omega_{n,t}$ shocks, increases in risk aversion are so severe that the desire of investors to borrow to smooth consumption dominates and short rates rise. Both shocks increase the conditional variance of equity returns but the effect of an $\omega_{n,t}$ shock dies out much more slowly than that of an $\omega_{p,t}$ shock. Finally, the variance premium persistently increases with an $\omega_{n,t}$ shock, and decreases slightly with an $\omega_{p,t}$ shock.

5.4 Endogenous predictability

Much of the asset pricing literature focuses on equity return predictability. Nevertheless, as we stress again, the return predictability evidence is rather weak. In Table 6, Panel A, we present some univariate statistics for regressions of excess equity returns on the short rate, the dividend yield and the variance premium. Neither the short rate, dividend yield nor the variance premium are significant predictors of future stock returns. The short rate in fact is the strongest predictor. 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 is simply uses $rvar_t$ the measure of conditional variance, $pvar_t$. In contrast, we use a projection of $rvar_t$ onto lagged several variables to identify $pvar_t$. The last column of Panel B show that the variance premium measured as in Bollerslev, Tauchen and Zhou (2009) indeed significantly predicts equity returns for our sample. Our structural model generates a modest amount of return predictability. We report the model-implied projection coefficients in brackets above the sample coefficients. All the signs are the same, but the magnitudes are somewhat smaller than in the data.

Panel B reports the expression for the equity premium in terms of the fundamental state variables and the model implied R^2, which is very modest at 25 basis points. Given the lack of strong predictability in the data, this would appear to be realistic. Nevertheless, the conditional Sharpe ratio for equity, the ratio of the conditional expected excess return to the conditional volatility, does vary substantially through time. Figure 4 plots the Sharpe ratio as a function of n_t and p_t. The Sharpe ratio is not very sensitive to p_t, and mostly remains well below an annualized 28 percent. However, the conditional Sharpe ratio is very sensitive to shocks to n_t and can become as high as 45% when n_t exceeds 0.13, about twice its unconditional mean. 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 and rare occurrence of return predictability (see Henkel, Martin and Nardari (2009)).

5.5 Higher order risk-neutral return moments

We have already shown in Table 4 that our model generates a positive variance risk premium, perhaps the most celebrated stylized fact about the risk-neutral distribution of equity returns. In Table 7, Panel A, we report some descriptive statistics for the higher order moments as well. These are the return distribution statistics under the model when the state vector is at its unconditional mean. Note that none of these moments were fit as part of the estimation. Moreover, the sample data to which we compare our model's implications are estimated by Figlewski (2009) from another data source. Figlewski uses option price data to empirically identify the complete risk-neutral distribution of returns for the S&P 500 over a time span similar to ours.[14] The model's implied risk-neutral skewness and kurtosis both suggest unrealistically large departures from normality when the state vector is at its unconditional mean. It is conceivable that this poor fit arises because the model tries to simultaneously explain quite benign consumption growth data and fairly dramatic asset price movements. We revisit these statistics when we re-estimate the BEGE model using alternative consumption statistics that are based on a longer consumption sample in Section 6.1.

[14] However, Figlewski uses 90-day options whereas we model 30-day options. We ignore the potential difference implied by the maturity difference for risk-neutral skewness and kurtosis.

In panel B, we report the correlation of changes in the risk neutral variance of returns with realized equity returns. The contemporaneous correlation in the data is significantly negative, which is matched quite well by the model. Further, because the change in risk-neutral variance is persistent, returns do not significantly forecast any subsequent changes in risk-neutral volatility. This feature of the data is also well-matched by the model.

Overall, Table 7 suggests a good fit between the BEGE model and the most salient stylized facts about the risk-neutral distribution of returns from the options pricing literature.

6 Robustness checks and model extensions

In this section, we first consider the problem that while sample only starts in 1990 because of the availability of the VIX data, consumption nonlinearities, the heart of the BEGE model, are much more evident in earlier time periods. Then, we describe a relatively straightforward extension to the current model that may further improve the fit with the data along some dimensions that are not the primary focus of this article.

6.1 A Longer-term perspective on consumption nonlinearities

Our main monthly data set, which extends from January 1990 through March 2009, covers a relatively mild period for consumption growth. Even the last twelve months of consumption growth, in the thick of the financial crisis of 2008 and 2009, show consumption falling only by about 4 basis points per month on average with volatility for the last 12 months of 33 basis points – just a bit higher than the overall sample volatility. Meanwhile, the upheaval in asset prices in 2008 and 2009 is more reminiscent of return dynamics during the Great Depression. Of course, it is possible that asset prices are simply foretelling more dramatic consumption dynamics (yet to come). It is reasonable, however, to ask whether our model results would differ materially if we instead took a longer view of consumption dynamics (it is not possible to examine asset price dynamics used in this paper over a longer sample given the limited availability of the VIX). For example, investors may have long placed some probability, albeit small, on the return of a regime like the Great Depression, but that belief is surely not represented by the statistics about vc_t and sc_t reported in Table 2 since they are based on very modest consumption dynamics exhibited in the 1990's. It follows that the preference parameters we estimate in Table 3 may also not be representative of investors' true preferences. This might also explain why the model has some trouble generating enough volatility in

$pvar_t$ under the physical measure (see Table 3) and why the model-implied skewness and kurtosis of the risk-neutral return density is so extreme in Table 7. To explore this issue, we first characterize consumption dynamics over a much longer time-span that encompasses the Great Depression. We then inject these consumption dynamics into our framework, and ask whether our structural model estimates differ materially from our main estimation.[15]

The monthly consumption source data used in this study extends back only to 1959. However, annual consumption data is available back to 1929 from the BEA in the NIPA accounts. To estimate monthly consumption dynamics back to the Great Depression era, we must interpolate intra-year consumption growth using another data source. The appendix shows how we use a bootstrapping procedure to sample from monthly consumption dynamics back to 1926 using the monthly growth rate of industrial production (which is available back to 1919) as an instrument. Based on these draws, we calculate bootstrapped statistics for Δc_t, vc_t, and sc_t in the same manner as we did for the short sample.

The median outcome for these statistics and standard errors over 10,000 draws are reported in Panel A of Table 8. The unconditional sample statistics for consumption growth are not too different from those for the short sample reported in Table 2, except that, not surprisingly, the volatility of consumption growth is higher in the longer sample. However, the properties of vc_t and sc_t are much more extreme. The column labeled *pvals* reports the median p-value for the significance of the regressions estimating vc_t and sc_t. The mean and volatility of vc_t are about three times higher for the long sample than the short one. For sc_t we find a much more negative unconditional mean and nearly fives times as much volatility. Figure 5 plots the median draw of vc_t and sc_t for the full sample. Not surprisingly, the more extreme consumption dynamics arise from the inclusion of the Great Depression in the long sample. However, the recent values taken on by vc_t and sc_t are more dramatic than any other economic downturn since the 1930's.

In Panel B of Table 8, we report results for the structural parameter estimates once we have replaced the sample statistics for Δc_t, vc_t and sc_t for those reported in Panel A of Table 8 using the long sample. All the other sample statistics to be matched remain the same (as reported in Tables 1 and 2). The structural model parameters are qualitatively similar to those in Table 3. In particular, the q_t dynamics are quite similar. Moreover, we find that p_t still has a large mean, indicating that $\omega_{p,t}$ shocks are typically quite Gaussian. However, n_t has significantly higher variance under the

[15] This is similar in spirit to the efforts of Barro, Nakamura, Steinsson and Ursua (2009) to obtain better estimates for the fundamentals of a rare disasters model using a large panel of cross-country data.

new parameters so that it is a more important driver of asset prices, and its mean is larger, suggesting it features less severe departures from normality than it did in the main estimation Unfortunately, we no longer find $\sigma_{qp} < 0$, suggesting that positive consumption shocks sometimes do not reduce risk aversion, which is inconsistent with the notion of habit. However, σ_{qn} is still quite large and positive, as before. The model retains the ability to fit all the asset price data quite closely and we do not report detailed statistics. Notably, with the more dramatic consumption dynamics, the model is now able to match the volatility of $pvar_t$ much more closely. One notable failure of this estimation is that the model does not match the correlation between changes in $pvar_t$ and $vprem_t$. This is occurring because n_t is the overwhelming driver of both $pvar_t$ and $vprem_t$. In Panel A of Table 8, we also report the model statistics for consumption dynamics. The model also fits the sc_t statistics somewhat more closely, but it still cannot generate enough volatility in sc_t. The other characteristics of vc_t and sc_t are fit near-perfectly.

For brevity, we do not reproduce the full set of model-implied dynamics analysis as we did for the main model in Tables 5 and 6. The results for the alternative estimation are quite similar. However, we do report the model's implications for the risk-neutral density of returns under the alternative estimation in Table 7. Note that under the alternate estimation, the model generates more modest mean risk-neutral conditional skewness and kurtosis of returns of -3.5 and 27.5 respectively. These values are quite close those reported by Figlewski (2009). However, the model does generate too much (negative) correlation between returns and changes in $qvar_t$ in the alternate estimation.

In summary, some of the few unrealistic features of the BEGE model reported for the main estimation are ameliorated when taking a longer view of consumption dynamics. In particular, the model matches option price data more closely. This is indirect evidence that the Great Depression and other periods of severe economic stress leave a lasting imprint on asset prices.

6.2 Model Extension

In presenting the BEGE model, we tried to stay as close as possible to the set-up in Campbell and Cochrane (1999), but introduced nonlinearities to allow the model to better fit option price dynamics. While the model is clearly successful in that dimension, it is too restrictive to match other salient features of the equity and risk free rate data. Specifically, we did not allow for conditional mean dynamics in the consumption growth process, and to model equity prices we priced a consumption claim as opposed to modelling equity dividends. This makes it harder to generate "flight-to-safety" effects, where bad consumption shocks cause interest rates to drop through a precautionary savings effect, while simultaneously making equities riskier and decreasing equity prices. With the current specification, we have essentially precluded the latter channel, as our equity claim is a claim to consumption, and there are no intricate cash flow dynamics present in the model.

It is rather straightforward to incorporate a more realistic dividend process, as shown by the following example. Instead of assuming that dividends equal consumption, assume that the logarithmic dividend-consumption ratio depends on p_t and n_t:

$$d_t - c_t = \overline{dc} + \kappa_{dp} p_t + \kappa_{dn} n_t \tag{34}$$

where d_t is the log level of dividends, and $\overline{dc}$, κ_{dp}, and κ_{dn} are parameters. Clearly, the dividend-consumption ratio is stationary under this specification, but may vary over the business cycle. The following lemma describes equity prices with this extension

For the economy described by Equations (1) through (11), and (34) the price-dividend ratio of equity is given by

$$\frac{P_t}{D_t} = \sum_{n=1}^{\infty} \exp\left(\widehat{A}_i + \widehat{B}_i p_t + \widehat{C}_i n_t + \widehat{D}_i q_t\right) \tag{35}$$

where the initial values of the parameter sequence are given by

$$\begin{aligned}
\widehat{A}_1 &= [\kappa_{dp}\overline{p} + \kappa_{dn}\overline{n}] + \ln\beta + (1-\gamma)\,\overline{g} + \gamma\left(1-\rho_q\right)\overline{q} \\
\widehat{B}_1 &= \left[\kappa_{dp}\left(\rho_p - 1 - \sigma_{pp}\right)\right] - a_p - \sigma_{cp} - \ln\left(1 - [\kappa_{dp}\sigma_{pp}] - a_p - \sigma_{cp}\right) \\
\widehat{C}_1 &= [\kappa_{dn}\left(\rho_n - 1 - \sigma_{nn}\right)] - a_n + \sigma_{cn} - \ln\left(1 - [\kappa_{dn}\sigma_{nn}] - a_n + \sigma_{cn}\right) \\
\widehat{D}_1 &= \gamma\left(\rho_{qq} - 1\right)
\end{aligned}$$

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

$$\begin{aligned}
\widehat{A}_i &= \widehat{A}_{i-1} + \widehat{B}_{i-1}\mu_p + \widehat{C}_{i-1}\mu_c + \widehat{D}_{i-1}\mu_q \\
\widehat{B}_i &\equiv \left(\left[\kappa_{dp}\left(\rho_p - 1 - \sigma_{pp}\right)\right] - a_p - \sigma_{cp} + \widehat{B}_{i-1}\left(\rho_p - \sigma_{pp}\right) - \widehat{D}_{i-1}\sigma_{qp}\right) \\
&\quad - \ln\left(1 - [\kappa_{dp}\sigma_{pp}] - a_p - \sigma_{cp} - \widehat{B}_{i-1}\sigma_{pp} - \widehat{D}_{i-1}\sigma_{qp}\right) \\
\widehat{C}_i &\equiv \left([\kappa_{dn}\left(\rho_n - 1 - \sigma_{nn}\right)] - a_n + \sigma_{cn} + \widehat{C}_{i-1}\left(\rho_n - \sigma_{nn}\right) - \widehat{D}_{i-1}\sigma_{qn}\right) \\
&\quad - \ln\left(1 - [\kappa_{dn}\sigma_{nn}] - a_n + \sigma_{cp} - \widehat{C}_{i-1}\sigma_{nn} - \widehat{D}_{i-1}\sigma_{qn}\right) \\
\widehat{D}_i &\equiv \widehat{D}_1 + \widehat{D}_{i-1}\rho_{qq}
\end{aligned}$$

The terms that are new relative to the equity pricing result in Section 2 are highlighted in brackets. They reflect pure cash-flow effects, and to the extent that p_t and n_t affect cash-flow expectations, they will drive a wedge between equity prices and the price of the real consol. We defer estimating such a model to future work.

7 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. In this paper, we appended the BEGE technology to the well-known Campbell–Cochrane (1999) 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 a lot of recent attention.

We do not want to propose the particular model explored in this paper as the new paradigm. Many realistic features are missing. The recent crisis reinforces the potential importance of Knightian uncertainty (see Drechsler (2009) and Epstein and Schneider (2007) for recent efforts) and learning (see Veronesi (1999) for example) 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. In particular, a very useful extension of our model would be to add a time –varying mean to the consumption growth process as in Bansal and Yaron (2004). The main advantage of such a model is that it allows expectations about the future state of the economy to be priced in financial market data. The current crisis again shows that anticipation of future bad economic conditions has marked implications on asset prices, yet, in our Campbell-Cochrane specification, fundamentals are only driven by ex-post shocks. 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 as in Bansal and Yaron (2004).

Moreover, the sample used in this article only witnessed a few mild recessions, with the current crisis likely not yet fully reflected in the data. A richer picture of the distribution of economic conditions can be gleaned using longer consumption growth data as is in the previous section, or from contemporaneous survey data. From the Survey of Professional Forecasters we can estimate the entire conditional distribution of real GDP growth. To do so, we combine information about probabilities from the survey with long-term data on GDP growth to compute the first three uncentered moments of real GDP growth (see the appendix for more details). Shaliastovich (2009) uses similar data to model expected consumption growth in a long-run risk model. Figure 6 plots the centered conditional moments of fundamentals growth based on the survey data using the methodology described in the appendix. While they refer to GDP, it is likely that the conditional distribution of real consumption growth follows similar patterns. The top panel plots the time series for the conditional mean of GDP growth. While the conditional mean typically fluctuates in a narrow band between 2 and 4 percent, low expected growth is evident around the recessions in the early 1990s, early 2000s and in early 2008. Moreover, in an exercise similar to that conducted for Table 2, we project these conditional moments on asset prices (the dividend yield, VIX, etc). The regressions overwhelmingly reject the null that there is no dependence between of the conditional second and third moments of GDP growth and asset prices. The middle panels plot the time-series of the conditional variance and volatility of growth. The decline in volatility previously referred to as the Great Moderation from the early 1980's through 2007 is clearly evident. However, the recent spike in volatility is near the all-time high for the series. The bottom two panels plot the uncentered third conditional moment of growth and conditional skewness. The conditional skewness plot shows interesting variation, with long periods of both positive and negative skewness. In particular, positive skewness which emerged in the early 2000s has given way to deeply negative skewness in 2008 and 2009. Overall, we interpret these results as consistent with strong time-variation in the higher-order moments of fundamentals growth over the business cycle. This is exactly the kind of variation we hope to capture with he BEGE model developed in this article. In the future, we hope to incorporate the survey data into a similar BEGE framework.

Our work is related to but quite different from Drechsler and Yaron (2009) and Bollerslev, Tauchen and Zhou (2009). Both articles feature equilibrium economies to attempt to explain the variance premium and its dynamics. Drechsler and Yaron essentially add jumps to the consumption growth technology in Bansal and Yaron (2004), whereas Bollerslev, Tauchen and Zhou introduce stochastic volatility of volatility of consumption growth in an Epstein-Zin (1989) framework. Neither article estimates structural parameters or comes as close as the BEGE model to fitting such a wide set of stylized facts.

8 Appendix

8.1 The General Model

We can write our model is general terms as follows

$$Y_{t+1} = \mu + AY_t + \Sigma_H \varepsilon_{t+1} + \Sigma_F \omega_{t+1} \tag{36}$$

Where $Y_t(nx1)$ is the state vector, $\mu(nx1)$ is the associated mean parameter vector, $A(nxn)$ is a transition parameter matrix, $\Sigma_H(nxq)$ is the conditional volatility matrix for normally-distributed shocks, $\varepsilon_{t+1}(qx1)$, and $\Sigma_F(nxp)$ is the conditional volatility matrix for the gamma-distributed shocks, $\omega_{t+1}(px1)$. The specific distributional assumptions for the shocks are

$$\begin{aligned} \varepsilon^i_{t+1} &\sim N(0,1),\ i = 1, ..., q \\ \omega^i_{t+1} &\sim \Gamma\left(k^i_t, 1\right) - k^i_t,\ k_t = \Phi Y_t,\ i = 1, ..., p \end{aligned} \tag{37}$$

and all the shocks are independent. The additive term in the ω^i_{t+1} definition, $-k^i_t$, sets the mean of the shock to zero. The parameter matrix $\Phi(pxn)$ is comprised of only zeros and ones and selects which elements of Y_t determine the "shape" parameter of each ω^i_{t+1} shock.[16] Let v be a $(nx1)$ parameter vector. For our main model, $Y_t = [p_t, n_t, \Delta c_t, q_t]'$, and the system matrices are:

$$\begin{aligned} \mu &= \left[\left(1-\rho_p\right)\overline{p}, \left(1-\rho_n\right)\overline{n}, \overline{g}, \left(1-\rho_q\right)\overline{q}\right]' \\ A &= diag\left(\left[\rho_p, \rho_n, 0, \rho_q\right]\right), \Sigma_H = 0 \\ \Sigma_F &= \begin{bmatrix} \sigma_{pp} & 0 \\ 0 & \sigma_{nn} \\ \sigma_{cp} & -\sigma_{cn} \\ \sigma_{qp} & \sigma_{qn} \end{bmatrix}, \Phi = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{aligned} \tag{38}$$

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

Lemma 1 *For the random variable Y_t in Equation (36) 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*

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

The physical measures of the expectation and variance of $v'Y_{t+1}$ are defined, respectively, as

$$\begin{aligned} &\frac{d}{ds}\left[E_t\left[\exp\left(sv'Y_{t+1}\right)\right]\right]_{s=0} \\ &\frac{d^2}{ds^2}\left[E_t\left[\exp\left(sv'Y_{t+1}\right)\right]\right]_{s=0} \end{aligned} \tag{39}$$

And are given by:

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

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

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

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

Lemma 2 *For the random variable Y_t in Equation (36), and a real pricing kernel, m_t, that is affine in current and lagged values of Y_t:*

$$m_t = m_0 + m_1'Y_t + m_2'Y_{t-1}, \tag{40}$$

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

$$E_t^Q\left[\exp\left(v'Y_{t+1}\right)\right] \equiv E_t\left[\exp\left(m_{t+1}\right)\right]^{-1} E_t\left[\exp\left(m_{t+1} + v'Y_{t+1}\right)\right] \tag{41}$$

and is given, using Lemma 1, by

$$\begin{aligned} & E_t^Q\left[\exp\left(v'Y_{t+1}\right)\right] \\ = \; & \exp\left(v'\mu + v'AY_t + \frac{1}{2}v'\Sigma_H\Sigma_H'v + m_1'\Sigma_H\Sigma_H'v - \left(v'\Sigma_F + \ln\left(\mathbf{1} - \cdot\frac{v'\Sigma_F}{1 - m_1'\Sigma_F}\right)\right)\Phi Y_t\right) \end{aligned} \tag{42}$$

where $\cdot\frac{a}{b}$ 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

$$\begin{aligned} & \frac{d}{ds}\left[E_t^Q\left[\exp\left(sv'Y_{t+1}\right)\right]\right]_{s=0} \\ & \frac{d^2}{ds^2}\left[E_t^Q\left[\exp\left(sv'Y_{t+1}\right)\right]\right]_{s=0} \end{aligned} \tag{43}$$

Upon evaluation, these reduce to:

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

For the third and fourth centered moments, straightforward calculations yield,

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

8.2 Unconditional moments of the state vector and endogenous variables

To calculate the unconditional moments of Y_t, we proceed as follows using the law of iterated expectations,

$$E\left[Y_{t+1}^n\right] = E\left[E_t\left[e'Y_{t+1}^n\right]\right] \tag{44}$$

where e is a vector selecting the appropriate element of Y_t. The inner expectation can be solved by recalling that Lemma 1 provides the moment-generating function for elements of Y_t. That is, by evaluating derivatives of $E_t\left[\exp\left(tY_{t+1}\right)\right]$:

$$E_t\left[Y_{t+1}^n\right] = \frac{\partial^n}{\partial m^n} E_t\left[\exp\left(me'Y_{t+1}\right)\right]|_{m=0} \tag{45}$$

Brute force algebra yields,

$$\begin{aligned}
E_t\left[e'Y_{t+1}^1\right] &= e'\mu + e'AY_t \\
E_t\left[e'\left(Y_{t+1} - E_tY_{t+1}\right)^2\right] &= e'\Sigma_H\Sigma_H'e + \left(e'\Sigma_F\right)^{\cdot 2}\Phi_FY_t \equiv Ve_t^2 \\
E_t\left[e'\left(Y_{t+1} - E_tY_{t+1}\right)^3\right] &= 2\left(e'\Sigma_F\right)^{\cdot 3}\Phi_FY_t \\
E_t\left[e'\left(Y_{t+1} - E_tY_{t+1}\right)^4\right] - 3Ve_t^2 &= \left(e'\Sigma_F\right)^{\cdot 4}\Phi_FY_t
\end{aligned} \tag{46}$$

all of which are linear in the state vector. To calculate unconditional moments, we simple condition down, replacing Y_t with $\overline{Y}$ in the above equations. All the asset prices and other endogenous variables in the model are linear functions of Y_t This is trivially true for Δc_t, and be seen in equations 4,16,22,53,24, and 27 for the other variables. For any endogenous variable affine in Y_t, it's unconditional moments follow trivially from the above equation. It follows that the unconditional moments of all the endogenous variables are nonlinear functions of the deep model parameters.

8.3 Log Linear Approximation of Equity Prices

In the estimation, we use a linear approximation to the price-dividend ratio. From Equation (**??**), we see that the price dividend ratio is given by

$$\begin{aligned}
\frac{P_t}{D_t} &= \sum_{i=1}^{\infty} q_{i,t}^0 \\
&= \sum_{i=1}^{\infty} \exp\left(b_i^0 + b_i'Y_t\right)
\end{aligned} \tag{47}$$

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

$$\overline{q}_i^0 = \exp\left(b_i^0 + b_i'\overline{Y}\right) \tag{48}$$

and note that

$$\frac{\partial}{\partial Y_t}\left(\sum_{i=1}^{\infty} q_{i,t}^0\right) = \sum_{i=1}^{\infty} \frac{\partial}{\partial Y_t} q_{i,t}^0 = \sum_{i=1}^{\infty} q_{i,t}^0 \cdot b_i' \tag{49}$$

Approximating,

$$\begin{aligned}
pd_t &\simeq \ln\left(\sum_{i=1}^{\infty} \overline{q}_i^0\right) + \frac{1}{\sum_{i=1}^{\infty} \overline{q}_i^0}\left(\sum_{i=1}^{\infty} \overline{q}_i^0 \cdot b_i'\right)\left(Y_t - \overline{Y}\right) \\
&= d_0 + d_1'Y_t
\end{aligned} \tag{50}$$

where d_0 and d' are implicitly defined. Similarly,

$$
\begin{aligned}
gpd_t \equiv \ln\left(1+\frac{P_t}{D_t}\right) &\simeq \ln\left(1+\sum_{i=1}^{\infty}\overline{q}_i^0\right)+\frac{1}{1+\sum_{i=1}^{\infty}\overline{q}_i^0}\left(\sum_{n=1}^{\infty}\overline{q}_i^0\cdot b_i'\right)\left(Y_t-\overline{Y}\right) \\
&= h_0+h_1'Y_t
\end{aligned} \tag{51}
$$

where h_0 and h' are implicitly defined. Note also that the dividend yield measure used in this study can be expressed as follows

$$
dp_t \equiv \ln\left(1+\frac{D_t}{P_t}\right) = gpd_t - pd_t \tag{52}
$$

so that it is also linear in the state vector under these approximations. Also, log equity returns can be represented follows. Using the definition of excess equity returns,

$$
\begin{aligned}
ret_{t+1} &= -pd_t+\Delta c_{t+1}+gpd_{t+1} \\
&\sim (h_0-d_0)+(e_c'+h_1')\,Y_{t+1}-d_1'Y_t \\
&= r_0+r_1'Y_{t+1}+r_2'Y_t
\end{aligned} \tag{53}
$$

where r_0, r_1' and r_2' are implicitly defined.

8.4 CMD Asymptotics

First note that the first order condition for our optimization is,

$$
\widehat{H}'\widehat{W}^{-1}\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}=0. \tag{54}
$$

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

$$
h\left(\widehat{\theta}\right)=h\left(\theta_0\right)+H_0\left(\widehat{\theta}-\theta_0\right). \tag{55}
$$

where $H_0=\nabla_\theta h\left(\theta_0\right)$ is the gradient of $h\left(\theta\right)$ at the true parameter value. Combining Equations (54) and (55), we have,

$$
\sqrt{T}H_0'\widehat{W}^{-1}H_0\left(\widehat{\theta}-\theta_0\right)=\sqrt{T}H_0'\widehat{W}^{-1}\left(\widehat{\theta}-p_0\right) \tag{56}
$$

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

$$
\sqrt{T}\left(\widehat{\theta}-\theta_0\right)\sim N\left(0,\widehat{V_\theta}\right) \tag{57}
$$

where $\widehat{V_\theta}=\left(\widehat{M}^{-1}\widehat{H}'\widehat{W}^{-1}\widehat{V}\widehat{W}^{-1}\widehat{H}\widehat{M}^{-1}\right)$, and $\widehat{M}=\widehat{H}'\widehat{W}^{-1}\widehat{H}$.

8.4.1 Overidentification Test

Under efficient CMD, a simple overidentification test is available,

$$
T\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}\widehat{V}^{-1}\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}\sim\chi^2_{ns-np} \tag{58}
$$

where ns and $n\theta$ 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

$$
\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}\widehat{W}^{-1}\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\}, \tag{59}
$$

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

$$\sqrt{T}\left\{\widehat{p}-h\left(\widehat{\theta}\right)\right\} = \sqrt{T}\widehat{p}-\sqrt{T}\left(h\left(\theta_0\right)+H_0\left(\widehat{\theta}-\theta_0\right)\right) \tag{60}$$

$$= \sqrt{T}\left(I-H_0\left(H_0'\widehat{W}^{-1}H_0\right)^{-1}H_0'\widehat{W}^{-1}\right)\left(\widehat{p}-p_0\right) \tag{61}$$

substitution into the objective function and rearrangement yields,

$$T\cdot Obj = \sqrt{T}\left(\widehat{p}-p_0\right)'\left(\widehat{W}^{-1}-\widehat{W}^{-1}H_0\left(H_0'\widehat{W}^{-1}H_0\right)^{-1}H_0'\widehat{W}^{-1}\right)\sqrt{T}\left(\widehat{p}-p_0\right) \tag{62}$$

$$= Z'\left(\widehat{W}^{-1}-\widehat{W}^{-1}H_0\left(H_0'\widehat{W}^{-1}H_0\right)^{-1}H_0'\widehat{W}^{-1}\right)Z \tag{63}$$

where Z is an ns dimensional random vector of normal distribution 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 we $\widehat{V}$ and $z\sim N\left(0,I\right)$, we obtain,

$$T\cdot Obj = z'Az \tag{64}$$

where $A=\widehat{V}^{1/2}\widehat{W}^{-1/2}\left(I-\widehat{W}^{-1/2\prime}\widehat{H}\left(\widehat{H}'\widehat{W}^{-1}\widehat{H}\right)^{-1}\widehat{H}'\widehat{W}^{-1/2}\right)\widehat{W}^{-1/2\prime}\widehat{V}^{1/2\prime}$. JW show that A has $(np-ns)$ positive eigenvalues. Moreover, $z'Az$ is easily simulated to derive critical values for $T\cdot Obj$.

8.5 Sampling monthly consumption data from 1926-1959

To begin, using the full monthly sample of consumption data spanning 1959-2008, we first demean both consumption and IP growth rates by their respective year-by-year average growth rates, denoted Δca_t and Δipa_t respectively. Then, we regress the demeaned consumption series on leads and lags of the demeaned IP series . Specifically, we use the following regression model:

$$\left(\Delta c_t-\Delta ca_t\right)=b_0\left(\Delta ip_t-\Delta ipa_t\right)+\sum_{i=1}^{lags}b_i^{lag}\left(\Delta ip_{t-i}-\Delta ipa_{t-i}\right)+\sum_{i=1}^{leads}b_i^{lead}\left(\Delta ip_{t+i}-\Delta ipa_{t+i}\right)+\varepsilon_t \tag{65}$$

We examined lead and lag lengths up to 4 months, but the usual BIC and AIC criteria both select one lag and no leads. Adopting this recommendation, estimation of this regression yields $\widehat{b}_0=0.0556$ and $\widehat{b}_1^{lag}=-0.0380$, with only the former statistically different from zero. While the R^2 from the model is modest (1.5 percent), it is only used to model the intra-year consumption growth variations pattern is available in the IP data. Specifically, we create draws for the monthly consumption series from 1929-1958 as follows

$$\Delta c_t^{draw}=\Delta ca_t+\widehat{b}_0\left(\Delta ip_t-\Delta ipa_t\right)+\widehat{b}_1^{lag}\left(\Delta ip_{t-1}-\Delta ipa_{t-1}\right)+\varepsilon_t^{draw} \tag{66}$$

where we draw ε_t^{draw} from a normal distribution with zero mean and variance equal to the sample variance of the residual, ε_t. Given a draw Δc_t^{draw}, we splice it with the actual consumption data from 1959-2008, and proceed to calculate vc_t^{draw} and sc_t^{draw} using the same methods as outlined in Section 2 for the shorter consumption series.[17] Finally, for each draw, we calculate sample statistics (and standard errors) for Δc_t^{draw}, vc_t^{draw} and sc_t^{draw} exactly as in Section 2.

[17] The available asset prices for the vc_t and sc_t projections are different, however, for the longer sample. We use the dividend yield, AAA and BAA corporate bond rates, and a measure of $rvar_t$ that is based on squared daily returns.

8.6 Survey data

We utilize survey data available from the Survey of Professional Forecasters currently conducted by the Federal Reserve Bank of Philadelphia. The data is available at the quarterly frequency. First, we use the one-quarter ahead GDP deflator inflation forecast from the SPF to deflate rf_t to form a measure of the risk free rate, rrf_t, maintaining the assuming that the inflation risk premium is zero at the quarterly horizon. The expected inflation data is available from the start of the SPF in 1968Q4. We also use the conditional distribution of four-quarter real GDP growth . Since 1981Q3, the SPF has asked respondents to fill in probabilities for histograms over real GDP growth outcomes for the coming year.[18] For instance, respondents are asked to fill in the probability that real GDP growth over the next year will fall into the "zero-to-one percent" bin. Unfortunately, the bins' boundaries have not been stable over the history of the SPF. To deal with this, we create "uber-bins" to which we can consistently assign probability from all the surveys. For instance, if a particular survey asked for separate probabilities for "zero-to-one" and "one-to-two" percent growth, we sum these probabilities for the "zero-to-two" uber-bin. For each uber-bin, we calculate the first, second and third uncentered empirical moments using historical US annual real GDP growth data from 1930-2008. For instance, conditional on GDP growth being less than -0.02, the first, second and third uncentered moments of historical GDP growth are -0.09 , 0.10^2 and -0.10^3 respectively. The below table summarizes these statistics for all the uber-bins.

	$(-\infty, -0.02]$	$(-0.02, 0.00)$	$[0.00, 0.02)$	$[0.02, 0.04]$	$(0.04, 0.06)$	$[0.06, \infty)$
$E_t\left[g_{t+1}\right]$	-0.09	-0.01	0.01	0.03	0.05	0.10
$E_t\left[g_{t+1}^2\right]^{1/2}$	0.10	0.01	0.01	0.03	0.05	0.10
$\lvert E_t\left[g_{t+1}^3\right]\rvert^{1/3}$	0.10	0.01	0.01	0.03	0.05	0.11

We use these conditional expectations together with the cross-sectional mean probabilities attached to each of the uber-bin to calculate the first three uncentered moments for the full distribution as:

$$E_t\left[g_{t+1}^i\right] = \sum_{b=1}^{4} prob_t\left(bin = b\right) \cdot E_t\left[g_{t+1}^i | bin = b\right] \tag{67}$$

where the summation runs over the six uber-bins shown in the above table and $prob_t\left(bin = b\right)$ is the cross-sectional mean probability attached to bin b in the survey dated t.

[18] In actuality, 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, etc.

References

[1] 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.

[2] Abel, A., 1990, Asset prices under habit formation and catching up with the Joneses, American Economic Review, 80, 38-42.

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

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

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

[6] Barro, R. J., E. Nakamura, J. Steinsson, J. Ursua, 2009, "Crises and Recoveries in an Empirical Model of Consumption Disasters, working paper.

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

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

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

[10] Bekaert, G., 1996, The time-variation of risk and return in foreign exchange markets: A general equilibrium perspective, Review of Financial Studies 9, 427-470.

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

[12] Bekaert, G., Engstrom, Eric., Xing, Y., 2009, Risk, uncertainty, and asset prices, Journal of Financial Economics, 91, 59-82.

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

[14] Bollerslev, T., Tauchen, G. E., and Zhou, H., 2008, Expected Stock Returns and Variance Risk Premia, Review of Financial Studies, Forthcoming

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

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

[17] Campbell, J. Y., Cochrane, J. H., 1995, By force of habit: A consumption based explanation of aggregate stock market behavior, NBER working paper.

[18] 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.

[19] 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.

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

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

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

[23] 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.

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

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

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

[27] 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.

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

[29] 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.

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

[31] 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.

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

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

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

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

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

[37] 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.

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

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

[40] Rietz, T.A., 1988,The equity risk premium a solution, Journal of Monetary Economics, 22, 117-131.

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

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

[43] Wachter, J., 2009, Can Time-varying Risk of Rare Disasters Explain Stock Market Volatility? working Paper.

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

[45] Wooldridge, 2002, Econometric Analysis of Cross Section and Panel Data, the MIT Press.

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

Table 1: Key Sample Statistics

Basic monthly series

	Δc_t	rrf_t	dp_t	ret_t	$pvar_t$	$qvar_t$	$qvar_t - pvar_t$
mean	0.0025	0.0010	−6.3948	0.0037	0.0021	0.0038	0.0017
	(0.0002)	(0.0003)	(0.0941)	(0.0042)	(0.0005)	(0.0006)	(0.0002)
std	0.0028	0.0012	0.3375	0.0433	0.0028	0.0037	0.0013
	(0.0002)	(0.0001)	(0.0410)	(0.0045)	(0.0008)	(0.0009)	(0.0001)
ac(1)	−0.1947	0.9839	0.9830	0.0612	0.7584	0.7599	0.8008
	(0.0941)	(0.1666)	(0.2214)	(0.0976)	(0.0869)	(0.0635)	(0.0909)
skew(Δc_t)	−0.1116	kurt(Δc_t)	3.7293				
	(0.1924)		(0.2964)				

Data are monthly from January 1990 through March 2009. All variables are expressed at a monthly rate. The variables include real nondurables and services consumption growth, Δc_t, the real short rate, rrf_t, the logarithmic dividend yield, dp_t, equity returns, ret_t, the conditional variance of returns under the physical and risk-neutral measures, $pvar_t$ and $qvar_t$ respectively, and the variance premium, $qvar_t - pvar_t$. GMM standard errors using 20 Newey-West (1987) lags are reported in parentheses.

Table 2: Estimating Higher Order Consumption Growth Dynamics

Panel A: Spanning higher Δc_t conditional moments

	pvals	mean	std	ac(1)
vc_t $(\times 10^4)$	(0.0002)	0.0784	0.0330	0.7791
		(0.0062)	(0.0063)	(0.0971)
sc_t $(\times 10^8)$	(0.0002)	−0.1851	3.1936	0.7599
		(0.5420)	(0.6286)	(0.0537)

Panel B: Correlations with higher Δc_t moments

	rrf_t	dp_t	$pvar_t$	$qvar_t$
vc_t	−0.3412	0.0145	0.8602	0.9268
	(0.1228)	(0.1904)	(0.0247)	(0.0244)
sc_t	0.4325	−0.4174	−0.7199	−0.6984
	(0.1077)	(0.1036)	(0.1686)	(0.1742)

In this table, we present results for the following system of regressions:

$$\begin{aligned} \Delta c_{t+1} &= \overline{g} + u^1_{t+1} \\ (\Delta c_{t+1} - \overline{g})^2 &= m_2 + x_t'\beta_2 + u^2_{t+1} \\ (\Delta c_{t+1} - \overline{g})^3 &= m_3 + x_t'\beta_3 + u^3_{t+1} \end{aligned}$$

and we take the fitted conditional variance and centered third moment to be, respectively,

$$\begin{aligned} vc_t &= m_2 + x_t'\beta_2 \\ sc_t &= m_3 + x_t'\beta_3 \end{aligned} \tag{68}$$

The top row of Panel A reports the p-value for the joint significance of β_2 and the second row reports the joint significance for β_3. The standard errors use 3 Newey-West (1987) lags. x_t includes the real short rate, rrf_t, the dividend yield, dp_t, physical and risk-neutral variance measures, $pvar_t$ and $qvar_t$, and exponentially-weighted (with parameter 0.1) moving averages of squared and cubed (demeaned) consumption growth. The subsequent columns report some sample statistics for vc_t and sc_t with GMM standard errors in parentheses. These standard errors use 20 Newey-West (1987) lags, but do not correct for first-stage estimation error in the $\beta's$. Panel B reports correlations of vc_t and sc_t with some of the instruments.

Table 3: Structural Model Estimates

p_t	$\overline{p}$	26.4742	ρ_p	0.6445	σ_{pp}	2.2556		
		(15.6575)		(0.1138)		(1.1517)		
n_t	$\overline{n}$	0.0659	ρ_n	0.9955	σ_{nn}	0.0190		
		(0.0132)		(0.0055)		(0.0141)		
Δc_t	$\overline{g}$	0.0026	σ_{cp}	0.0005	σ_{cn}	0.0033		
		(0.0002)		(0.0001)		(0.0003)		
q_t	$\overline{q}$	1.0000	ρ_q	0.9948	σ_{qp}	−0.0018	σ_{qn}	0.1438
		($fixed$)		(0.0017)		(0.0006)		(0.0288)
m_t	$\ln(\beta)$	−0.0003	γ	2.3241				
		($fixed$)		(0.7056)				
Jstat	45.91							
pval	(0.0044)							

The model being estimated is summarized by the equations

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

Estimation uses the Classical Minimum Distance method using our monthly sample from Jan 1990 through March 2009. Standard errors are in parentheses. The parameters matched by CMD are those in Table 1 and Panel A of Table 2. The Jstat statistic is the test of over-identifying restrictions, the distribution of which is described in the Appendix.

Table 4: Model Performance

Panel A: Key sample statistics

	Δc_t	rrf_t	dp_t	ret_t	$pvar_t$	$qvar_t - pvar_t$
mean	[0.0026]	[0.0009]	[−6.4227]	[0.0042]	[0.0016]	[0.0017]
	0.0025	0.0010	−6.3948	0.0037	0.0021	0.0017
	(0.0002)	(0.0003)	(0.0941)	(0.0042)	(0.0005)	(0.0002)
std	[0.0028]	[0.0013]	[0.3594]	[0.0398]	[0.0010]	[0.0013]
	0.0028	0.0012	0.3375	0.0433	0.0028	0.0013
	(0.0002)	(0.0001)	(0.0410)	(0.0045)	(0.0008)	(0.0001)
ac(1)	[0.0000]	[0.9726]	[0.9944]	[−0.0029]	[0.9954]	[0.6508]
	−0.1947	0.9839	0.9830	0.0612	0.7584	0.6986
	(0.0941)	(0.1666)	(0.2214)	(0.0976)	(0.0869)	(0.1644)
skew(Δc_t)	[0.1254]	kurt(Δc_t)	3.9318	$corr\,(\Delta pvar_t, \Delta vprem_t)$		[0.5180]
	−0.1101		3.7293			0.1529
	(0.1924)		(0.2964)			(0.1922)

Panel B: Sample statistics for higher Δc_t moments

	mean	std	ac(1)
vc_t $(x10^4)$	[0.0774]	[0.0408]	[0.6508]
	0.0784	0.0330	0.7791
	(0.0062)	(0.0063)	(0.0971)
sc_t $(x10^8)$	[0.2704]	[0.5476]	[0.7922]
	−0.1851	3.1936	0.7599
	(0.5420)	(0.6286)	(0.0537)

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 and corresponding standard errors are reproduced from Tables 1 and 2. Panel B reports on the model-implied versus sample statistics for the conditional variance and centered third moment of consumption growth, vc_t and sc_t, respectively.

Table 5: Factor Loadings

	rrf_t	dp_t	$eqprem_t$	$pvar_t$	$qvar_t - pvar_t$
p_t	-0.0001	-0.00004	0.0001	0.0001	-0.00003
n_t	-0.0765	-5.1607	0.0448	0.0192	0.0252
q_t	0.0120	1.6244	-0.0010	0.0000	0.0000

This table reports the loadings of various endogenous variables on the state vector, $Y_t = [p_t, n_t, q_t]'$ for the model and point estimates reported in Table 3.

Table 6: Equity Return Predictability

Panel A: Return predictability in the model and data sample

	rrf_t	dp_t	$qvar_t - cvar_t$	$qvar_t - rvar_t$
	[0.4052]	[0.0047]	[1.5018]	–
	3.8341	0.0077	2.4373	3.3484
	(2.4051)	(0.0085)	(2.2651)	(1.3376)
sample R^2	0.0110	0.0036	0.0051	0.0267

Panel B: Equity risk premium dynamics under the model

	$E_t\,(ret_{t+1} - rrf_t)$	$VAR_t\,(ret_{t+1} - rrf_t)$	$\frac{VAR(E_t(ret_{t+1}-rrf_t))}{VAR(ret_{t+1}-rrf_t)}$
analytic	$0.0011 + 0.0001p_t + 0.0448n_t - 0.0010q_t$	$0.12^{e-4}p_t + 0.19^{e-1}n_t$	
$Y_t = \overline{Y}$	0.0033	0.0016	0.0024

Panel A reports the univariate predictability of one-period head excess equity returns with respect to instruments listed in columns. The coefficient implied by the model is 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 statistics are listed below. Panel B reports the dependence of the conditional mean and variance of excess equity returns implied by the structural model at the parameters estimated in Table 4.

Table 7: The Conditional Distribution of Equity Returns

Panel A: Univariate statistics for risk-neutral return moments

	$\overline{qvar_t^{1/2}}$	$\overline{qsk_t}$	$\overline{qkt_t}$
BEGE main estimation	[0.20]	[-6.6]	[78.1]
Data	0.20	-2.4	20.5
BEGE long cons estimation	{0.20}	{-3.5}	{27.5}

Panel B: Correlations: $\Delta qvar_t$

	ret_t	ret_{t-1}
BEGE main estimation	[-0.5141]	[0.0027]
Data	-0.6291	0.0748
	(0.0759)	(0.1000)
BEGE long cons estimation	{-0.9991}	{0.0024}

Panel A reports on the univariate properties of the higher order moments of returns under the risk-neutral measures when the state vector is at its unconditional mean. The row labeled "Data" in Panel A (only) reproduces results from Table 3 of Figlewski (2009). The bottom row reports BEGE model-implied moments estimated using long-term consumption growth data as described in Section 6. Panel B reports on the correlations with changes in the risk neutral variance and realized returns. In both panels, model-implied moments are in brackets. Sample data are reported with GMM standard errors, when available, (20 Newey West lags) below in parentheses.

Table 8: Long-Term Perspective of Consumption Growth Dynamics

Panel A: Sample statistics for consumption growth

	pvals	mean	std	ac(1)	skew	kurt
Δc_t		[0.0027]	[0.0047]	[0.0000]	[−0.1334]	[4.2188]
		0.0027	0.0046	0.0153	−0.1305	4.0214
		(0.0003)	(0.0002)	(0.0714)	(0.1569)	(0.2847)
vc_t $(\times 10^4)$		[0.2200]	[0.1205]	[0.9847]		
	(< 0.0001)	0.2155	0.1183	0.9865		
		(0.0168)	(0.0202)	(0.0212)		
sc_t $(\times 10^8)$		[−1.3771]	[4.7607]	[0.9859]		
	(< 0.0001)	−1.2999	14.0113	0.9757		
		(1.8820)	(3.2215)	(0.0191)		

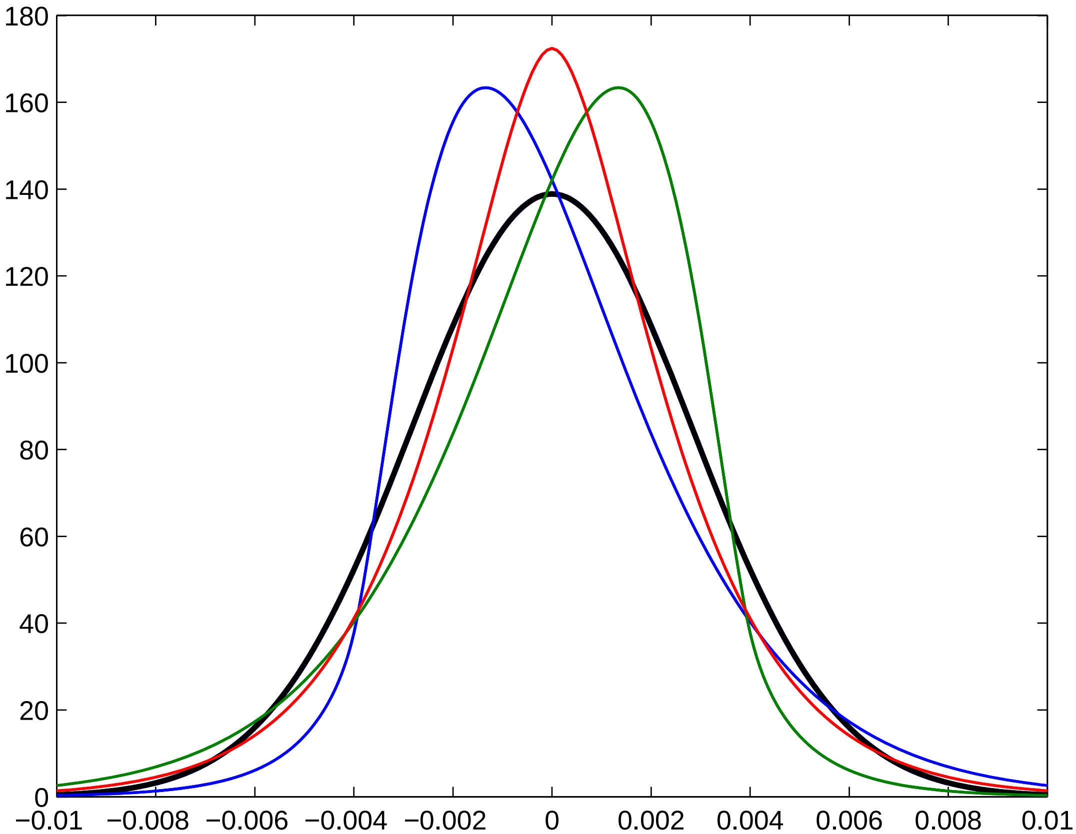
Panel B: Point Estimates

p_t	$\overline{p}$	30.2652	ρ_p	0.9884	σ_{pp}	0.6069		
n_t	$\overline{n}$	0.2374	ρ_n	0.9859	σ_{nn}	0.0903		
Δc_t	$\overline{g}$	0.0027	σ_{cp}	0.0008	σ_{cn}	0.0044		
q_t	$\overline{q}$	1.0000	ρ_q	0.9898	σ_{qp}	0.0003	σ_{qn}	0.0783
m_t	$\ln(\beta)$	−0.0003	γ	2.6114				

Panel A reports on the properties of monthly consumption growth based on a sample extending back to 1929. The text describes our methodology for sampling consumption data for this sample. The variables vc_t and sc_t refer to the conditional second and third centered moments respectively. The statistics in square brackets are the model-implied moments, computed using the structural parameters reported in Panel B.

This table reports results from calibration of the structural model using the longer-term consumption data reported in Table 8. The model estimated is the same as reported in Table 3. In Panel A, we report the calibrated statistics for the model.

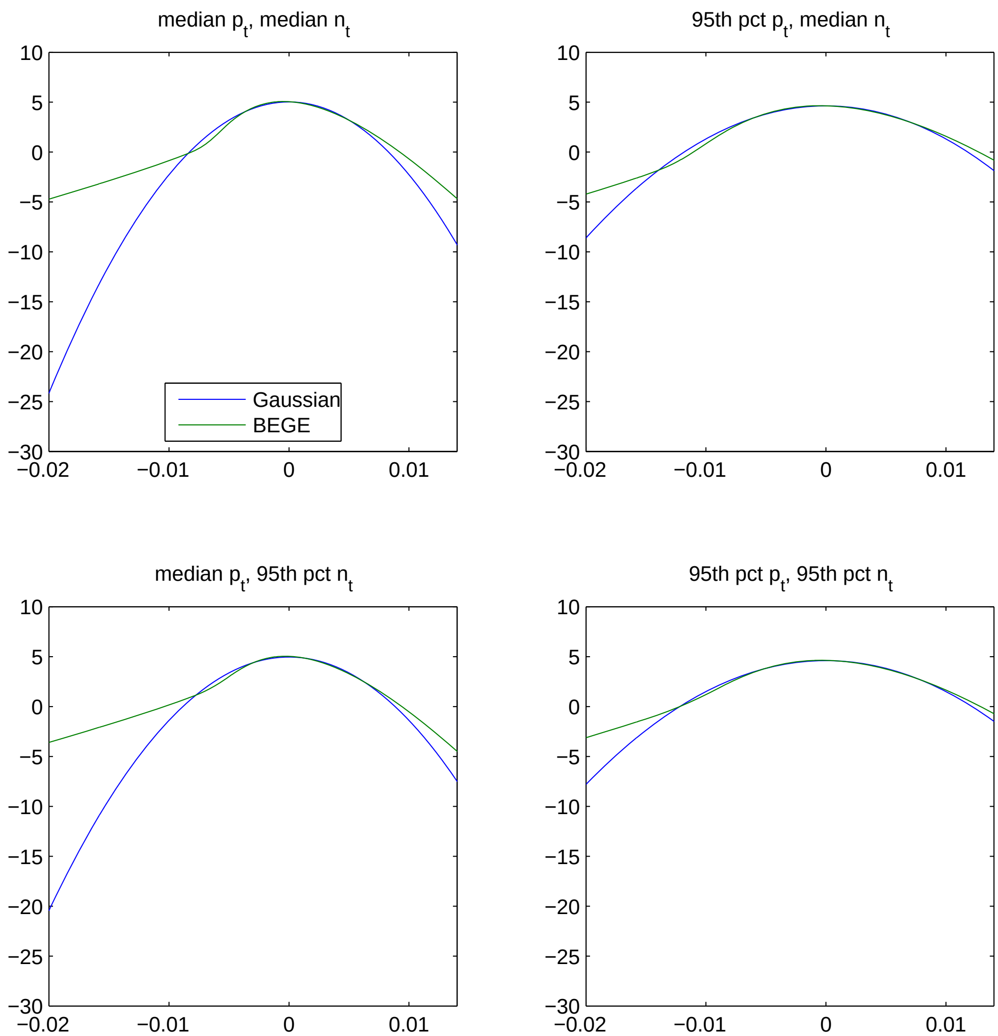
Figure 1: Examples of the BEGE distribution



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.

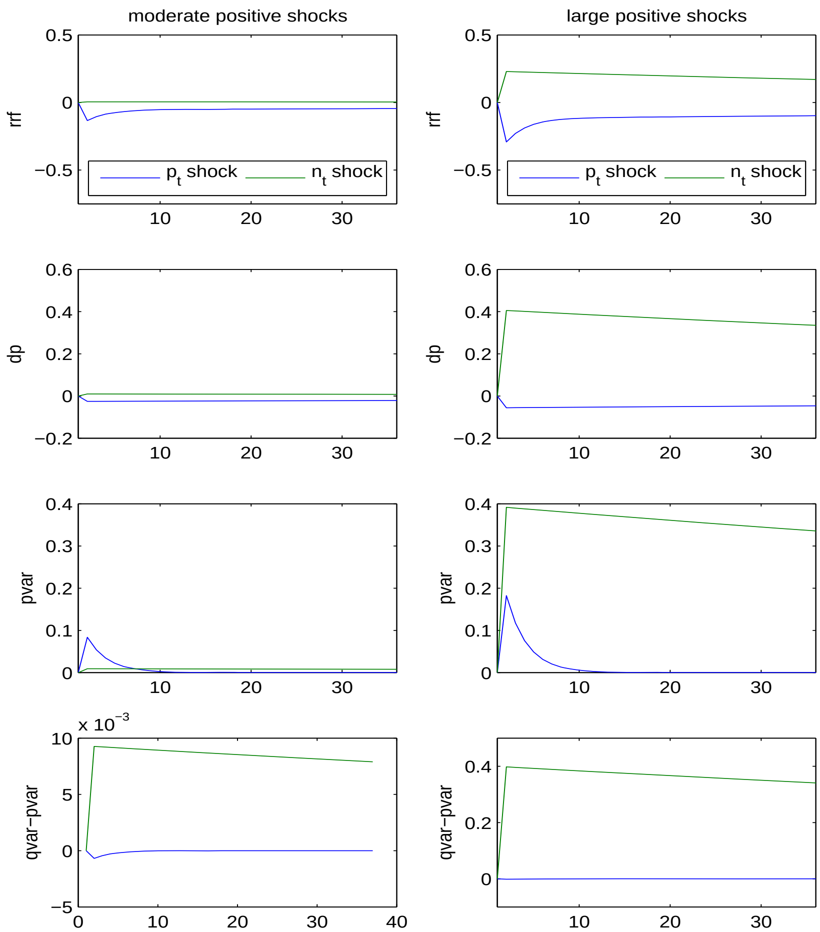
	p_t	n_t	σ_{cp}	σ_{cn}
black	40	40	0.0003	0.0003
red	2	2	0.0014	0.0014
green	.4	3	0.0016	0.0016
blue	3	.4	0.0016	0.0016

Figure 2: Estimated Log Density of Δc_t



This figure plots the log 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 normal log densities with the same mean and variance as the BEGE density for at each configuration of p_t and n_t.

Figure 3: Impulse Responses under the Structural Model

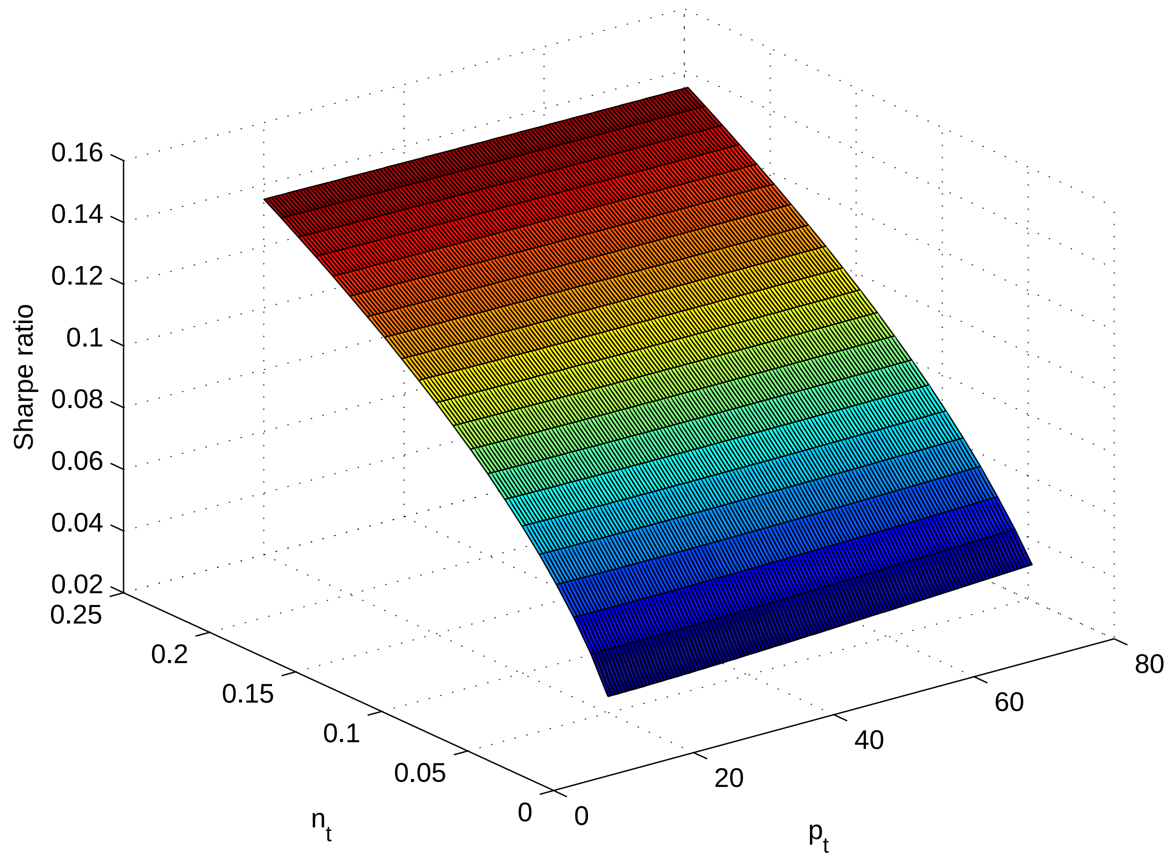


This figure shows the impulse response of rrf_t, dp_t, $pvar_t$, and $qvar_t - pvar_t$ to shocks to p_t and n_t. For all variables, the units on the vertical axis are unconditional standard deviations. In all panels, the shocks occur at month 1 and the horizontal axis runs from 0 months (prior to the shock) through 36 months. In the left column, impulse responses to 90th percentile shocks to p_t and n_t are reported. In the right column, response to 99th percentile shocks are reported. For p_t, the 90th and 99th percentile shock values are 3.16 and 6.88 respectively. For n_t the 90th and 99th percentile shock values are 0.025 and 1.074 respectively. Note that the scale in the bottom left panel has been expanded for visibility. The response of each endogenous variable in j periods, iz_{t+j}, is given by

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

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

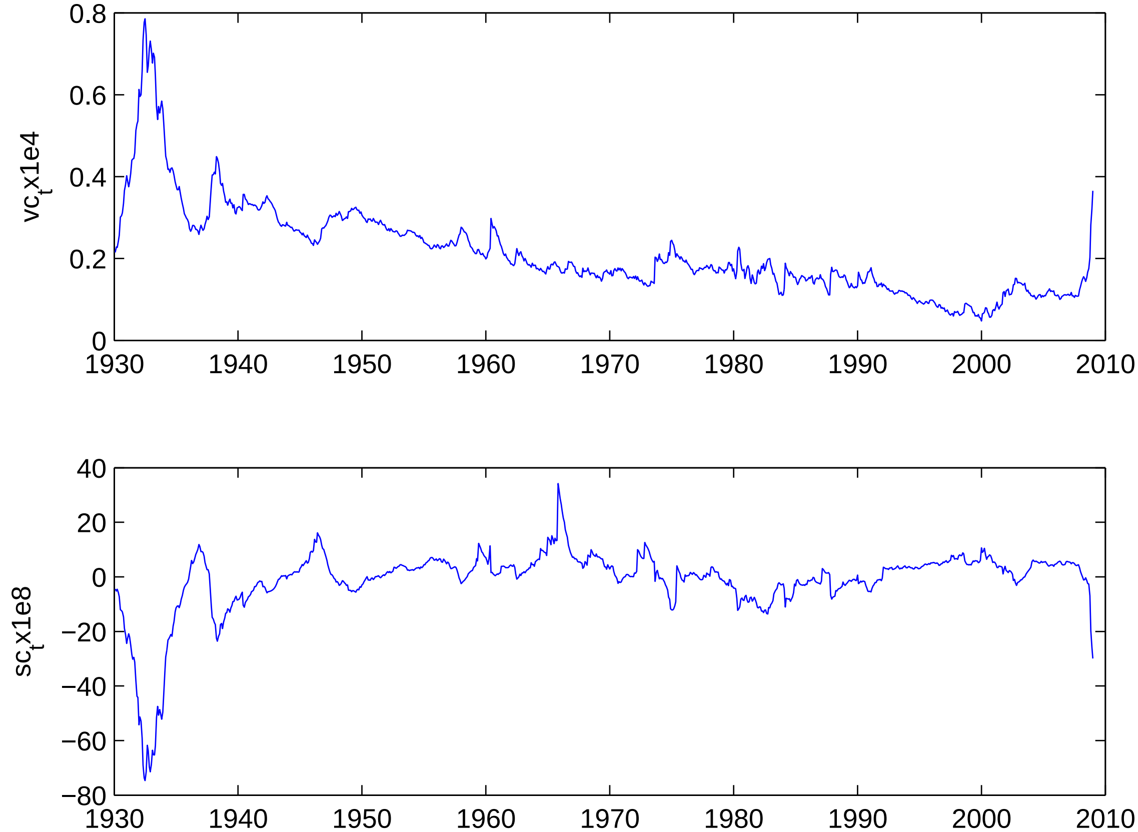
Figure 4: Dependence of the Equity Conditional Sharpe Ratio



This figure reports the monthly Sharpe ratio for one-month ahead equity returns under the structural model and point estimates in Table 4 calculated as

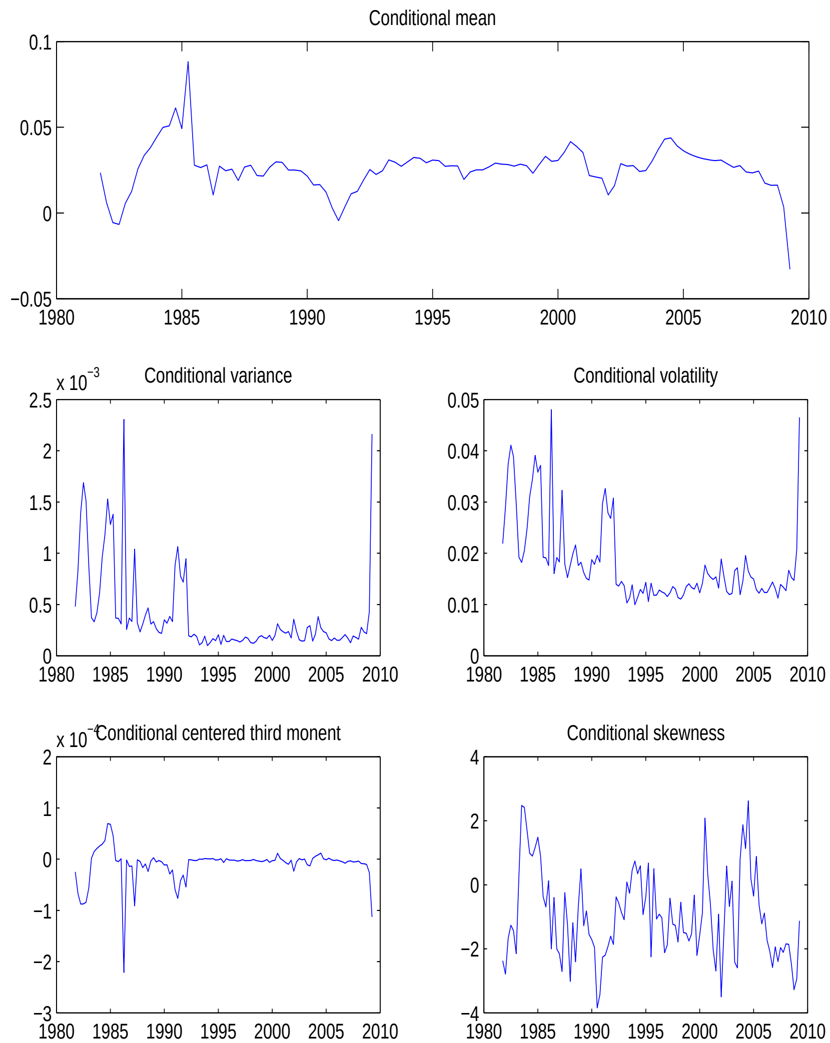
$$Sharpe\ ratio = \frac{E_t\left[ret_{t+1} - rrf_t\right]}{VAR_t\left[ret_{t+1}\right]^{1/2}}$$

Figure 5: Long-term perspective on Conditional Consumption Moments



This figure presents the median draws of vc_t and sc_t as described in Section 6.

Figure 6: Conditional Moments of real GDP Growth from Survey Data



This figure presents the conditional expectation of four-quarter GDP growth from survey data in the SPF. Data are quarterly from 1981Q1 through 2009Q1. The appendix describes the construction of these series.