

Rational Pessimism, Rational Exuberance, and Asset Pricing Models

RAVI BANSAL, A. RONALD GALLANT
Fuqua School of Business, Duke University

and

GEORGE TAUCHEN
Duke University

First version received May 2004; final version accepted February 2007 (Eds.)

The paper estimates and examines the empirical plausibility of asset pricing models that attempt to explain features of financial markets such as the size of the equity premium and the volatility of the stock market. In one model, the long-run risks (LRR) model of Bansal and Yaron, low-frequency movements, and time-varying uncertainty in aggregate consumption growth are the key channels for understanding asset prices. In another, as typified by Campbell and Cochrane, habit formation, which generates time-varying risk aversion and consequently time variation in risk premia, is the key channel. These models are fitted to data using simulation estimators. Both models are found to fit the data equally well at conventional significance levels, and they can track quite closely a new measure of realized annual volatility. Further, scrutiny using a rich array of diagnostics suggests that the LRR model is preferred.

1. INTRODUCTION

Asset market models provide alternative explanations for a wide range of asset markets anomalies. Campbell and Cochrane (1999) and Bansal and Yaron (2004) use calibration to verify that their models can account for the equity premium, risk-free rate, and asset price volatility puzzles. The empirical plausibility of the key risk channels, which are the slow-moving habits in the Campbell–Cochrane model and the long-run risks (LRR) in the Bansal and Yaron model, is not tested against data. Hence, it is not clear if aggregate economic data support the risk channels highlighted in these papers. Our simulation-based estimation approach provides data-driven estimates of structural parameters and a rich array of diagnostics that help discriminate between these models.

The primary contribution of this paper is to estimate, using simulation techniques, the Campbell and Cochrane (1999) and Bansal and Yaron (2004) models and evaluate if there is any empirical support for the risk channels developed in these papers. We provide estimates of the structural parameters and highlight the differences between these models. We also impose in estimation the cointegration restriction between consumption and dividends, unlike Campbell and Cochrane (1999) and Bansal and Yaron (2004). Imposing this restriction is economically well motivated because aggregated dividends and consumption cannot permanently deviate from each other and financial wealth cannot permanently deviate from aggregate wealth.

Below, we set forth the LRR model, which is patterned after Bansal and Lundblad (2002) and Bansal and Yaron (2004). In the LRR model, growth rates of cash flows contain a statistically

small but economically important slow-moving component. As is well-known from the familiar Gordon growth formula, small, near-permanent movements in the forecasted growth rates of cash flows can be expected to generate large movements in asset valuations relative to current cash flows. The LRR model couples these cash flow dynamics with preferences of the Epstein–Zin–Weil type. One departure from the Bansal and Yaron (2004) specification is the inclusion of cointegration restriction between aggregate stock market dividends and consumption.

We also present a model with habit persistence (HAB) that follows Campbell and Cochrane (1999) rather closely. Their dynamics must be modified slightly to be compatible with the cointegration relationships in the data. The HAB model presumes quite simple cash flow dynamics, in contrast to the LRR model, but it assumes an involved preference structure with a time-varying habit stock that evolves with conditionally heteroskedastic innovations. Like the LRR model, the HAB model also generates long-term swings in stock market valuations.

In addition to these models, we also provide evidence on a short-run risks (SRR) model, which is similar to the type considered in Hall (1978), Hansen and Singleton (1982), and Mehra and Prescott (1985). This model is the traditional baseline which emphasizes SRR in consumption as the key risk channel in the economy.

We undertake econometric estimation of both the LRR and the HAB models using a simulation-based procedure similar to that of Smith (1993). Intuitively, the estimation technique minimizes a statistical measure of distance between a vector autoregression (VAR) model fitted to observed data and simulated data at the annual frequency. The simulations are generated by operating the models in monthly time, numerically solving for the equilibrium, and then aggregating appropriately to the annual frequency. The estimation is a type of generalized method of moments (GMM) estimation, and it therefore delivers a chi-squared measure of fit to the data. We also undertake extensive additional diagnostic assessments using reprojected (Gallant and Tauchen, 1998, 2007). As part of this effort, we develop a new measure of realized annual volatility, which is patterned after the measure developed by Andersen, Bollerslev and Diebold (2007), and we assess the models' capabilities to track this new variable.

The empirical work, which simultaneously imposes restrictions on consumption and asset market data, shows that the SRR model is sharply rejected in the data. Further, very large persistence is required in the habit process of the HAB model to account for the volatility of the dividend yield and the equity premium. Our estimation finds support for the key channels highlighted in the LRR model—low-frequency movements in consumption growth and in movements in consumption volatility. The preference parameters for both models are estimated at plausible values. Model specification tests do not reject the HAB and LRR models.

While the omnibus chi-square statistic provides support for both models, our diagnostics permit discrimination across them. One channel for discrimination is the role that high-frequency movements in consumption play in accounting for the risk–return relation across the two models. The consumption beta of the market return in the HAB model is 4.19, for the LRR model it is 0.52. In the observed data, the consumption beta is 0.79. Unlike the HAB model, high-frequency consumption movements are not significantly compensated in the LRR model. Kiku (2006) uses this intuition and the LRR model to account for the value premium puzzle and the general failure of the standard consumption beta model. The role for high-frequency consumption movements in accounting for risk in the HAB model and the role of low-frequency movements in consumption to account for the risk–return relation in the LRR provide a clear distinction between the two models. The data, as we document, do not support a significant role for the high-frequency consumption movements in accounting for risk premia. In this sense, we find greater support for the LRR model.

To further underscore the different channels across the HAB and LRR models, we also consider the valuation of hypothetical contingent claims on the market index (price of dividend

claim) and aggregate wealth claims (price of aggregate consumption claim). The contingent claim prices on the market index are very similar across both models, suggesting that the underlying market return distribution is very similar across these models. However, there are big differences in the contingent claim prices for aggregate wealth index. The aggregate wealth contingent claim prices are very large in the HAB model relative to the LRR model. Further, the contingent claim prices for the market index and the aggregate wealth index are very similar to each other in the HAB model, implying that return distribution for both claims is very similar in the HAB model. This suggests that consumption insurance, from the perspective of HAB model, is very expensive relative to the LRR model.

Section 2 below sets out the LRR model, and Section 3 sets forth HAB model. Section 4 develops the observation equations along with the new measure of realized annual volatility. Section 5 describes the data and the cointegration analysis. Section 6 contains the estimation results and is followed by Section 7 which evaluates the fitted models' performance. Section 8 contains concluding remarks. Details of the simulation-based methods are described in a technical appendix.

2. LONG-RUN RISKS

We develop an asset pricing model that is an extension of Bansal and Yaron (2004). Some key features of the model are that consumption and dividends are separate stochastic processes but are tied together by a long-run cointegrating relationship. Other features include preferences of the Epstein and Zin (1989) and Weil (1989) type along with time-varying stochastic volatility.

2.1. Dynamics of driving variables

Let $d_t = \log(D_t)$ and $c_t = \log(C_t)$ denote log real *per capita* values of the stock dividend and the consumption endowment, respectively. The log endowment c_t is assumed to be generated as

$$c_t = c_{t-1} + \mu_c + x_{t-1} + \epsilon_{ct}, \quad (1)$$

where μ_c is the average growth rate of c_t , x_{t-1} is the stochastic part of the conditional mean of consumption growth, that is, the LRR process, and ϵ_{ct} is the error. Let $\Delta c_t = c_t - c_{t-1}$. We impose (and later test) the restriction that log consumption and log dividends are cointegrated,

$$d_t - c_t = \mu_{dc} + s_t,$$

where μ_{dc} is a constant and s_t is an I(0) process. Note that the cointegration coefficient is set at one and the deviation $d_t - c_t$ is strictly stationary; that is, dividends and consumption share the same deterministic and stochastic trends. Finally, we introduce a stochastic volatility factor v_t and below adopt a standard exponential stochastic volatility process for the dynamics of the variables of interest. For compactness, collect the four variables into the vector

$$q_t = \begin{pmatrix} \Delta c_t \\ x_t \\ s_t \\ v_t \end{pmatrix}, \quad (2)$$

which we assume is a VAR(1) process with stochastic volatility that takes the form

$$q_t = a + Aq_{t-1} + \exp(\Lambda_t)\Psi z_t, \quad (3)$$

where z_t is a 4×1 standard normal random variable; the parameters of the VAR are

$$a = \begin{pmatrix} \mu_c \\ 0 \\ -\mu_{dc} \\ \mu_\sigma \end{pmatrix}, \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & \rho_x & 0 & 0 \\ 0 & \lambda_{sx} & \rho_s & 0 \\ 0 & 0 & 0 & \rho_v \end{bmatrix}, \quad (4)$$

with volatility structure

$$\Lambda_t = \begin{bmatrix} b_{cc}v_t & 0 & 0 & 0 \\ 0 & b_{xx}v_t & 0 & 0 \\ 0 & 0 & b_{ss}v_t & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \Psi = \begin{bmatrix} \Psi_{cc} & \Psi_{cx} & \Psi_{cs} & 0 \\ 0 & \Psi_{xx} & 0 & 0 \\ 0 & 0 & \Psi_{ss} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (5)$$

From the matrix A , it is seen that the LRR factor x_t is Granger causally prior and the volatility factor v_t evolves autonomously. The zero restrictions within A , along with the diagonality of Λ_t , and the zero restrictions within Ψ shut down many paths of cross-variable feedback in the interests of parsimony. On these assumptions, the minimal state vector is

$$u_t = \begin{pmatrix} x_t \\ s_t \\ v_t \end{pmatrix}, \quad (6)$$

knowledge of which suffices to simulate the system one step forward.

Note that cointegration between c_t and d_t implies that $\Delta d_{t+1} = \Delta c_{t+1} + \Delta s_{t+1}$. Hence, given equation (3) and the dynamics of s_{t+1} , the implied dividend growth rate specification is identical to Bansal and Yaron (2004) save for the fact that s_t is needed in addition to x_t to forecast dividend growth rates. Cointegration between d_t and c_t ensures that the long-run variances of dividend and consumption growth rates are equal to each other and the levels of these processes cannot deviate from each other permanently. This in turn ensures that the present value of the consumption stream (aggregate wealth) and the present value of the dividend stream (financial wealth) are cointegrated and cannot permanently deviate from each other.

In the LRR model, innovations in the x_t correspond to low-frequency risks or LRR, and the ϵ_{ct} correspond to high-frequency risks in consumption. Both these risks carry a distinct and different risk compensation, with the risk compensation associated with shocks to x_t being critical for accounting for asset prices. The common volatility process across consumption and dividends, v_t , as in Bansal and Yaron, is for parsimony. This variation in volatility is important for capturing time variation in risk premia.

2.2. Asset pricing

Let P_{ct} denote the price of an asset that pays the consumption endowment and let

$$v_{ct} = \frac{P_{ct}}{C_t} \quad (7)$$

denote the corresponding price–dividend ratio. The Epstein–Zin–Weil utility function is

$$U_t = \left[(1 - \delta)C_t^{\frac{1-\gamma}{\theta}} + \delta(\mathcal{E}_t U_{t+1}^{1-\gamma})^{\frac{\theta}{\theta-1}} \right]^{\frac{\theta-1}{\theta}}, \quad (8)$$

where γ is the coefficient of risk aversion,

$$\theta = \frac{1 - \gamma}{1 - 1/\psi}, \tag{9}$$

and ψ is the elasticity of inter-temporal substitution. As noted in Campbell (2002), Bansal and Yaron (2004), and elsewhere, the first-order conditions derived in Epstein and Zin (1989, 1991) imply that the price–dividend ratio, v_{ct} , is the solution to the non-linear expectational equation

$$v_{ct} = \mathcal{E}_t\{\delta^\theta \exp[-(\theta/\psi)\Delta c_{t+1} + (\theta - 1)r_{c,t+1}](1 + v_{c,t+1})\exp(\Delta c_{t+1})\}, \tag{10}$$

where

$$r_{c,t+1} = \log \left[\frac{1 + v_{c,t+1}}{v_{ct}} \exp(\Delta c_{t+1}) \right] \tag{11}$$

is the geometric return on the asset. Evidently, the one-period marginal rate of substitution is

$$M_{t,t+1} = \delta^\theta \exp[-(\theta/\psi)\Delta c_{t+1} + (\theta - 1)r_{c,t+1}]. \tag{12}$$

The price–dividend ratio $v_{dt} = P_{dt}/D_t$ on the asset that pays D_t is the solution to

$$v_{dt} = \mathcal{E}_t\{\delta^\theta \exp[-(\theta/\psi)\Delta d_{t+1} + (\theta - 1)r_{d,t+1}](1 + v_{d,t+1})\exp(\Delta d_{t+1})\}.$$

The one-step-ahead risk-free rate r_{ft} is the solution to

$$e^{-r_{ft}} = \mathcal{E}_t\{\delta^\theta \exp[-(\theta/\psi)\Delta c_{t+1} + (\theta - 1)r_{c,t+1}]\}.$$

The outcome variables of interest are the price–dividend ratios, the risk-free rate, and returns. These are obtained by evaluating the pricing functions

$$\begin{aligned} v_{ct} &= v_c(u_t), \\ v_{dt} &= v_d(u_t), \\ r_{ft} &= r_f(u_t), \end{aligned} \tag{13}$$

which are determined as functions of the state vector (6) using the solution method described in the Appendix, and then generating returns via

$$\begin{aligned} r_{ct} &= \log \left[\frac{(1 + v_{ct}) \exp(\Delta c_t)}{v_{c,t-1}} \right], \\ r_{dt} &= \log \left[\frac{(1 + v_{dt}) \exp(\Delta d_t)}{v_{d,t-1}} \right]. \end{aligned} \tag{14}$$

Cecchetti, Lam and Mark (1993) consider asset price implications of consumption and/or dividends models that incorporate regime shifts in the context of time separable utility. Similarly, Evans (1998) uses regime shift in dividends to interpret the behaviour of dividend yields across time. Models of regime shifts in growth rates may provide an alternative way to capture predictable variation in fundamentals, such as consumption growth.

3. HABIT PERSISTENCE

We now set forth a model of HAB that follows the specification of Campbell and Cochrane (1999) with the exception that we impose cointegration on their driving variables.

The agent is assumed to maximize the utility function

$$\mathcal{E}_t \sum_{i=0}^{\infty} \delta^i \frac{(C_{t+i} - X_{t+i})^{1-\gamma} - 1}{1-\gamma},$$

where C_t denotes real consumption, X_t denotes the habit stock, δ is a discount rate, and γ is a risk aversion parameter, although relative risk aversion itself depends on X_t as well and will be much higher than γ for plausible values of X_t . Define the surplus ratio

$$H_t = \frac{C_t - X_t}{C_t}$$

and assume that

$$h_t = \log(H_t)$$

evolves as

$$h_{t+1} = (1 - \rho_h)\bar{h} + \rho_h h_t + \lambda(h_t)\epsilon_{c,t+1},$$

where ρ_h and \bar{h} are parameters, $\epsilon_{c,t+1}$ is the consumption innovation defined in (1) above, and the sensitivity function $\lambda(h_t)$ is

$$\lambda(h_t) = \begin{cases} \frac{1}{\bar{H}} \sqrt{1 - 2(h_t - \bar{h})} - 1, & h_t \leq h_{\max}, \\ 0, & h_t > h_{\max}, \end{cases}$$

where

$$\bar{H} = \sigma_{\epsilon_c} \sqrt{\frac{\gamma}{1 - \rho_h}}$$

and σ_{ϵ_c} is the standard deviation of ϵ_{ct} . The value h_{\max} is the value at which $\lambda(h_t)$ first touches zero; by inspection,

$$h_{\max} = \bar{h} + \frac{1}{2}[1 - (\bar{H})^2],$$

and, in the continuous time limit, h_{\max} is an upper bound on h_t . These dynamics for h_t imply, among other things, that the risk-free rate is constant to a first approximation.

Under the assumption of external habit, the inter-temporal marginal rate of substitution is

$$M_{t+1} = \delta \left(\frac{H_{t+1} C_{t+1}}{H_t C_t} \right)^{-\gamma}$$

or, equivalently,

$$M_{t+1} = \delta \exp[-\gamma (\Delta h_{t+1} + \Delta c_{t+1})], \quad (15)$$

where $\Delta h_{t+1} = h_{t+1} - h_t$ and Δc_{t+1} is as defined in Section 2 above. Thus, in notation consistent with the preceding sections, the expectational equation for the price-dividend ratio $v_{ct} = P_{ct}/C_t$ for the asset that pays the consumption endowment C_t is

$$v_{ct} = \mathcal{E}_t\{\delta \exp[-\gamma (\Delta h_{t+1} + \Delta c_{t+1})](1 + v_{c,t+1}) \exp(\Delta c_{t+1})\}, \tag{16}$$

and the expectational equation for the price–dividend ratio $v_{dt} = P_{dt}/D_t$ for the asset that pays the dividend D_t is

$$v_{dt} = \mathcal{E}_t\{\delta \exp[-\gamma (\Delta h_{t+1} + \Delta c_{t+1})](1 + v_{d,t+1}) \exp(\Delta d_{t+1})\}. \tag{17}$$

The risk-free rate r_{ft} is the solution to

$$\exp(-r_{ft}) = \mathcal{E}_t\{\delta \exp[-\gamma (\Delta h_{t+1} + \Delta c_{t+1})]\}. \tag{18}$$

The above three equations correspond directly to their counterparts above for Epstein–Zin–Weil preferences. The geometric return is

$$r_{c,t+1} = \log \left[\frac{1 + v_{c,t+1}}{v_{ct}} \exp(\Delta c_{t+1}) \right] \tag{19}$$

on the consumption asset and

$$r_{d,t+1} = \log \left[\frac{1 + v_{d,t+1}}{v_{dt}} \exp(\Delta d_{t+1}) \right] \tag{20}$$

on the equity asset.

Campbell and Cochrane (1999) assume that log consumption c_t and the log dividend d_t processes are each I(1) with the same growth rates and correlated innovations. We can easily modify the dynamics (3) of Section 2 to retain the essential features of their model while incorporating cointegration between consumption and dividends. We retain the equation

$$d_t = \mu_{dc} + c_t + s_t.$$

We set $x_t = 0$ for all t ; so that

$$\begin{aligned} \Delta c_t &= \mu_c + \epsilon_{ct}, \\ \Delta d_t &= \mu_c + (\rho_s - 1)s_{t-1} + \epsilon_{ct} + \epsilon_{st}, \end{aligned} \tag{21}$$

and turn off stochastic volatility, that is, $v_t = 0$ for all t , so that ϵ_{ct} and ϵ_{st} immediately above are i.i.d. Gaussian. This set-up, with correlation between ϵ_{ct} and ϵ_{st} , is a direct extension of Campbell and Cochrane (1999). If $\rho_s = 1$, the set-up reduces to that of Campbell and Cochrane (1999) specification of i.i.d. dividend growth; in this case, dividends and consumption are not cointegrated and have separate stochastic trends. However, the imposition of unit cointegration implies that dividend growth rates are not i.i.d., as they are predictable via the error-correction variable s_t .

The outcome variables of interest are the same as for the LRR model above. Using the same solution technique, we solve for the pricing functions

$$\begin{aligned} v_{ct} &= v_c(u_t), \\ v_{dt} &= v_d(u_t), \\ r_{ft} &= r_f(u_t) \end{aligned} \tag{22}$$

and then generate returns. The minimal state vector for HAB is s_t and h_t . Here, we augment it to

$$u_t = \begin{pmatrix} s_t \\ h_t \\ \lambda(h_t) \end{pmatrix} \quad (23)$$

for numerical reasons.

4. TIME AGGREGATION AND THE OBSERVATION EQUATIONS

The LRR and HAB models defined in Sections 2 and 3 operate in monthly time while we observe annual consumption data and annual summary measures from the financial markets. To simulate the model, we run the VAR with stochastic volatility (3) at the monthly frequency, burn off the transients, and form the monthly levels of consumption C_t and dividends D_t . The annual aggregate consumption and dividends, C_t^a and D_t^a , are 12-month moving sums sampled annually, which are then converted to logs:

$$\begin{aligned} c_t^a &= \log(C_t^a), \quad t = 12, 24, 36, \dots, \\ d_t^a &= \log(D_t^a), \quad t = 12, 24, 36, \dots \end{aligned}$$

The annual price–dividend ratio is computed as

$$p_{dt}^a - d_t^a = \log\left(v_{dt} \frac{D_t}{D_t^a}\right), \quad t = 12, 24, 36, \dots$$

We also compute the annual geometric returns

$$\begin{aligned} r_{dt}^a &= \sum_{k=0}^{11} r_{d,t-k}, \\ r_{ft}^a &= \sum_{k=0}^{11} r_{f,t-k}. \end{aligned}$$

As an additional diagnostic on the asset pricing models, which to our knowledge is new, we employ ideas from realized variance literature (Andersen, Bollerslev and Diebold, 2007). Specifically, we compute log of the within-year realized variance

$$q_t^a = \log\left(\sum_{k=0}^{11} r_{d,t-k}^2\right) \quad (24)$$

as a measure of within-year variation (Andersen, Bollerslev and Diebold, 2007). A closely related volatility measure is

$$\text{std}_t^a = \sqrt{\sum_{k=0}^{11} r_{d,t-k}^2} \quad (25)$$

which we shall often use when reporting results.

In results reported in Section 6, we estimate models using an observation equation comprised of the four variables

$$y_t = \begin{pmatrix} d_t^a - c_t^a \\ c_t^a - c_{t-12}^a \\ P_{dt}^a - d_t^a \\ r_{dt}^a \end{pmatrix} \quad (26)$$

annually. If the observations were monthly, inclusion of r_{dt} in the observation equation (26) would induce a non-linear inter-temporal redundancy; however, at an annual frequency, the aggregation protocol just described implies that including r_{dt}^a adds additional information. We evaluate the LRR and HAB models by studying their implications for the dynamics of these four variables along with their implications for the dynamics of realized volatility, which is nowhere used in estimation.

5. DATA

5.1. Raw data

Our data set consists of annual observations 1929–2001. All variables, except population and consumption, are computed using monthly data from the Center for Research in Security Prices and then converted to the annual frequency. Annual real (\$1996) *per capita* consumption, non-durables and services, along with the mid-year population data are taken from the Bureau of Economic Analysis (BEA) web site.

To construct the annual *per capita* stock market valuation series, we start with the month-end combined nominal capitalizations of the New York Stock Exchange (NYSE) and the American Stock Exchange (AMEX), convert to \$1996 using the monthly Consumer Price Index (CPI), take the year-end value, and divide by the BEA population figure. To compute the annual dividend series, we use the difference between the nominal value-weighted return and the capital return (*i.e.* the return excluding dividend) to compute an implied monthly nominal dividend yield on the NYSE + AMEX. Applying this dividend yield to the preceding month's market capitalization gives an implied monthly nominal dividend series. This series is converted to a real (\$1996) monthly dividend series using the monthly CPI, aggregated over the year, and then divided by the BEA population figure. To compute the annual real return series, we use the monthly nominal value-weighted return on the NYSE + AMEX and the CPI to compute a monthly real geometric return, which is then cumulated over the year to form a real annual geometric return. The annual quadratic variation is the sum of the monthly squared real geometric returns.

For consistency with the presentation of the model, we let t denote the time index in months, so that $P_{dt}^a, t = 12, 24, \dots$, denotes the end-of-year *per capita* stock market value observations; $D_t^a, t = 12, 24, \dots$, denotes the annual aggregate *per capita* dividend observations; $C_t^a, t = 12, 24, \dots$, denotes the annual *per capita* consumption observations; $r_{dt}^a, t = 12, 24, \dots$, denotes the annual real geometric return observations; and $q_t^a, t = 12, 24, \dots$, denotes the log annual realized variation observations. While in principle it would be possible to use mixed monthly and annual data, this would entail a more complicated likelihood and additional assumptions. Our preference here is to use a more standard estimation approach.

Figure 1 shows time series plots of the annual observations on the logged series $p_{dt}^a = \log(P_{dt}^a)$, $d_t^a = \log(D_t^a)$, $c_t^a = \log(C_t^a)$, r_t^a , and $q_t^a = \log(Q_t^a)$. The three series p_{dt}^a , d_t^a , and c_t^a , are upwards trending series, while r_t^a and q_t^a appear stationary.

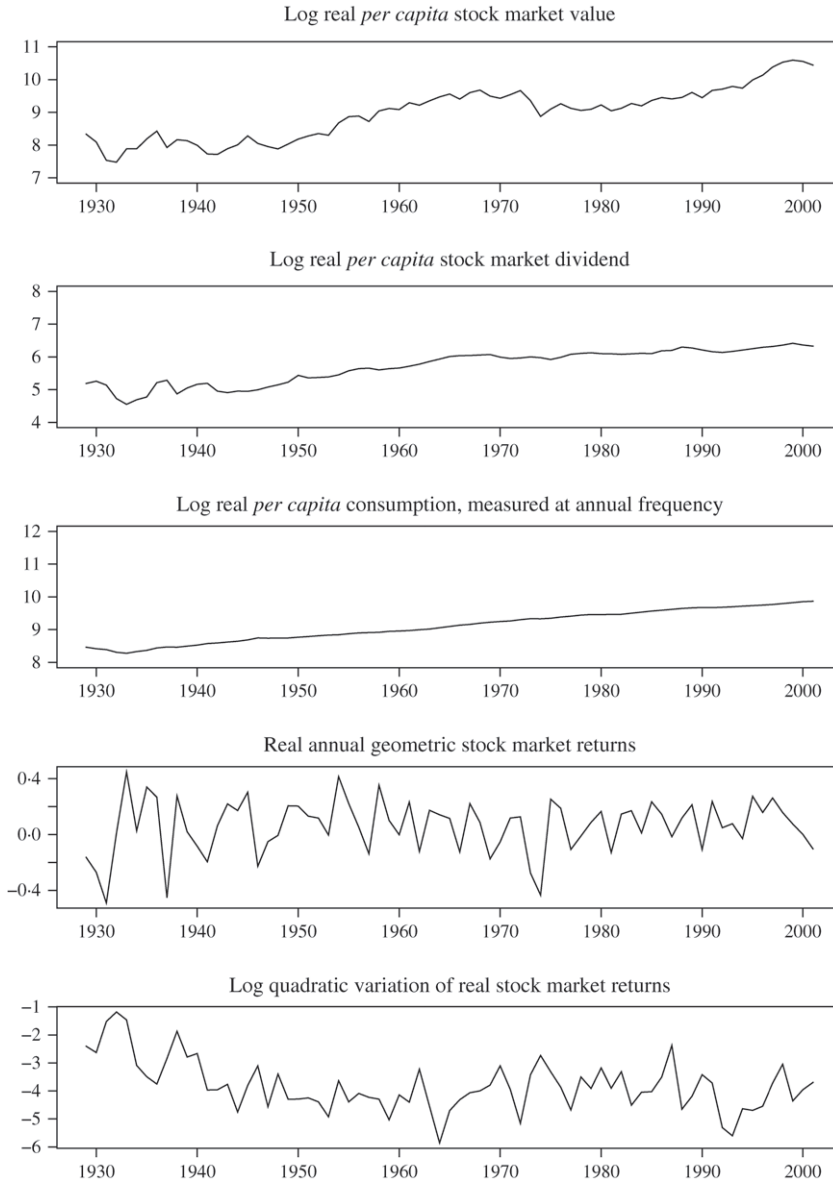


FIGURE 1

Raw data. In the notation of the text, the variables shown are $p_{dt}^a = \log(P_{dt}^a)$, $d_t^a = \log(D_t^a)$, $c_t^a = \log(C_t^a)$, r_{dt}^a , and $q_t^a = \log(Q_t^a)$, in order

5.2. Cointegrating relationships

One would expect there to be cointegrating relationships among the three trending variables. A simple regression of $(p_{dt}^a \ d_t^a \ c_t^a)$ on $(p_{d,t-12}^a \ d_{t-12}^a \ c_{t-12}^a)$, annual data 1930–2001, yields an autoregressive matrix with one eigenvalue nearly exactly equal to unity and two others about 0.90 in magnitude. Two eigenvalues separated from unity suggest that there are two cointegrating

relationships among the three variables. One relationship, which has been explored extensively in the literature and which we imposed *a priori* in the development of the LRR and HAB models above, is that the log price–dividend ratio $v_{dt}^a = p_{dt}^a - d_t^a$ is stationary.

It is natural to presume that there is also a cointegrating relationship between the log dividend and the log consumption variable. Thus, we conjecture the relationship

$$d_t^a - \lambda_{dc} c_t^a = I(0),$$

where λ_{dc} is a parameter that should be unity. We first constrain $\lambda_{dc} = 1$ and do a heteroskedasticity-robust augmented Dickey–Fuller test for a unit root in $d_t^a - c_t^a$. The Bayesian Information Criterion suggests that one lag in the augmented Dickey–Fuller equation is better than none or two, and for that model the unit-root t -statistic is 3.56 (p -value = 0.009), which is rather strong evidence for a cointegrating relationship between the log dividend and the log consumption variables with a coefficient of unity.

Another strategy is to estimate λ_{dc} by running a reduced-rank regression model (Anderson, 2001). We did this and the estimate was $\lambda_{dc} = 0.9497$, which is consistent with the unit-root test above.

In what follows, we shall take the series $p_{dt}^a - d_t^a$, $d_t^a - c_t^a$, and $c_t^a - c_{t-12}^a$ as the three jointly $I(0)$ variables embodied in the three trending series. Of course, the above analysis only identifies the two-dimensional subspace of \mathfrak{R}^3 that determines the two cointegrating relationships among (p_{dt}^a, d_t^a, c_t^a) . Put another way, given any non-singular 2×2 matrix times the vector

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{pmatrix} p_{dt}^a \\ d_t^a \\ c_t^a \end{pmatrix}$$

could, on statistical grounds, also be taken as the $I(0)$ linear combinations of the three trending series. However, the normalization above leads to very simple and easy-to-interpret variables: the log price–dividend ratio, the log dividend–consumption ratio, and consumption growth, all jointly $I(0)$.

6. MODEL PARAMETER ESTIMATES

In order to estimate the models described in Sections 2 and 3, one must handle the fact that the *ex ante* real risk-free rate of interest is not directly observable. Campbell (2002) notes that any reasonable asset pricing model must incorporate the indirect evidence that the mean risk-free rate r_{ft} is very low with low volatility. Campbell’s evidence suggests that the mean risk-free rate for the U.S. is 0.896% per annum. We want to impose Campbell’s empirical evidence on the parameter estimates; to do so, we impose a Campbell’s restriction on $\mathcal{E}(r_{ft})$ within the monthly simulations. In practical terms, we constrain $\mathcal{E}(r_{ft})$ to lie within a small band of approximately 50 annual basis points around 0.896% per annum. The reason for treating the constraint on the risk-free rate as a small inequality constraint is numerical: doing so gives the optimization software a small amount of slack and it makes feasible the task of finding start values for the numerical optimization. Treating it as an exact constraint is numerically infeasible. It bears noting that Campbell’s auxiliary evidence sharply restricts the parameter space and thereby challenges any model so constrained. Indeed, at the start of a run, we found it difficult and tedious to find start values for the parameters that generate simulated data consistent with Campbell’s evidence.

In results reported below, we use as the auxiliary model a four-variable VAR(1) at the annual frequency

$$y_t = b_0 + B y_{t-12} + e_t, \quad (27)$$

where

$$y_t = \begin{pmatrix} d_t^a - c_t^a \\ c_t^a - c_{t-12}^a \\ p_{d_t}^a - d_t^a \\ r_t^a \end{pmatrix}. \quad (28)$$

The definition of y_t above corresponds exactly with that of the observation equation (26) for the theoretical model. The period of the dependent variable in (27) is 1931–2001. We do not include the log realized variance q_t^a , but we do check how well the asset pricing model can match the dynamics of that variable.

In some initial work, we used an unrestricted 4×4 VAR(1). However, many of the estimates of the off-diagonal elements of B are statistically insignificant, and the diagonal elements clearly dominate; rather than adopting some arbitrary scheme for dropping variables, we use a VAR that is diagonal in both location and scale for estimation. The intercept and AR(1) coefficients are shown in the columns labelled “observed” in Table 4, which is discussed more fully below. The simulation estimator described in the Appendix remains consistent and asymptotically normal with this choice of the auxiliary model. The estimation begins with a simplified version where the LRR factor of the LRR model is turned off and all risks are presumed short run; this version corresponds closely to the model of Hall (1978). In this version, labelled the SRR model, the x_t process is set to zero, so consumption growth is a random walk with drift, and the driving processes in (3) are homoskedastic, so that the risk premia are constant. The restrictions are implemented by setting $\rho_x = 0$, by setting the elements of Ψ corresponding to x_t to zero, and by setting to zero the parameters governing the stochastic volatility in (5). The assumption that consumption growth is i.i.d. is quite common, for example, Campbell and Cochrane (1999), as is the assumption of homoskedastic driving variables. We use the term SRR because the only exposure of the dividend to consumption risk arises through innovation correlations. We use linear functional forms for the solution functions of the SRR model in the numerical solution of the model as described in the Appendix; we could find little indication of a need for quadratic terms. The use of a linear form is typical in the literature. A major difference is that we maintain cointegration between the consumption and the dividend process.

The second model, labelled the LRR model, is more general with underlying dynamics given by (3) above. This model incorporates a stochastic mean, $\mu_c + x_t$, for consumption growth, and it includes stochastic volatility as embodied in (5). A normalization for identification is that $b_c = b_s = b_x = 0$. Initial estimation was done with both linear and quadratic terms in the solution functions. Inspection of the results indicated that the only important quadratic term was the interaction term of x_t with the log-volatility variable, v_t , and that other quadratic terms could be set to zero without altering the results.

The initial work suggested that freely estimating the LRR model using an unconstrained 4×4 VAR(1) as the auxiliary model gives results almost exactly the same as those reported below, except that consumption volatility tends to be underestimated somewhat and therefore returns’ volatility is underestimated. We elected to fix the scale parameters of the consumption process in (4) and (5) to values that, after some experimentation, would give rise to consumption

TABLE 1

Parameter estimates: four-variable score

Parameter	SRR model		LRR model	
	Estimate	S.E.	Estimate	S.E.
μ_c	0.002178	0.000290	0.001921	0.000332
ρ_s	0.9892	0.0080	0.9583	0.1032
λ_{sx}			2.5004	16.9346
ρ_x			0.9871	0.0088
b_{cc}			0.1432	c
b_{ss}			0.8188	0.2788
b_{xx}			0.1100	c
ψ_{cc}	0.00646	0.00458	0.0034	c
ψ_{cs}	0.00146	0.01746		
ψ_{ss}	0.02487	0.00809	0.4109×10^{-06}	0.2203×10^{-05}
ψ_{xx}			0.000120	c
ρ_σ			0.9866	0.0011
δ	0.997488	0.004369	0.999566	0.000343
θ	-195.79	233.59	-12.2843	7.6243
ψ	2.00	c	2.00	c
γ	98.8969	116.7968	7.1421	3.8122
μ_{dc}	-3.3965	0.0428	-3.3857	0.0540
	$\chi^2(5) = 41.051(0.9 \times 10^{-7})$		$\chi^2(3) = 10.501(0.0148)$	

Notes: LRR, long-run risks; SRR, short-run risks. γ is a derived parameter computed from θ and ψ and "c" indicates a calibrated parameter.

volatility matching that of the observed data. These values are

$$\begin{aligned}
 b_{cc} &= 0.14320, \\
 b_{xx} &= 0.11000, \\
 \Psi_{cc} &= 0.00340, \\
 \Psi_{xx} &= 0.00012,
 \end{aligned} \tag{29}$$

and they remain fixed throughout the estimation, so they are calibrated parameters; Ψ_{ss} is a free parameter while the off-diagonal elements of Ψ in (5) are set to zero. Also, estimates of the elasticity of substitution parameter, ψ , typically land in the region 1.50–2.50, but the objective function is very flat in ψ , so we constrain $\psi = 2.00$. The remaining parameters, including the dynamic parameters ρ_x and ρ_σ and the parameter λ_{sx} , are freely estimated. Note that ρ_σ affects the dynamics of the price to dividend ratio and stock returns which identifies ρ_σ as can be seen from the expressions for the price to dividend and stock returns given in Bansal and Yaron (2004).

Table 1 displays parameter estimates for the SRR and LRR models along with the chi-squared measure of fit. The chi-squared statistics suggest that the SRR model is rejected, as might be expected, though the LRR model appears to give an adequate fit given the well-known tendency of this type of chi-squared test to over reject. Both point estimates of ρ_s are close to unity, indicating that shocks to the log dividend–consumption ratio are highly persistent in both models. The estimate of ρ_x is close to unity in the LRR model, which is evidence for a very persistent component x_t to consumption growth, though the small conditional standard deviation suggests that this component is small. Likewise, the estimate of ρ_σ is very close to unity, a finding consistent with all other empirical evidence on volatility persistence. Interestingly, the estimate

TABLE 2
Parameter estimates for the habit persistence (HAB) model

Parameter	Estimate	S.E.
μ_c	0.002116	0.000250
ρ_s	0.9719	0.0154
ψ_{cc}	0.006151	0.000896
ψ_{ss}	0.036503	0.007716
ρ_h	0.9853	0.002597
δ	0.9939	0.000526
γ	0.8386	0.2463
μ_{dc}	-3.3587	0.0380

$\chi^2(5) = 7.109(0.213)$

of ρ_σ close to unity is obtained only using mean dynamics without regard to the log realized variance process, q_t^a , as that process was not included in the VAR on which LRR was estimated.

Perhaps, the sharpest difference between SRR and LRR is the estimates of the risk aversion parameter γ , which is huge in the SRR model but much more reasonable in the LRR model. The difference can be traced directly to very different dynamics of the driving variables across the models. In the SRR model, consumption shocks are completely transient, while in the LRR model the transient consumption shocks are superimposed with the very persistent process x_t . Thus, in the latter model, the asset that pays the consumption endowment C_t is much riskier and thereby commands a higher risk premium, $\mathcal{E}_t(r_{c,t+1} - r_{ft})$, other things being equal. In addition, the near-permanent shocks to x_t affect the mean growth rate of the dividend in (4), the amount $\lambda_{s,x}$, estimated to be about 2.5. Thus, the asset paying the dividend stream D_t is likewise much riskier in the LRR model and commands a higher risk premium, $\mathcal{E}_t(r_{d,t+1} - r_{ft})$. Taken together, these characteristics of the dynamics of the driving processes imply that the LRR model can generate higher equity risk premia without appeal to an (implausible) high coefficient of risk aversion γ .

Another model we considered was one with long-run dynamics but with the utility function constrained to be the constant relative risk aversion (CRR) instead of the more general Epstein–Zin–Weil form considered in the LRR model. This estimation is achieved by setting $\theta = 1$ in (8) and estimating risk aversion, γ , which corresponds to $1/\psi$, along with the other free parameters. The empirical failures of CRR utility have been extensively documented in the literature, so we do not report the results in detail. The model is sharply rejected (p -value = 4.98×10^{-13}). Interestingly, the fitted model estimates a modest value for the risk aversion parameter, $\hat{\gamma} = 7.285$ (2.223), but it also estimates a very small value for μ_c , and thereby predicts an absurdly low value for mean annual average consumption growth. The low estimate of μ_c is the only way this model can satisfy the real interest constraint; of course, it cannot then generate a sufficient equity premium, as has been long understood. Table 2 shows parameter estimates and the chi-squared statistic for the model of HAB described in Section 3 above. Initial work suggested that Ψ_{cs} is difficult to estimate as a free parameter, so we constrain $\Psi_{cs} = 0$. Keep in mind that the consumption and dividend growth innovations are still correlated in (21); the parameter estimates imply correlations very close to 0.20, which is the value imposed by Campbell and Cochrane (1999). As seen in the table, the HAB model does well on the χ^2 criterion and the point estimates appear reasonable. As expected, the estimates of ρ_h and ρ_s are close to unity as is δ . The risk aversion parameter γ is estimated to be close to unity and is reasonably precisely estimated. A conventional Wald confidence interval would include a range of values above unity but would exclude the value 2.00 imposed by Campbell and Cochrane (1999).

7. CONTRASTS BETWEEN MODELS

7.1. Unconditional moments

Table 3 shows unconditional means and standard deviations computed from the data and the predicted values under the models SRR, LRR, and HAB. In this, and in subsequent tables, the predicted values should be regarded as population values implied by the model at the estimated parameter values corrupted by very small Monte Carlo noise. All three models agree rather closely with the data on the means and standard deviations of the log dividend–consumption ratio and also for consumption growth, except that the HAB model under predicts consumption growth volatility somewhat. Also, the three models agree quite closely with the data on the mean of the price–dividend ratio, though the SRR model seriously underestimates its volatility, while the LRR and HAB models are much closer, though still below that of the data.

All three models predict an annual return on the stock market, that is, the dividend asset, at just over 6% per year, which is consistent with the data. Likewise, since the mean of the risk free rate is tightly constrained, the three models also predict about the same equity premium. The mechanisms, however, are quite different. The unrealistically high degree of risk aversion of the SRR model generates an expected return on the consumption asset, $\mathcal{E}(r_{ct})$, of about 5.60% per year, which, via the correlation between consumption innovations and dividend innovations, is stepped up to a predicted value of 6.15% per year for the dividend asset. On the other hand, for the LRR model, the risk aversion parameter is much smaller, but the inherent higher riskiness of the consumption asset implies an expected return $\mathcal{E}(r_{ct})$ on the consumption asset of about 2.33% per year. Cointegration, innovation correlation, and the effects of x_t innovations implied by a positive estimate of λ_{sx} (see (3)) increase the expected return on the equity $\mathcal{E}(r_{dt})$ to the predicted value of 6.29% per year. Finally, as explained at length in Campbell and Cochrane (1999), the dynamics of the surplus consumption ratio make stocks rather unappealing to investors who therefore require a relatively large premium over cash. As seen from the table, the unconditional first two moments of the dividend asset and the consumption asset are about the same under the HAB model, and for the dividend asset, are in close agreement with the data, which is largely consistent with Campbell and Cochrane (1999, Table 2, p. 225).

One quite remarkable finding in Table 3 is how well both the LRR and the HAB models do in terms of matching the unconditional moments of the volatility measure defined by the square root of the quadratic variance variable (25). Both models are nearly right on the observed values despite the fact that this variable is not used in the estimation.

7.2. Conditional first and second moments

Table 4 shows observed and predicted univariate AR(1) models. It includes an AR(1) model for the log quadratic variance variable, q_t^a defined in (24), which was not used in estimation but is still available for model assessment of return volatility dynamics. The shortcomings of the SRR model are readily apparent. This model under predicts the persistence in annual consumption growth, $c_t^a - c_{t-12}^a$, and it misses very badly on volatility dynamics.

The LRR model, on the other hand, matches quite well nearly all the univariate features of the data. Interestingly, this model captures the serial correlation properties of return volatility quite well while only slightly underestimating the level of volatility. Overall, the HAB does just about as well, though there are two exceptions worth noting: first, the HAB model, just like the SRR model, predicts a value of 0.25 for the first-order autocorrelation in consumption growth, while the observed value is about 0.45. This under prediction can be traced directly to the presumption that consumption growth process lacks the long-term component x_t of the LRR model. Second,

TABLE 3
Comparison of model predictions with observed unconditional moments

	Observed		Predicted-SRR		Predicted-LRR		Predicted-HAB	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
log dividend-consumption ratio	-3.399	0.162	-3.414	0.163	-3.384	0.154	-3.37	0.148
Consumption growth (% per year)	$d_t^a - c_t^a$							
Price-dividend ratio	$100(\times c_t^a - c_t^a - 12)$		2.58	1.91	2.34	2.36	2.52	1.76
Return (% per year), dividend asset	$\exp(v_{dt}^a)$		28.86	4.18	27.47	8.01	27.75	7.037
	$100 \times r_{dt}^a$		6.15	2.82	6.29	16.00	6.54	16.89
$100 \times \sqrt{\text{quadratic variation}}$	$100 \times \text{std}_t^a$		3.22	0.64	14.95	8.21	14.41	9.69
Risk-free rate (% per year)	$100 \times r_{ft}^a$		0.39	0.00	0.78	0.41	1.07	0.99
Return (% per year), consumption asset	$100 \times r_{ct}^a$		5.60	2.33	2.34	3.95	6.60	16.99
Equity premium (% per year)	$100 \times r_{dt}^a - r_{ft}^a$		5.76	2.82	5.51	16.09	5.46	17.14

Notes: HAB, habit persistence; LRR, long-run risks; SRR, short-run risks. Observed values are sample statistics computed from annual data, 1930-2001; predicted values are computed from a long simulation from the indicated model.

TABLE 4

AR(1) models for each series

z_t	Observed		Predicted (SRR)		Predicted (LRR)		Predicted (HAB)	
	α_0	α_1	α_0	α_1	α_0	α_1	α_0	α_1
$d_t^a - c_t^a$	-0.6879	0.7992	-0.30281	0.91128	-1.0358	0.6939	-0.6784	0.7985
	0.3542	0.1034						
$c_t^a - c_{t-12}^a$	0.0116	0.4495	0.01894	0.2666	0.0119	0.4916	0.0185	0.2645
	0.0031	0.0909						
$p_{dt}^a - d_t^a$	0.4926	0.8542	0.3467	0.8966	0.7880	0.7598	0.54924	0.8326
	0.3678	0.1077						
r_{dt}^a	0.0675	0.0088	0.0582	0.0546	0.0673	-0.0694	0.0689	-0.0536
	0.0639	0.1454						
q_t^a	-0.8972	0.5352	-6.7936	0.0177	-1.73212	0.5621	-1.2400	0.7114
	0.2545	0.1275						

Notes: HAB, habit persistence; LRR, long-run risks; SRR, short-run risks. Each line of the table shows the coefficients for a regression of the form $z_t = \alpha_0 + \alpha_1 z_{t-1} + \epsilon_{zt}$ for the variable named in the first column. The regressions in the columns labelled observed are for the annual data where the span of the dependent variable is 1931–2001. The regressions in the columns labelled predicted are for a simulation of the model.

the persistence parameter of the realized volatility q_t^a is too large for the HAB model relative to the data whereas it is comparable to the data in the LRR model.

7.3. Data tracking

Figures 2–4 show time series plots of the observed data along with the predicted values from an unrestricted VAR estimate and the restricted VAR implied by the SRR, LRR, HAB models. The unrestricted VAR is a four-variable VAR estimated on y_t in (28) augmented by the log quadratic variance variable q_t^a , but set up to be block diagonal in (y_t, q_t^a) . The restricted VAR is the same thing except it is estimated on a long simulation (based on 50,000 months) from the model evaluated at the fitted parameters and is therefore the reprojected VAR; the one-step-ahead predictions are the reprojected data series. To ease interpretation, the values for q_t^a are converted from logs to levels and annualized.

The figures are quite conclusive on the abilities of the models to track the data. The reprojected series from the SRR model are a disaster, as seen in Figure 2. The reprojected series from the LRR and HAB models, however, are seen in Figures 3 and 4 to track very closely the predictable components of the log dividend–consumption ratio, consumption growth, the log price–dividend ratio, and the equity return. Indeed, the LRR and HAB predict well ahead of time the drop in equity returns that occurs between 1999 and 2001. Also, the LRR and HAB models track rather well the conditional volatility of the return as is evident from the bottom panels of these figures.

It seems quite striking that the two models with such different internal structures as the LRR and HAB can come to such a close agreement on the data, at least on a one-step-ahead basis; we thus investigate further the multi-step characteristics of these two models using selected pairwise projections of financial and macro-variables analogously to Wachter (2002).

7.4. Predictability regressions

Table 5 shows projections of the end-of-year log price–dividend ratio on contemporaneous and five (annual) lags of the consumption growth variable. The table suggests a rather weak link between the log price–dividend ratio and the history of consumption in both the observed data

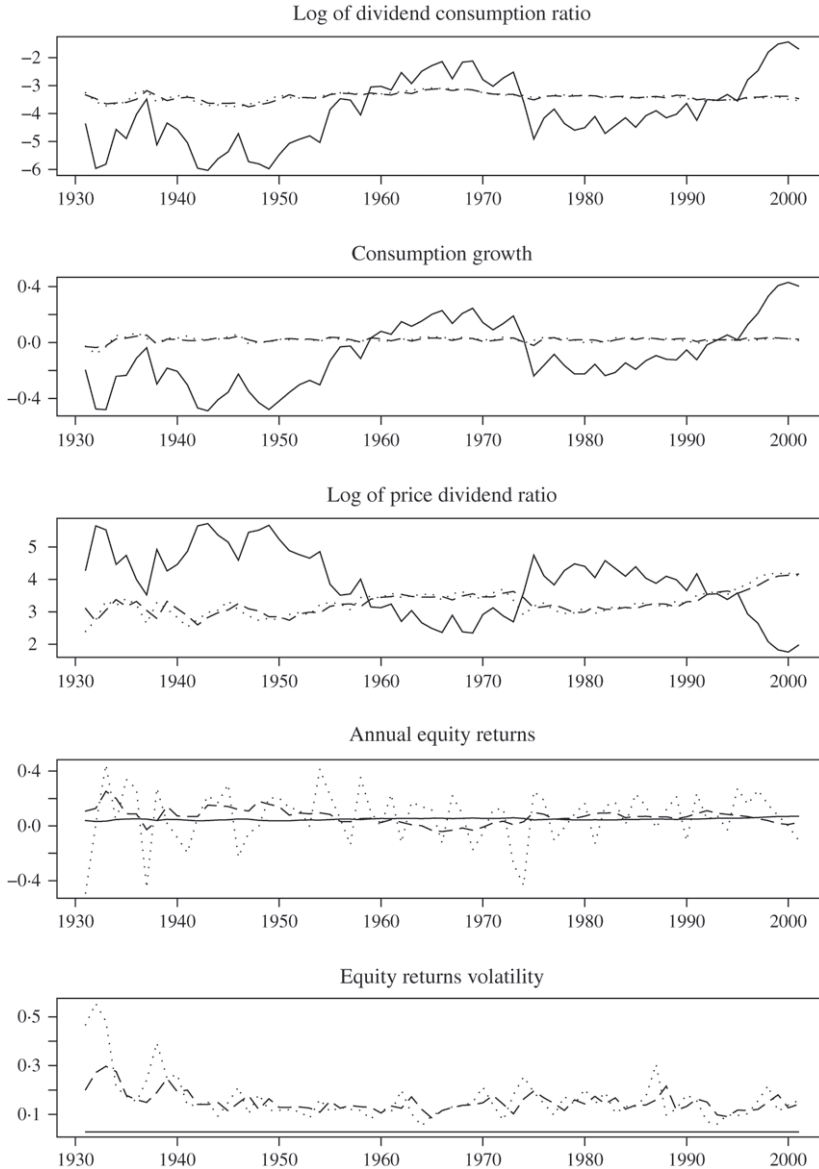


FIGURE 2

One-step-ahead forecasts of the data confronted by the short-run risks (SRR) model. The variables shown are $d_t^a - c_t^a$, $c_t^a - c_{t-12}^a$, $p_{dt}^a - d_t^a$, r_{dt}^a , and q_t^a of the text, in order. The dotted line is the data. The dashed line is a one-step-ahead forecast of a VAR fitted to the data. The solid line is the one-step-ahead forecast of a VAR computed from a simulation from the SRR model

and under the LRR model: the R^2 's are nearly negligible. On the other, the HAB model suggests a much tighter link than is consistent with the data. The link between consumption (its lags) and the price-dividend ratio helps discriminate across the two models.

Table 6 shows linear projections of cumulative n -year-ahead geometric stock returns on the log price-dividend ratio, for $n = 1, 2, \dots, 5$ years out. The R^2 's increase with the horizon for all three sets of projections. For those done on the observed data, the increase is rather mild,

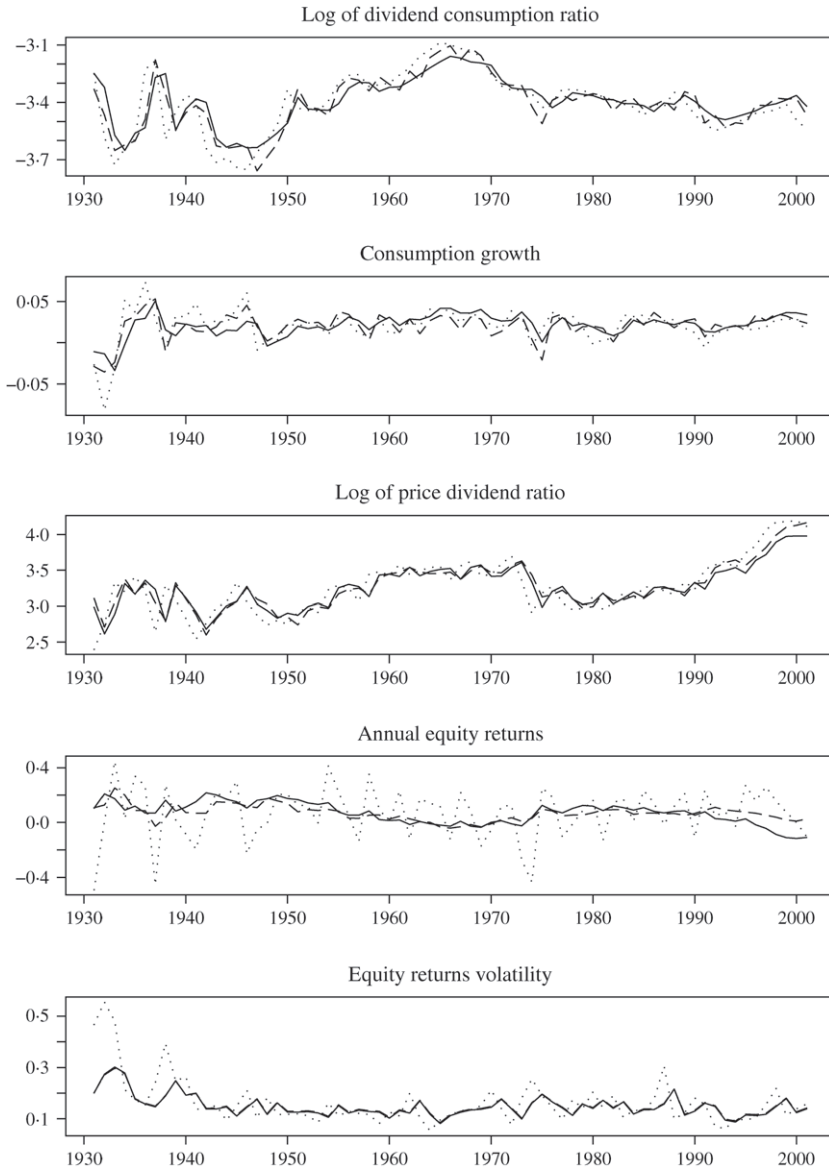


FIGURE 3

One-step-ahead forecasts of the data confronted by the long-run risks (LRR) model. The variables shown are $d_t^a - c_t^a$, $c_t^a - c_{t-12}^a$, $p_{dt}^a - d_t^a$, r_{dt}^a , and q_t^a of the text, in order. The dotted line is the data. The dashed line is a one-step-ahead forecast of a VAR fitted to the data. The solid line is the one-step-ahead forecast of a simulation from the LRR model

while for those done for the LRR model there is a steeper increase with horizon, while for the HAB model, there is an even steeper increase with horizon. Table 5, and to some extent Table 6, appears to contain evidence supportive of the LRR model over the HAB model. Recent papers in the literature also entertain other returns forecasting variables. Lettau and Ludvigson (2001) consider an empirical proxy for the consumption–wealth ratio and show that it has some ability

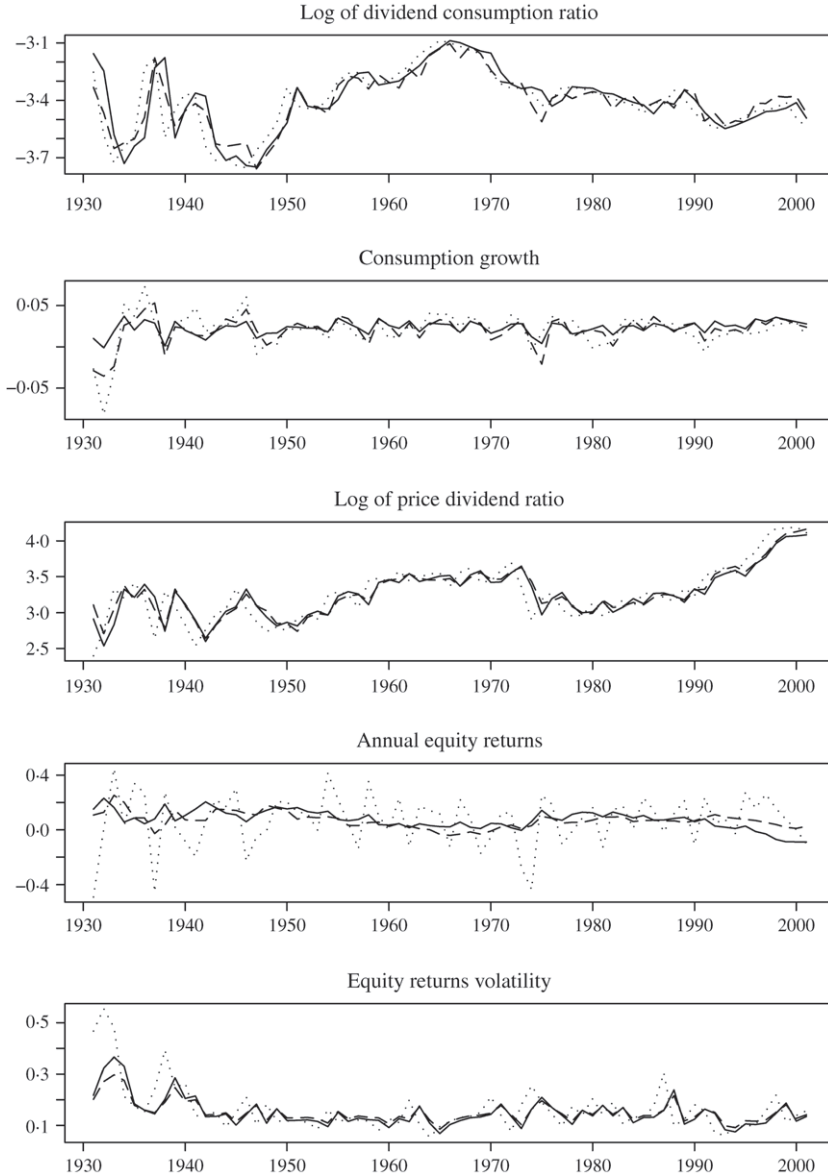


FIGURE 4

One-step-ahead forecasts of the data confronted by the habit persistence (HAB) model. The variables shown are $d_t^a - c_t^a$, $c_t^a - c_{t-12}^a$, $p_{dt}^a - d_t^a$, r_{dt}^a , and q_t^a of the text, in order. The dotted line is the data. The dashed line is a one-step-ahead forecast of a VAR fitted to the data. The solid line is the one-step-ahead forecast of a VAR computed from a simulation from the HAB model

to forecast returns. We evaluate this in the context of our models. For the LRR and HAB models, regressions of cumulative stock returns on the log wealth–consumption ratio, which is available in a theoretically consistent form directly within each model, reveal no appreciable differences in the R^2 's across LRR and HAB models and are similar in pattern and magnitude to those reported for the price–dividend ratio in Table 6. This suggests that regressions of cumulative stock returns on the log wealth–consumption ratio do not distinguish between the two models.

TABLE 5
Predictability projections: price–dividend ratios

	Observed		Predicted	
	Coefficients	S.E.	LRR	HAB
Intercept	3.4168	0.0055	3.2520	2.7183
Δc_t	-0.3021	0.1744	1.0716	6.7533
Δc_{t-24}	-1.0560	0.1926	0.0383	3.7708
Δc_{t-48}	-0.6209	0.1925	0.2521	3.4854
Δc_{t-60}	-0.3480	0.1926	-0.0176	3.1036
Δc_{t-72}	-0.1263	0.1926	-0.0442	2.4550
Δc_{t-84}	-0.2006	0.1744	-0.0857	2.7694
R^2	0.036		0.011	0.422

Notes: Shown above are the linear projections of the log price–dividend ratio, v_d on contemporaneous, and five annual lags of log consumption growth Δc . The period for the observed projection is 1935–2001. The predicted values are from long simulations from the long-run risks (LRR) model and the habit persistence (HAB) model.

TABLE 6
Long horizon predictability projections: cumulative future return on the price–dividend ratio

Horizon (years)	R^2		
	Observed	LRR	HAB
1	0.038	0.088	0.084
2	0.060	0.129	0.154
3	0.071	0.172	0.213
4	0.071	0.202	0.262
5	0.070	0.229	0.299

Notes: The table shows R^2 's from projections of cumulative annual geometric returns for 1, 2, ..., 5 years ahead onto the log price–dividend ratio for the observed data, 1935–2001, and for long simulations from the long-run risks (LRR) model and the habit persistence (HAB) model.

Predictability of consumption and dividend growth rates is also an issue of considerable interest. Papers by Ang and Bekaert (2001), Bansal, Khatchatrian and Yaron (2005), and Lettau and Ludvigson (2005b) show that dividend growth rates are predictable, particularly at long horizons. We consider cumulative n -year-ahead geometric growth rates of consumption and dividends on the log consumption–wealth ratio and the price–dividend ratio and that there are no substantive differences across models. For the consumption regressions, the R^2 are low (about 1% at the five-year horizon) in the case of the LRR model, as discussed in Bansal and Yaron (2004), because the R.H.S. variable is affected by several other state variables that diminish its ability to forecast future growth rates. In the HAB model, the lack of predictability is due to i.i.d. consumption growth. For the dividend regressions, one must keep in mind that we have imposed cointegration on all models, including the HAB model, see (21), unlike Campbell and Cochrane (1999). Consequently, we find that dividend growth rates are forecastable, even in the HAB model. Specifically, the error-correction variable s_t in (21) is a state variable that drives price–dividend ratios, and hence this valuation ratio can forecast dividend growth rates to some extent; for example, at the five-year horizon, the R^2 is about 20% for both models. Note that the error-correction state variable s_t captures similar intuition to the business cycle variable considered in Lettau and Ludvigson (2005b) for the predictability regressions.

One potential source of discrimination across models is the variance decomposition of the price–dividend ratio into portions attributable to variations in expected dividends and expected

TABLE 7
Consumption betas: annual frequency

	Data	LRR	HAB
$\hat{\beta}$	0.790	0.521	4.19
$\hat{\sigma}_{\hat{\beta}}$	1.124	0.176	0.17
\bar{R}^2	-0.006	0.006	0.19

Notes: HAB, habit persistence; LRR, long-run risks. For the observed data, the period of the dependent variable for the beta regression is annual data, 1929–2001, while the other two regressions are estimated on long annual simulated data sets from the estimated models.

returns (see Cochrane, 1991, for example). We find that even on this score, there is no great discrimination across models either. In our simulation, we find that 65% of the price–dividend variability is due to expected returns and about 35% due to dividend growth rates for the HAB model. The numbers for the LRR model are about 60% and 40% for expected returns and dividends, respectively—quite similar to those reported in Bansal and Yaron (2004). For the HAB model, as discussed above, cointegration implies that dividend growth rates are predictable in contrast to the HAB model considered in Campbell and Cochrane (1999). In the data, the point estimate for the expected return contribution to the long price–dividend variation is around 100%; however, these estimates in the data have fairly large standard errors (of the order of 40%), and the variance decompositions above can be regarded as consistent with the data.

7.5. Consumption–return relationships

Table 7 summarizes consumption beta regressions, that is, regressions of the stock market return on consumption growth estimated on the observed data and simulated data from the models. The contrasts between the LRR and the HAB models are quite sharp in this respect. In the data, the stock market’s consumption beta is 0.79 and the projection adjusted R^2 is zero. The HAB predicts a much higher beta of 4.19 and an adjusted R^2 of 19%. In contrast, the LRR predicts values quite comparable to those observed in the data. (The estimated standard errors using observed and simulated data are not comparable because the simulated sample sizes are far larger.) The evidence suggests that the HAB predicts an overly tight relationship between stock returns and consumption growth. Kiku (2006) uses the risk and return links in the LRR model to explain the empirical failure of standard consumption CAPM and the market CAPM in accounting for the differences in mean returns across assets.

A related exercise is to evaluate the ability of the two models to account for the findings of Hansen and Singleton (1982) regarding the empirical implausibility of the power utility specification for the stochastic discount factor. Table 8 summarizes the results of using GMM to estimate the subjective discount factor and risk aversion parameters using simulated data from each model at both the monthly and the annual frequencies. This is a form of encompassing, that is, a check on whether a maintained model can explain previous empirical findings. There are three Euler equation errors corresponding to the risk-free rate, the dividend asset, and the consumption asset. GMM was used to minimize the length of the vector of average pricing errors using the appropriate covariance matrix to form the over identification test, which is chi-squared on one degree of freedom. Not surprisingly, both models produce rather large estimates of risk aversion. Note, however, that LRR generates far larger chi-squared test statistics than HAB, suggesting that it accounts better for previous empirical findings using power utility. While this exercise is similar

TABLE 8
GMM estimation of the power utility model

	LRR	HAB
Annual		
$\chi^2(1)$	242.65	2.61
<i>p</i> -value	0.00	0.11
$\hat{\beta}$	1.38	2.39
$\hat{\gamma}$	16.62	50.40
Monthly		
$\chi^2(1)$	26.28	1.49
<i>p</i> -value	0.00	0.22
$\hat{\beta}$	1.01	1.04
$\hat{\gamma}$	9.60	23.22

Notes: HAB, habit persistence; LRR, long-run risks. β and γ represent the subjective time preference and curvature parameter for power utility, respectively. Parameter estimates and test statistics are based on simulations from the estimated models.

in spirit to that conducted in Lettau and Ludvigson (2005a), there are important differences: our test assets differ from theirs, and the model implementations are different. For example, we impose cointegration which affects predictability of dividend growth rates and returns as discussed above.

The consumption beta and the power utility-based diagnostic highlight that the HAB model, through the habit function, emphasizes the risks associated with high-frequency movements in consumption to account for the risk and return relation. The data do not support this channel. The LRR model on the other hand emphasizes the low-frequency movements in consumption as the key channel for risks in asset markets; these risks are handsomely compensated while high-frequency risks do not receive significant risk compensation. This feature of the LRR explains its smaller consumption beta and the large chi-square statistic discussed above. Bansal, Dittmar and Lundblad (2005) and Kiku (2006), for example, utilize this feature of the LRR model to examine cross-sectional differences in risk premia between assets.

7.6. Derivatives prices

We now examine the implications of the two models for derivatives pricing, and for simplicity we restrict the analysis to standard European call and put options written on the stock and consumption assets. The computations are numerically intensive but conceptually very straightforward because the functional form of each model's asset pricing kernel is generated as a by-product of the estimation. These call and put derivatives do not directly correspond to actual real-world traded securities, and so the computed prices cannot be compared to actual data. Nonetheless, the contrasts between the two models' derivative prices can provide additional insight into their internal structures.

As is well understood, the price P_{gt} at time t of a financial instrument with pay-off G_{t+k} at $t+k$ is given by

$$P_{gt} = \mathcal{E}_t \left[\exp \left(\sum_{j=1}^k m_{t+j-1,t+j} \right) G_{t+k} \right], \quad (30)$$

where $m_{t,t+1}$ is the log of the one-period marginal rate of substitution. Define the pricing operator

$$\mathcal{V}_{tk}(\cdot) = \mathcal{E}_t \left[\exp \left(\sum_{j=1}^k m_{t+j-1,t+j} \right) (\cdot) \right], \quad (31)$$

so the valuation can be expressed simply as $P_{gt} = \mathcal{V}_{tk}(G_{t+k})$.

Consider a European call option written on the stock price process P_{dt} with strike price X and expiration k steps ahead. The pay-off $G_{d,t+k} = \max(P_{d,t+k} - X, 0)$ expressed in relative terms is

$$\frac{G_{d,t+k}}{P_{dt}} = \max \left(\frac{P_{d,t+k}}{P_{dt}} - x_t^*, 0 \right),$$

where $x_t^* = X/P_{dt}$ denote the strike-to-underlying ratio of the call. Using the pricing operator (31), the relative price of the call is

$$v_{d,kt} = \mathcal{V}_{tk} \left(\frac{G_{d,t+k}}{P_{dt}} \right),$$

which is computed using the solution technique described in the Appendix.

Evidently, the computations proceed in analogous manner for a European put option starting from

$$G_{p,t+k} = \max(X - P_{d,t+k}, 0).$$

A put–call parity relationship holds for these options, but the formula is awkward because of the stochastic dividend stream.

Table 9 shows for both the LRR and the HAB models the unconditional means of the implied relative call and put prices written on the dividend asset P_{dt} . The expiration dates range from 1 to 12 months ahead and the strike-to-underlying ratios are 0.99, 1.00, and 1.01. The derivative prices are in percent relative to the level of the stock price at the time the option is written. For example, under the LRR model, a one-month call option with strike-to-underlying ratio of 1.00 costs 1.65% of the price of a share of the stock, with a similar interpretation for the other reported average prices. The average put prices might seem high relative to the average call prices but one must keep in mind the dividend, which tends to increase the value of puts relative to calls. Overall, the average prices shown in Table 9 exhibit the usual properties of call and put options. There does not seem to be much difference between the prices computed under the LRR and HAB models, indicating that the models are generally in agreement on the prices of call and put derivatives written on the dividend asset.

For derivatives written on the consumption asset, P_{ct} , it is helpful to keep in mind that P_{ct} is just wealth in the economies considered above, so derivatives written on P_{ct} represent derivatives on wealth. Consider a European call option written on $P_{c,t+k}$. It has pay-off $G_{c,t+k} = \max(P_{c,t+k} - X, 0)$ that can be normalized as

$$\frac{G_{c,t+k}}{P_{ct}} = \max \left(\frac{P_{c,t+k}}{P_{ct}} - x_t^*, 0 \right),$$

where now x_t^* represents the wealth strike-to-underlying of the call. The relative price of this call is

$$v_{w,kt} = \mathcal{V}_t \left(\frac{G_{t+k}}{P_{dt}} \right)$$

TABLE 9
Average put and call prices on the stock

Strike-to-underlying	Call			Put		
	0.99	1.00	1.01	0.99	1.00	1.01
Long-run risks (LRR)						
Expiration (months)						
1	2.19	1.65	1.22	1.38	1.85	2.41
2	2.79	2.28	1.84	2.18	2.67	3.23
3	3.25	2.75	2.31	2.83	3.33	3.89
4	3.61	3.13	2.69	3.38	3.89	4.45
5	3.93	3.45	3.01	3.88	4.40	4.96
6	4.21	3.73	3.30	4.34	4.86	5.42
7	4.46	3.99	3.56	4.76	5.29	5.85
8	4.69	4.23	3.80	5.15	5.68	6.25
9	4.90	4.44	4.01	5.54	6.07	6.64
10	5.08	4.63	4.21	5.92	6.45	7.02
11	5.25	4.80	4.38	6.28	6.82	7.39
12	5.40	4.96	4.54	6.64	7.19	7.76
Habit persistence (HAB)						
Expiration (months)						
1	2.06	1.51	1.12	1.31	1.76	2.37
2	2.63	2.09	1.67	2.13	2.60	3.17
3	3.05	2.52	2.08	2.82	3.29	3.85
4	3.39	2.87	2.42	3.41	3.88	4.43
5	3.68	3.16	2.70	3.97	4.44	4.98
6	3.94	3.42	2.96	4.50	4.97	5.51
7	4.18	3.66	3.20	5.02	5.50	6.03
8	4.40	3.88	3.42	5.54	6.02	6.54
9	4.59	4.08	3.61	6.02	6.50	7.03
10	4.78	4.27	3.80	6.50	6.98	7.50
11	4.95	4.44	3.97	6.96	7.44	7.97
12	5.10	4.60	4.13	7.44	7.92	8.45

Notes: Prices are percent of the stock price at the time the option is written.

and would be interpreted as the fraction of wealth that must be paid for a call with pay-off $\max\left(\frac{P_{c,t+k}}{P_{ct}} - x_t^*, 0\right)$. The analysis is completely analogous for wealth puts starting from the scaled pay-off $\max\left(x_t^* - \frac{P_{c,t+k}}{P_{ct}}, 0\right)$.

Table 10 shows the computed average relative prices for puts and calls written on the consumption asset P_{ct} expressed as a percentage of total wealth P_{ct} . Thus, the LRR model implies that, on average, the cost of a 12-month at-the-money put option on wealth is 1.26% of current wealth. By way of contrast, the HAB model implies that the same put option would cost 7.93% of current wealth.

There are sharp differences between the wealth derivative implications of the two models. The HAB model implies substantially higher costs across the board for both the call and the put derivatives. A factor accounting for the differences appears to be the different volatilities of the return on the consumption asset under the two models. From Table 3 above, the HAB model implies that the annual volatility on the return on the consumption asset is 18.12% per year, which is about the same as that on the dividend asset. In contrast, the LRR model implies the much lower value of 3.95% per year for the volatility of the consumption asset. Thus, the HAB model generates about the same values for derivatives written on either the dividend or the consumption asset, while the LRR model generates much smaller values for derivatives written on the consumption asset. Put another way, using derivatives to protect or hedge the value of the

TABLE 10
Average put and call prices on wealth

Strike-to-underlying	Call			Put		
	0.99	1.00	1.01	0.99	1.00	1.01
Long-run risks (LRR)						
Expiration (months)						
1	1.18	0.43	0.10	0.11	0.36	1.03
2	1.35	0.63	0.22	0.22	0.50	1.09
3	1.51	0.80	0.35	0.32	0.61	1.15
4	1.65	0.95	0.47	0.40	0.70	1.21
5	1.79	1.09	0.59	0.48	0.78	1.27
6	1.92	1.23	0.71	0.55	0.85	1.32
7	2.05	1.36	0.82	0.62	0.92	1.38
8	2.18	1.48	0.93	0.68	0.98	1.43
9	2.30	1.61	1.04	0.75	1.05	1.48
10	2.42	1.72	1.15	0.81	1.12	1.54
11	2.53	1.84	1.26	0.88	1.18	1.60
12	2.65	1.95	1.37	0.96	1.26	1.66
Habit persistence (HAB)						
Expiration (months)						
1	2.07	1.52	1.13	1.33	1.77	2.38
2	2.64	2.10	1.68	2.15	2.61	3.19
3	3.06	2.53	2.09	2.84	3.31	3.86
4	3.40	2.88	2.43	3.43	3.91	4.45
5	3.69	3.17	2.72	3.99	4.47	5.01
6	3.95	3.44	2.97	4.53	5.00	5.54
7	4.19	3.68	3.21	5.05	5.53	6.06
8	4.41	3.90	3.43	5.57	6.05	6.58
9	4.61	4.10	3.62	6.07	6.55	7.07
10	4.79	4.28	3.81	6.54	7.03	7.55
11	4.96	4.45	3.98	7.01	7.49	8.01
12	5.12	4.61	4.14	7.48	7.97	8.49

Notes: Prices are percent of wealth at the time the option is written.

stock index would cost about the same in either model, but using derivatives to protect wealth would be much cheaper in an LRR economy than in an HAB economy.

8. CONCLUSION

A simulated method of moments method proposed by Smith (1993) was used to estimate consumption-based rational expectations asset pricing models using aggregate data on consumption, dividends, and stock prices from 1929 to 2001, taking cognizance of the cointegrating relationships among these variables both with respect to the data and with respect to the dynamics of the models themselves.

The first model, patterned after Bansal and Yaron (2004) and termed LRR, has Epstein–Zin–Weil preferences and slowly time-varying growth dynamics for the consumption and dividend driving processes. A special case of this model, termed SRR, is patterned after Hall (1978), and has the same preferences but constant growth dynamics for the consumption and dividend driving processes. The second, patterned after Campbell and Cochrane (1999) and termed HAB, has power utility preferences with external habit formation and constant growth dynamics for the consumption and dividend driving processes. For these models, a practicable simulation-based numerical method for pricing dividend flows, consumption flows, and derivatives on the stock market and on wealth is developed and implemented in the paper.

The SRR was overwhelmingly rejected by the test of overidentifying restrictions, and it did a poor job of matching conditional and unconditional moments of the data. It was dismissed from further consideration.

On the other hand, both LRR and HAB did well on the tests of overidentification and did an excellent job of matching conditional moments, unconditional moments, and first-order dynamics in general. The models account for the equity premium, the long swings in stock market valuations relative to underlying cash flows, the high level of stock market volatility, and the persistence of long-term swings in stock market volatility. The models were further validated by considering the match of predicted to realized conditional and unconditional moments used in estimation and also to a realized variance measure, newly developed in this paper, that had been held out from the data used for estimation. Also of interest is the fact that both models predicted the stock market downturn around 2000 whereas VAR forecasts did not. According to conventional determinants of the adequacy of asset pricing models, both LRR and HAB appear successful; one can conclude that there is no need to abandon rationality, and behavioural models might be unnecessary.

A choice between LRR and HAB must be made on the basis of a more extensive scrutiny of their structural characteristics. Regressions designed to reveal differences in multi-step-ahead dynamics were undertaken. These are regressions of the log price–dividend ratio on current and lagged consumption growth and regressions of future stock returns on the price–dividend ratio. Regressions for LRR simulations agreed more closely with regressions for the data than did regressions for HAB simulations. Also, consumption beta regressions and encompassing GMM estimations suggested that LRR better reflects the comovements between consumption and returns. Finally, analysis of prices of put and call option written on the stock market and wealth suggests that HAB makes little distinction between stock market wealth and overall wealth, while LRR finds a sharp distinction. A bottom-line decision is very difficult to make, but on the basis of the more extensive scrutiny one concludes that probably the LRR model is preferred to the habit model.

APPENDIX: MODEL SOLUTION TECHNIQUE AND ESTIMATION METHODOLOGY

We use a simulation-based methodology based on Smith (1993) and implemented using a score-based estimation strategy proposed by Gallant and Tauchen (1996). Central to the methodology is the use of long simulations to compute unconditional expectations of the functions of the state vector involving one or more leads and lags. In particular, an expectation of a random variable of the form $g(u_{t+1}, u_t)$, which depends on a contemporaneous and a one-step-ahead value of the state vector, can be computed from a simulation $\{\hat{u}_t\}_{t=1}^{N+1}$ as

$$\mathcal{E}(g) \doteq \frac{1}{N} \sum_{t=1}^N g(\hat{u}_{t+1}, \hat{u}_t)$$

to any desired degree of accuracy by taking N sufficiently large.

Because one must bear the cost of generating the simulation $\{\hat{u}_t\}_{t=1}^{N+1}$ for the purpose of estimation anyway, it becomes convenient to use this same simulation to determine the pricing functions $v_c(u)$, $v_d(u)$, and $r_f(u)$ using a Bubnov–Galerkin method (Miranda and Fackler, 2002, pp. 152–153). Consider, for example, the determination of $v_c(u)$ for the LRR model, which, like all other policy functions in this paper, can be adequately approximated by the quadratic $v_c(u) \doteq a_0 + a_1' u + u' A_2 u$. Using (10) and (11), $v_c(u)$ must satisfy the conditional expression

$$\mathcal{E}\{v_c(u_t) - M(u_{t+1}, u_t)[1 + v_c(u_{t+1})] \exp(\Delta c_{t+1}) \mid u_t = u\} = 0 \tag{32}$$

for all $u \in \mathfrak{R}^3$, where

$$M(u_{t+1}, u_t) = \delta^\theta \exp[-(\theta/\psi)\Delta c_{t+1} + (\theta - 1)r_c(u_{t+1}, u_t)],$$

$$r_c(u_{t+1}, u_t) = \log \left[\frac{1 + v_c(u_{t+1})}{v_c(u_t)} \exp(\Delta c_{t+1}) \right].$$

Let

$$g(u_{t+1}, u_t) = \{v_c(u_t) - M(u_{t+1}, u_t)[1 + v_c(u_{t+1})] \exp(\Delta c_{t+1})\} Z_t,$$

where $Z_t = [1, u_t', \text{vech}'(u_t u_t')]'$ and $\text{vech}(A)$ denotes the vector comprised of the elements of the upper triangle of a symmetric matrix A . Then, by the law of iterated expectations, the conditional expression (32) implies that the unconditional expression $\mathcal{E}g = 0$ must also be satisfied by $v_c(u)$. Let $g(u_{t+1}, u_t, \alpha_c)$ denote $g(u_{t+1}, u_t)$ with $v_c(u)$ replaced by its quadratic approximation and let $\alpha_c = [a_0, a_1', \text{vech}'(A_2)]'$ denote the 10 distinct coefficients of the quadratic approximation. Because both α_c and Z_t are vectors of length 10, the unconditional expression $\mathcal{E}g(\cdot, \alpha_c) = 0$ becomes a system of 10 non-linear equations that can be solved for α_c . As integrals are computed by averaging over a long simulation $\{\hat{u}_t\}_{t=1}^{N+1}$, it is actually the system

$$\frac{1}{N} \sum_{t=1}^N g(\hat{u}_{t+1}, \hat{u}_t, \alpha_c) = 0 \quad (33)$$

that is to be solved for α_c . One solves for $v_d(u)$ and $r_f(u)$ similarly. As the marginal rate of substitution $M(u_{t+1}, u_t)$ depends on $v_c(u_{t+1})$ and $v_c(u_t)$, one must solve for $v_c(u)$ first. The solution strategy for HAB is exactly the same using (15) as the marginal rate of substitution and (23) as the state vector. Because $v_c(u)$ is needed to price macro-risks as described later, we compute it for the habit model. At this stage of the computations, both the simulation $\{\hat{u}_t\}_{t=1}^{N+1}$ and the values $v_c(\hat{u}_t)$, $v_d(\hat{u}_t)$, and $r_f(\hat{u}_t)$ at the monthly frequency become available. The corresponding values for $\{\hat{y}_t\}_{t=12,24,\dots}^{N/12}$ at the annual frequency can now be computed by applying the expressions for aggregation set forth in Section 4.

Let the transition density of the VAR be written as $f(y_t | y_{t-12}, \theta)$, $t = 12, 24, \dots$; let $\tilde{\theta}$ denote the maximum likelihood estimate of θ computed from the data $\{\tilde{y}_t\}_{t=12,24,\dots}^{n/12}$. Collect the parameters of one of the structural models, e.g. LRR, into a vector denoted ρ . Simulations from the structural model will follow a stationary density that we denote by $p(y_t, y_{t-12} | \rho)$. Consider the unconditional moment functions

$$m(\rho, \tilde{\theta}) = \int \int \frac{\partial}{\partial \theta} \log f(y_t | y_{t-12}, \tilde{\theta}) p(y_t, y_{t-12} | \rho) dy_t dy_{t-12}. \quad (34)$$

The integral in (34) is computed by specifying ρ , simulating the structural model, and averaging over the simulation as discussed above. Using these moment conditions, ρ can be estimated by minimizing the GMM criterion

$$s(\rho) = m'(\rho, \tilde{\theta}) (\tilde{X})^{-1} m(\rho, \tilde{\theta}), \quad (35)$$

where

$$\tilde{X} = \sum_{t=12,24,\dots}^{n/12} \left\{ \frac{\partial}{\partial \theta} \log f(\tilde{y}_t | \tilde{y}_{t-12}, \tilde{\theta}) \right\} \left\{ \frac{\partial}{\partial \theta} \log f(\tilde{y}_t | \tilde{y}_{t-12}, \tilde{\theta}) \right\}'. \quad (36)$$

The Efficient Method of Moments (EMM) estimator is $\hat{\rho}$ that maximizes $s(\rho)$. It is \sqrt{n} -consistent and asymptotically normally distributed as proved in Gallant and Tauchen (1996) under regularity conditions set forth there. The most important of these conditions is an identification condition that, excepting pathological examples, will be satisfied when $\dim(\theta) \geq \dim(\rho)$, which is essentially the requirement that the number of moments used to define a method of moments estimator must equal or exceed the number of parameters estimated. Gouriéroux, Monfort and Renault (1993) show that, for the particular choice of auxiliary model $f(y_t | y_{t-12}, \theta)$ used here, equation (27), the EMM estimator is asymptotically equivalent to Smith's (1993) estimator. The Gallant and Tauchen regularity conditions do not require that the data follow the distribution $f(y_t | y_{t-12}, \theta)$ because the proof strategy relies only on quasi-maximum likelihood estimation theory (Gallant, 1987). However, when using (36) as the weighting matrix, it is a good practice to subject $f(y_t | y_{t-12}, \theta)$ to specification tests; we used BIC. With respect to setting off diagonal terms of $f(y_t | y_{t-12}, \theta)$ to zero in Section 6, a good analogy is to the problem of bias caused by weak instruments. Removing the moments corresponding to the poorly estimated off-diagonal moments is like removing weak instruments and therefore relatively conservative. One would expect it to reduce bias rather than increase it. We possibly pay a price in efficiency.

The EMM estimator is a GMM estimator whence $ns(\hat{\rho})$ is asymptotically distributed as a (non-central) chi-square random variable on $\dim(\theta) - \dim(\rho)$ degrees of freedom. When $m(\rho^o, \theta^o) = 0$, where θ^o and ρ^o are the almost sure limits of $\tilde{\theta}$ and $\hat{\rho}$, expressions for which are given in Gallant and Tauchen (1996), then $ns(\hat{\rho})$ follows the central chi-square distribution. If the data do, in fact, follow the structural model $p(y_t, y_{t-12} | \rho)$, then $m(\rho^o, \theta^o) = 0$. Thus, comparing $ns(\hat{\rho})$ to the chi-square critical value, as we do in Section 6, is a test of model specification. This logic is exactly the same as the logic of the GMM test of overidentifying restrictions.

Acknowledgements. This work was supported by the National Science Foundation Grant SES 0438174.

REFERENCES

- ANDERSEN, T. G., BOLLERSLEV, T. and DIEBOLD, F. X. (2007), "Parametric and Nonparametric Volatility Measurement", in L. P. Hansen and Y. Aït-Sahalia (eds.) *Handbook of Financial Econometrics* (Amsterdam: North Holland, forthcoming).
- ANDERSON, T. W. (2001), "Reduced Rank Regression in Cointegrated Models", *Journal of Econometrics*, **106**, 203–216.
- ANG, A. and BEKAERT, G. (2001), "Stock Return Predictability: Is it There?" (NBER Working Paper No. 8207, National Bureau of Economic Research).
- BANSAL, R., DITTMAR, R. and LUNDBLAD, C. (2005), "Consumption, Dividends, and the Cross Section of Equity Returns", *Journal of Finance*, **60**, 1639–1972.
- BANSAL, R., KHATCHATRIAN, V. and YARON, A. (2005), "Interpretable Asset Markets?", *European Economic Review*, **49**, 531–560.
- BANSAL, R. and LUNDBLAD, C. (2002), "Market Efficiency, Asset Returns, and the Size of the Risk Premium in Global Equity Markets", *Journal of Econometrics*, **109**, 195–237.
- BANSAL, R. and YARON, A. (2004), "Risks For the Long Run: A Potential Resolution of Asset Pricing Puzzles", *Journal of Finance*, **59**, 1481–1509.
- CAMPBELL, J. Y. (2002), "Consumption Based Asset Pricing" (Working Paper, Department of Economics, Harvard, Cambridge, MA).
- CAMPBELL, J. Y. and COCHRANE, J. (1999), "By Force of Habit: A Consumption-based Explanation of Aggregate Stock Market Behavior", *Journal of Political Economy*, **107**, 205–251.
- CECCHETTI, S. G., LAM, P.-S. and MARK, N. C. (1993), "The Equity Premium and the Risk-Free Rate: Matching the Moments", *Journal of Monetary Economics*, **31**, 21–45.
- COCHRANE, J. (1991), "Explaining the Variance of Orice-Dividend Ratios", *Review of Financial Studies*, **5**, 243–280.
- EPSTEIN, L. G. and ZIN, S. (1989), "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica*, **57**, 937–969.
- EPSTEIN, L. G. and ZIN, S. (1991), "Substitution, Risk Aversion and the Temporal Behavior of Consumption and Asset Returns II: An Empirical Analysis", *Journal of Political Economy*, **99**, 1263–1286.
- EVANS, M. (1998), "Dividend Variability and Stock Market Swings", *Review of Economic Studies*, **65**, 711–740.
- GALLANT, A. R. (1987) *Nonlinear Statistical Models* (New York: Wiley).
- GALLANT, A. R. and TAUCHEN, G. (1996), "Which Moments to Match?", *Econometric Theory*, **12**, 657–681.
- GALLANT, A. R. and TAUCHEN, G. (1998), "Reprojecting Partially Observed Systems with Application to Interest Rate Diffusions", *Journal of the American Statistical Association*, **93**, 10–24.
- GALLANT, A. R. and TAUCHEN, G. (2007), "Simulated Score Methods and Indirect Inference for Continuous Time Models", in L. P. Hansen and Y. Aït-Sahalia (eds.) *Handbook of Financial Econometrics* (Amsterdam: North Holland, forthcoming).
- GOURIÉROUX, C., MONFORT, A. and RENAULT, E. (1993), "Indirect Inference", *Journal of Applied Econometrics*, **8**, S85–S118.
- HALL, R. E. (1978), "Stochastic Implications of the Life Cycle Permanent Income Hypothesis: Theory and Evidence", *Journal of Political Economy*, **86**, 971–988.
- HANSEN, L. P. and SINGLETON, K. J. (1982), "Generalized Instrumental Variable Estimation Estimators of Nonlinear Rational Expectations Models", *Econometrica*, **50**, 1269–1286.
- KIKU, D. (2006), "Is the Value Premium a Puzzle?" (Working Paper, Wharton Business School, University of Pennsylvania, Philadelphia, PA).
- LETTAU, M. and LUDVIGSON, S. C. (2001), "Consumption, Aggregate Wealth, and Expected Stock Returns", *Journal of Finance*, **56**, 815–849.
- LETTAU, M. and LUDVIGSON, S. C. (2005a), "Euler Equation Errors" (Working Paper, Stern School of Business, New York University, New York, NY).
- LETTAU, M. and LUDVIGSON, S. C. (2005b), "Expected Returns and Expected Dividend Growth", *Journal of Financial Economics*, **76**, 583–626.
- MEHRA, R. and PRESCOTT, E. C. (1985), "The Equity Premium: A Puzzle", *Journal of Monetary Economics*, **15**, 145–161.
- SMITH, A. A. (1993), "Estimating Nonlinear Time Series Models Using Simulated Vector Autoregressions", *Journal of Applied Econometrics*, **8**, S63–S84.
- WACHTER, J. A. (2002), "Habit Formation and Returns on Bonds and Stocks" (Working Paper, Stern School of Business, New York University, New York, NY).
- WEIL, P. (1989), "The Equity Premium Puzzle and the Risk Free Rate Puzzle", *Journal of Monetary Economics*, **24**, 401–421.