Journal of Monetary Economics 31 (1993) 21–45. North-Holland

The equity premium and the risk-free rate

Matching the moments*

Stephen G. Cecchetti
Ohio State University, Columbus, OH 43210, USA
National Bureau of Economic Research, Cambridge, MA 02138, USA

Pok-sang Lam and Nelson C. Mark
Ohio State University, Columbus, OH 43210, USA

Received June 1991, final version received December 1992

We investigate the ability of a representative agent model with time-separable utility to explain the first and second moments of the risk-free rate and the return to equity. We generalize the standard calibration methodology by accounting for the uncertainty in both the sample moments to be explained and the estimated parameters to which the model is calibrated. We find that the first moments of the data can be matched for a wide range of preference parameter values but the model is unable to generate both first and second moments of returns that are statistically close to those in the sample.

1. Introduction

A primary goal of financial economists is to understand the dynamics of asset price movements. Recent research has focused on measuring and explaining both the degree of serial correlation and the size and variation of asset returns. In an earlier paper, Cecchetti, Lam, and Mark (1990), we studied the first of

Correspondence to: Stephen G. Cecchetti, Department of Economics, Ohio State University, Columbus, OH 43210, USA.

*This is a revised version of NBER working paper no. 3752, June 1991. We thank Andrew Abel, George Constantinides, Lawrence Christiano, Paul Evans, James Hamilton, G.S. Maddala, N. Gregory Mankiw, Tony Smith, Robert Stambaugh, Alan Viard, Stanley Zin, seminar participants at the 1989 NBER Summer Institute, Claremont, the Federal Reserve Bank of Minneapolis, University of Illinois, Indiana University, Queens, Purdue, Vanderbilt, University of Virginia, University of Washington, and the editors and referee for their comments and suggestions. Cecchetti acknowledges the financial support of the National Science Foundation.

these issues.[1] Here, we examine the second. We show that a representative agent model based on Lucas (1978) calibrated to historical consumption and dividend growth jointly can explain the first but not the second moments of the equity premium and the risk-free rate found in the data.

The empirical issues that concern the work here were first discussed by Mehra and Prescott (1985). They show that for plausible values of the discount factor and the coefficient of relative risk aversion, a simple representative agent model that is calibrated to certain features of historical consumption data implies values of the equity premium that are 'too low' together with values of the risk-free that are 'too high'. The difficulty that Mehra and Prescott found in using a frictionless, pure exchange Arrow–Debreu economy to match the first moments of the equity premium and the risk-free rate is what has come to be known as the 'equity premium puzzle'.

This paper has two features that distinguish it from previous studies of equity returns.[2] First, we explicitly separate consumption from dividends. It is common in the literature to set consumption and dividends equal, and then calibrate the model to estimates of a univariate consumption process. But this practice ignores the fact that equities are actually levered claims to firms' production. Recently, both Kandel and Stambaugh (1990a, 1990b, 1991) and Benninga and Protopapadakis (1990) report success in matching the first and second moments of returns data using models with leverage. Those papers treat the leverage ratio – the ratio of debt to the market value of the firm – as a free parameter.[3] These authors are implicitly allowing the share of dividends to consumption to vary in order to match the moments of returns. But the data provide a precise guide as to what that share should be. The payments to equity holders represent only a very small fraction of total consumption – during the twentieth century total dividends have averaged between 3% and 5% of aggregate consumption. When this ratio is imposed, the model cannot fully explain the data.

The second salient feature of this paper is that we develop a testing framework to measure the ability of the model to match the data. This addresses a common

[1]That paper shows how the negative serial correlation in asset returns found in sample is consistent with the equilibrium model described in section 2 below.

[2]Numerous solutions have been proposed to this puzzle. A partial list includes Mankiw's (1986) suggestion that the high risk premium is the consequence of nondiversifiable risk, Reitz's (1988) examination of big crashes, Nason's (1988) study of the consequences of assuming that consumption has some lower bound, Abel's (1988) work on heterogeneous beliefs, Weil's (1989) and Epstein and Zin's (1990) use of nonexpected utility, Constantinides' (1988) and Abel's (1990) models based on habit formation, Labadie's (1989) monetary model, and Mankiw and Zeldes' (1991) separation of stockholders from nonstockholders.

[3]In their original paper, Mehra and Prescott (1985) do note that equity is the residual claim to output after labor has been paid, but in contrast to Benninga and Protopapadakis (1990), they find that it does not help in generating a large equity premium. Brainard and Summers (1990), who also examine the impact of bonds, claim that the Mehra and Prescott analysis is misleading.

problem in the calibration literature. Previous authors generally fail to provide a well-articulated criterion for evaluating the models they examine. To understand the problem, let ψ_T be a vector of sample moments and $\mu(\theta; \phi)$ be the corresponding implied moments from a completely specified economic model with parameter vectors θ, representing technology, and ϕ, representing tastes. Inspired first by Kydland and Prescott (1982) and then Mehra and Prescott (1985), recent asset pricing and business cycle research has explored various parameterizations in an attempt to set $[\psi_T - \mu(\theta; \phi)] = 0$. In this calibration method, the parameters of the technology, θ, are estimated in order to conform to certain features of the actual economic environment. The investigator then searches over 'plausible' values of the preference parameters, ϕ, in an attempt to find implied moments of the economic model that are 'close' to the sample moments. But this ignores two sources of uncertainty. Since ψ_T is an estimator for the moments of interest and θ is set equal to θ_T, an estimator of the parameters of the technology, the comparison can be thought of as testing to see if the difference between two jointly distributed random variables is zero.

By explicitly accounting for uncertainty that arises from the fact that ψ_T and θ_T are estimated, we can calculate the distribution of $[\psi_T - \mu(\theta_T, \phi)]$, conditional on a particular choice of ϕ, the taste parameters. This allows us to formulate a test statistic and apply standard inference procedures to evaluate the fit of the model.[4]

Our starting point is an equilibrium asset pricing model based on Lucas (1978), generalized to incorporate nontraded assets. We make assumptions about preferences and the stochastic process governing endowments that yield a closed form solution for asset prices. The utility function is time-separable and in the constant relative risk aversion class, while the endowment obeys a form of Hamilton's (1989) Markov-switching model.

We go on to assume that *dividends* represent the flow that accrues to the owner of the equity, and that these are discounted by the intertemporal marginal rate of substitution defined over *consumption*. This approach mirrors reality exactly in that equity prices are based solely on the flow that accrues to their owner.

This bivariate model requires that we estimate a stochastic process for consumption and dividend growth rates jointly. We then calibrate the model by setting the parameters of the endowment process equal to estimates of the Markov-switching model using annual observations on U.S. real consumption

[4]In a recent paper, Hansen and Jagannathan (1990) suggest an alternative method for evaluating whether an asset pricing model is capable of matching the unconditional moments of the data. They examine the ability of various preference specifications to generate intertemporal marginal rates of substitution that match those implied by asset returns data. They find that time-separable utility functions require substantial curvature to meet their criterion. See section 5 below for a more detailed comparison of our results with those of Hansen and Jagannathan.

and dividend growth from 1892 to 1987. We proceed to examine the ability of the bivariate consumption–dividends model to explain the equity premium puzzle, i.e., the first moments of the equity premium and the risk-free rate. As first suggested by Constantinides (1990), we also study the ability of the models to match the covariance matrix of the returns data. While leverage allows us to match the mean equity premium and risk-free rate fairly easily, we find that our attempt to match the second moments fails.

The remainder of this paper is organized into five parts. Section 2 describes the asset pricing model and derives the closed form solution for asset prices for the bivariate consumption–dividends model, assuming that endowment growth follows the Markov–switching process. In section 3 we report estimates of the parameters of the processes constructed using Hansen's (1982) Generalized Method of Moments (GMM) procedure. Section 4 discusses the methodology for evaluating the performance of the models. Section 5 examines the ability of the bivariate consumption–dividends model to match the first moments of the equity premium and the risk-free rate alone, as well as the first and second moments of the returns data together. Section 6 contains concluding remarks.

2. The model

This section presents the model of asset pricing we use and derives the solution for returns. We consider a variant of the Lucas model in which a single nonstoreable consumption good is made available through an exogenous endowment process.[5] Throughout, we assume that the endowment can be described by the Markov-switching model first introduced by Hamilton (1989).

We begin by assuming that the consumption good is generated by two distinct processes. Call the first process *dividends*, and let the claim to dividends be called *equity*. The price of equity is determined in a competitive market. The claim to the second process, which can be thought of as labor income, is not traded. Total consumption in any period is the sum of dividends and labor income. The economy is populated by a large number of identical individuals who are aggregated into a representative agent. This model is presented in section 2.1. Sections 2.2 and 2.3 describe the stochastic model for the endowment and an explicit solution for returns.

[5]The model presented here generalizes our earlier results presented in Cecchetti, Lam, and Mark (1990). We follow the standard practice of the aggregate asset pricing literature and study an endowment economy. Since we are unconcerned with consumption decisions themselves, this makes little difference. In principle, we could specify a production economy and derive the stochastic process for technology that would be required to yield the consumption process we assume. All that is important for our work is that the consumption process fits the data.

2.1. Bivariate consumption–dividends: The investors

We begin with the first-order conditions for the generalized Lucas economy in which consumption and dividends are not necessarily equal. These are

$$P_t^e = \beta \mathrm{E}_t \left\{ \frac{U'(C_{t+1})}{U'(C_t)} [P_{t+1}^e + D_{t+1}] \right\}, \tag{1}$$

$$P_t^f = \beta \mathrm{E}_t \frac{U'(C_{t+1})}{U'(C_t)}, \tag{2}$$

where

P_t^e = real price of the traded asset, or equity
P_t^f = real price of the risk-free asset,
C_t = per capita real consumption,
D_t = dividend from owning one unit of equity,
U' = marginal utility of the representative agent,
β = subjective discount factor, $0 < \beta$, and
E_t = mathematical expectation conditional on information at time t.

Let preferences be given by

$$U(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}, \tag{3}$$

where $0 < \gamma < \infty$ is the coefficient of relative risk aversion.

Now substitute (3) into (1) and (2) to obtain

$$P_t^e C_t^{-\gamma} = \beta \mathrm{E}_t C_{t+1}^{-\gamma} (P_{t+1}^e + D_{t+1}), \tag{4}$$

$$P_t^f = \mathrm{E}_t \beta \left(\frac{C_{t+1}}{C_t} \right)^{-\gamma}. \tag{5}$$

We note here that in the empirical computations below β is allowed to exceed unity. Kocherlakota (1990a) has shown that a unique solution to the asset pricing problem exists in economies where the discount factor is greater than one.[6]

[6]There are a number of ways to understand values of β that are greater than one. For example, it can be thought of as a simple, but crude, way of approximating habit formation behavior of the type described in Constantinides (1990) and Abel (1990). In their models, marginal utility is an increasing function of the level of past consumption. This implies behavior similar to that implied by a discount factor greater than one.

2.2. The endowment process

We assume that consumption and dividends are governed by a bivariate version of Hamilton's (1989) Markov-switching model.[7] Our earlier paper [Cecchetti, Lam, and Mark (1990)] demonstrates the empirical usefulness of the Markov-switching process for modeling consumption and dividend growth. In particular, we showed that this model is able to characterize the significant negative skewness and excess kurtosis found in the consumption and dividend growth data. Furthermore, the Markov-switching model admits a closed form solution to the asset pricing problem.

Let $c_t \equiv \ln C_t$ and $d_t \equiv \ln D_t$. We assume that $\{c_t, d_t\}$ is governed by the following bivariate random walk with two-state Markov drift:

$$\begin{pmatrix} c_t \\ d_t \end{pmatrix} = \begin{pmatrix} c_{t-1} \\ d_{t-1} \end{pmatrix} + \begin{pmatrix} \alpha_0^c \\ \alpha_0^d \end{pmatrix} + \begin{pmatrix} \alpha_1^c \\ \alpha_1^d \end{pmatrix} S_t + \begin{pmatrix} \varepsilon_t^c \\ \varepsilon_t^d \end{pmatrix}, \tag{6}$$

where $\begin{pmatrix} \varepsilon_t^c \\ \varepsilon_t^d \end{pmatrix}$ is i.i.d. normal with mean zero and covariance matrix $\Sigma = \begin{pmatrix} \sigma_c^2 & \sigma_{cd} \\ \sigma_{cd} & \sigma_d^2 \end{pmatrix}$, and S_t is a Markov random variable that takes on values of 0 or 1 with transition probabilities

$$\begin{aligned} \Pr[S_t = 1 | S_{t-1} = 1] &= p, \\ \Pr[S_t = 0 | S_{t-1} = 1] &= 1 - p, \\ \Pr[S_t = 1 | S_{t-1} = 0] &= 1 - q, \\ \Pr[S_t = 0 | S_{t-1} = 0] &= q. \end{aligned} \tag{7}$$

The model of the endowment process requires estimation of nine parameters: $(\alpha_0^c, \alpha_0^d, \alpha_1^c, \alpha_1^d, p, q, \sigma_c^2, \sigma_d^2, \sigma_{cd})$.[8]

As a normalization, we restrict the α_1's to be negative. Consequently the economy will be in a good state when $S_t = 0$ and in a bad state when $S_t = 1$. The

[7]We model consumption, rather than labor income, jointly with dividends to maintain tractability of the model. We note that this formulation places implicit restrictions on the production technology that we do not investigate. See footnote 2 above.

[8]The fact that both consumption and dividend growth have normally distributed innovations implies that the difference between their levels, $(C_t - D_t)$, can take on a negative value with nonzero probability. But, because consumption is nearly 20 times the size of dividends in the data, the probability that the model will ever imply a negative value for labor income is vanishingly small, even though σ_d^2 is nearly 10 times σ_c^2. (See table 1 for details.) As a result, we ignore this potential complication in the presentation below.

parameter q is the probability of remaining in the good state next period given that the economy is currently in the good state, while p is the probability of remaining in the bad state given that the economy is currently in the bad state. The transition probabilities between the two states are $(1-q)$ and $(1-p)$.

The Markov components of the dividend and consumption processes are assumed to be perfectly correlated, and so dividends and consumption are in the good or the bad state simultaneously. The mean consumption and dividend growth rates are α_0^c and α_0^d in the good state and $(\alpha_0^c + \alpha_1^c)$ and $(\alpha_0^d + \alpha_1^d)$ in the bad state.

The bivariate Markov-switching model generalizes the Markov-growth process of Mehra and Prescott in three ways. First, consumption and dividends are modeled jointly. Second, the continuous random variable ε is included. Third, the transition matrix in (7) is permitted to be asymmetrical. We obtain the Mehra and Prescott endowment process by setting $C = D$, $p = q$, and $\Sigma = \mathbf{0}$.

2.3. The solution for returns

Assuming that consumption and dividends follow the bivariate process given by (6) and (7), we obtain the closed form solution for the price of a share of equity and the price of the risk-free asset by the method of undetermined coefficients.[9] Conjecture the following solution:

$$P_t^e = \rho(S_t) D_t. \tag{8}$$

The problem is to verify that (8) solves (4) and to find the function $\rho(S_t)$. To do this, first substitute (8) into (4) to obtain

$$\rho(S_t) D_t C_t^{-\gamma} = \beta \, \mathrm{E}_t C_{t+1}^{-\gamma} D_{t+1} [\rho(S_{t+1}) + 1]. \tag{9}$$

Next, write (6) in levels,

$$\begin{bmatrix} C_{t+1} \\ D_{t+1} \end{bmatrix} = \begin{bmatrix} C_t \mathrm{e}^{(\alpha_0^c + \alpha_1^c S_{t+1} + \varepsilon_{t+1}^c)} \\ D_t \mathrm{e}^{(\alpha_0^d + \alpha_1^d S_{t+1} + \varepsilon_{t+1}^d)} \end{bmatrix}. \tag{10}$$

Now substitute (10) into (9) and note that ε_t^c and ε_t^d are i.i.d. normal with covariance matrix Σ to obtain

$$\rho(S_t) = \beta \, \mathrm{e}^{[\alpha_0^d - \gamma\alpha_0^c + \frac{1}{2}(\gamma^2\sigma_c^2 - 2\gamma\sigma_{cd} + \sigma_d^2)]} \, \mathrm{E}_t \mathrm{e}^{(\alpha_1^d - \gamma\alpha_1^c) S_{t+1}} [\rho(S_{t+1}) + 1]. \tag{11}$$

[9]The technique presented here can also be used to obtain a solution for the general case in which the endowment follows an n-state Markov-switching process in the mean, the variance, or both.

Because S_t can take on only two values, 0 or 1, (11) is a system of two linear equations in $\rho(0)$ and $\rho(1)$. Solving these two equations yields

$$\rho(0) = \frac{1 - (p + q - 1)\tilde{\alpha}_1 \tilde{\beta}}{1 - \tilde{\beta}(p\tilde{\alpha}_1 + q) + \tilde{\beta}^2 \tilde{\alpha}_1 (p + q - 1)} - 1, \tag{12}$$

$$\rho(1) = \frac{1 - (p + q - 1)\tilde{\beta}}{1 - \tilde{\beta}(p\tilde{\alpha}_1 + q) + \tilde{\beta}^2 \tilde{\alpha}_1 (p + q - 1)} - 1, \tag{13}$$

where

$$\tilde{\beta} = \beta \mathrm{e}^{[\alpha_0^d - \gamma\alpha_0^c + \frac{1}{2}(\gamma^2 \alpha_c^2 - 2\gamma\sigma_{cd} + \sigma_d^2)]},$$

$$\tilde{\alpha}_1 = \mathrm{e}^{\alpha_1^d - \gamma\alpha_1^c}.$$

This establishes that (8) is the solution to (4).[10]

The price of the risk-free asset is obtained using (6) and (7) to evaluate (5). That is,

$$P_t^{\mathrm{f}} = \beta \mathrm{e}^{(-\gamma\alpha_0^c + \frac{1}{2}\gamma^2\sigma_c^2)} \varphi(S_t), \tag{14}$$

where $\varphi(0) = q + (1 - q)\mathrm{e}^{-\gamma\alpha_1^c}$ and $\varphi(1) = p\mathrm{e}^{-\gamma\alpha_1^c} + (1 - p)$. The implied rates of return to holding the equity and risk-free assets from date t to $t + 1$ are

$$\begin{aligned} R_t^{\mathrm{e}} = R^{\mathrm{e}}(S_{t+1}, S_t, \varepsilon_{t+1}^d) &= \frac{P_{t+1}^{\mathrm{e}} + D_{t+1}}{P_t^{\mathrm{e}}} - 1 \\ &= \frac{\rho(S_{t+1}) + 1}{\rho(S_t)} \mathrm{e}^{(\alpha_0^d + \varepsilon_{t+1}^d + \alpha_1^d S_{t+1})} - 1, \end{aligned} \tag{15}$$

$$R_t^{\mathrm{f}} = R^{\mathrm{f}}(S_t) = \frac{1}{P_t^{\mathrm{f}}} - 1. \tag{16}$$

Next, integrate ε_{t+1}^d out of the expression for the equity return to obtain

$$R^{\mathrm{e}}(S_{t+1}, S_t) = \frac{\rho(S_{t+1}) + 1}{\rho(S_t)} \mathrm{e}^{(\alpha_0^d + \frac{1}{2}\sigma_d^2 + \alpha_1^d S_{t+1})} - 1. \tag{17}$$

[10]The solution for the equity price can also be obtained by iterating the stochastic difference equation (8) forward and exploiting results in Hamilton (1989) to evaluate the resulting geometric series. The conditions required for this series to converge also guarantee that $\rho(S_t)$ is nonnegative. The nonnegativity of $\rho(S_t)$ is always imposed in the empirical work that follows.

Finally, the implied means of the risky and the risk-free rates of return are computed by summing over the probabilities:

$$\mu^{e} = \mathrm{E}[R^{e}(S_{t+1}, S_t)]$$

$$= \sum_{s_{t+1}=0}^{1} \sum_{s_t=0}^{1} \Pr(S_{t+1} = s_{t+1} \mid S_t = s_t) \Pr(S_t = s_t) R^{e}(s_{t+1}, s_t), \tag{18}$$

$$\mu^{f} = \mathrm{E}[R^{f}(S_t)] = \sum_{s_t=0}^{1} \Pr(S_t = s_t) R^{f}(s_t), \tag{19}$$

where $\Pr(S_t = s_t)$ is the unconditional probability that $S_t = s_t$. For the good state, $s_t = 0$, this is $\Pr(S_t = 0) = (1-p)/(2-p-q)$. The unconditional probability of a bad state is $\Pr(S_t = 1) = 1 - \Pr(S_t = 0) = (1-q)/(2-p-q)$.

The expected equity premium follows as

$$\mu^{p} = \mu^{e} - \mu^{f}. \tag{20}$$

In order to compute the implied covariance of the premium and the risk-free rate, first use (15) and (16) to define

$$R_t^{p} = R^{p}(S_{t+1}, S_t, \varepsilon_{t+1}^{d}) = R^{e}(S_{t+1}, S_t, \varepsilon_{t+1}^{d}) - R^{f}(S_t). \tag{21}$$

Next, integrate ε_{t+1}^{d} out of (21) to obtain the equivalent of (17),

$$R^{p} = R^{p}(S_{t+1}, S_t).$$

Using these, we can write the covariance matrix of the equity premium and the risk-free rate:

$$\operatorname{var}\begin{pmatrix} R_t^{p} \\ R_t^{f} \end{pmatrix} = \sum_{s_{t+1}=0}^{1} \sum_{s_t=0}^{1} \Pr(S_{t+1} = s_{t+1} \mid S_t = s_t) \Pr(S_t = s_t),$$

$$\times \begin{bmatrix} \int (R_t^{p} - \mu^{p})^2 \, \mathrm{d}F(\varepsilon_{t+1}^{d}) & (R^{p} - \mu^{p})(R^{f} - \mu^{f}) \\ (R^{p} - \mu^{p})(R^{f} - \mu^{f}) & (R^{f} - \mu^{f})^2 \end{bmatrix}, \tag{22}$$

where $F(\varepsilon_{t+1}^{d})$ is the cumulative normal distribution function, with mean zero and variance σ_d^2.

3. Estimation of the endowment process

The next step in deducing the behavior of the joint equity premium–risk-free rate process (R_t^{p}, R_t^{f}) implied by the model of section 2 is to estimate the parameters of the stochastic process for the endowment.

Estimates are computed by Hansen's (1982) Generalized Method of Moments (GMM), using annual data on real dividend growth for the Standard and Poor's index, and real per capita consumption growth, together with the returns reported in table 2 of section 5 below.[11] The moments used in the GMM procedure were chosen to match the maximum likelihood estimates of the bivariate consumption–dividends Markov-switching model as closely as possible.[12] The appendix describes the moment conditions used for the estimation and the criterion for their choice.

All the data sources are described in detail in the data appendix to Cecchetti, Lam, and Mark (1990). Briefly, the dividends data are the nominal figures from Campbell and Shiller (1987), deflated by the annual average CPI from Wilson and Jones (1988).[13] The real consumption series begins with the Kendrick data in 1889, reported in Balke and Gordon (1986), and continues in 1929 with the NIPA series for real personal consumption expenditure.[14] The consumption data are divided by population estimates from the *Historical Statistics of the United States* and the *Economic Report of the President* to obtain per capita observations.[15]

The results are reported in table 1. In addition to the estimates of the Markov-switching model reported in the first column of the table, we include estimates of two nested alternatives: a random walk, where $q = 1$, $p = 0$ and $\alpha_1^c = \alpha_1^d = 0$; and a Mehra–Prescott style model, estimated by matching the means and variances of consumption and dividend growth, and the first-order autocorrelation of consumption growth, where $\sigma_c^2 = \sigma_d^2 = \sigma_{cd} = 0$ and $p = q$.[16] Standard errors for these Mehra–Prescott style estimates are obtained by GMM on the exactly identified model.

In our earlier paper we provide a number of tests that demonstrate the statistical superiority of the asymmetric Markov-switching model over alterna-

[11]By using data on the S&P index we are assuming that the growth in S&P dividends, as well as the total return to holding the index, accurately mirrors the behavior of economy-wide dividend growth and equity returns.

[12]Ideally we would prefer to estimate all of the parameters of the endowment and the moments of returns jointly using a maximum likelihood procedure. But this would require that we evaluate the likelihood for returns, and force the economic model to match all the aspects of the returns data. There is no point to pursuing this strategy, as it is unlikely that our simple model can explain all aspects of the data.

[13]The method and deflator were chosen so that our results in table 2 match the values in Mehra and Prescott's table 1 as closely as possible.

[14]None of the results reported in sections 3, 4, and 5 are changed in any noticeable way if the total consumption series is replaced by consumption of nondurables and services alone as used by Grossman and Shiller (1981).

[15]Since the Mehra–Prescott process requires matching a first-order autocorrelation of consumption growth, we estimate all the models over the period 1892 to 1987.

[16]The estimation of the Mehra–Prescott model uses the procedure suggested on page 154 of their paper.

Table 1

GMM estimates of the bivariate Markov-switching model, 1892–1987 (asymptotic t-ratios in parentheses).

	Bivariate consumption–dividends		
	Markov-switching	Random walk	Mehra–Prescott
q	0.9604 (20.49)	1.0	0.4723 (5.62)
p	0.0994 (0.60)	–	0.4723 (5.62)
α_0^c	0.0218 (6.06)	0.0183 (4.93)	0.0206 (7.53)
α_1^c	– 0.0832 (3.19)	–	– 0.0908 (5.70)
α_0^d	0.0255 (1.69)	0.0118 (0.93)	0.1334 (7.13)
α_1^d	– 0.3208 (4.99)	–	– 0.2431 (7.29)
σ_c	– 0.0333 (8.67)	0.0372 (9.08)	0.0
σ_d	0.1030 (5.00)	0.1216 (7.29)	0.0
σ_{cd}	0.2732 (1.53)	0.443 (4.49)	0.0
$\Pr(S = 1)$	0.0425	0.0	0.50

tives of the type considered here. Briefly, we first estimate the Markov-switching model and the random walk model by maximum likelihood and construct a likelihood ratio test. The result allows us to reject the random walk model in favor of the Markov-switching model at the 1% level or better.[17] Furthermore, it is clear from the size of the estimates of the components of Σ, as well as the fact that p is significantly smaller than q, that the Markov-switching model dominates the Mehra–Prescott style estimates. Finally, we prefer the Markov-switching model because of its ability to produce the negative serial correlation in equity returns discussed at length in our earlier paper.

[17]As noted in Cecchetti, Lam, and Mark (1990), the Markov-switching model is not identified under the null hypothesis that the random walk is correct. Nevertheless, Monte Carlo experiments demonstrated that the critical values for the pseudo-likelihood ratio test were close to those of a chi-squared distribution with degrees of freedom slightly larger than the maximum number of constraints one could possibly count. For the bivariate consumption–dividends model considered here, twice the difference in the log-likelihood is 20.58. This is beyond the 99.5 percent critical value for a χ_6^2, a very conservative standard considering that the test has at most 4 constraints.

For the purposes of computing the equity premium, an important thing to notice is the size of the drop in consumption and dividends in the bad state, the $(\alpha_0 + \alpha_1)$'s. During a boom, the growth rates of dividends and consumption are estimated to be 2.6 percent and 2.2 percent, respectively. But in a downturn these fall to -29.5 percent for dividends and -6.2 percent for consumption. In the bad state, dividends crash. As one would expect, consumption is quite a bit smoother.

To help understand the implications of these estimates, notice that the unconditional probability of a crash is $\Pr(S_t = 1) = (1 - q)/(2 - p - q) = 0.043$, and so we expect real dividends to fall by 30 percent in approximately 4 of the 96 years of the sample. While this may seem surprising, it is consistent with the historical experience. The model estimates imply that, given that the economy is in the bad state ($S_t = 1$), the asymptotic 95 percent confidence interval for the growth in real dividends is $(-0.50, -0.09)$. The same confidence interval for the good state ($S_t = 0$) is $(0.23, -0.18)$. Consequently, if dividends fall by 20 percent or more, we can be fairly certain that $S_t = 1$. Of the 96 years in our sample, 6 meet this criterion – real dividends fell by more than 20 percent.[18]

We note that the GMM criterion function for the bivariate Markov-switching model is flat for variations in p, the probability of remaining in a crash state given that the economy is currently in that state. This is not surprising considering the asymmetric behavior of dividends over the business cycle. Downturns tend to be short-lived, lasting 4 to 6 quarters. This makes it difficult to obtain a good estimate of p using annual observations.

4. A generalized calibration methodology

In this section we generalize the standard calibration methodology to incorporate statistical inference. Calibration, as it is usually practiced, seeks a parameterization of the model for which the implied moments exactly match the sample moments. Two sources of uncertainty are ignored by the standard procedure, however.[19]

The first source involves the uncertainty in the sample moment vector, ψ_T, as an estimator of the population moment vector, ψ, since

$$\sqrt{T}(\psi_T - \psi) \overset{A}{\sim} \mathrm{N}(0, \Omega_\psi). \tag{23}$$

[18]Real aggregate dividends fell by more than 30 percent during 3 years, from 20 to 30 percent in 3 years, and by 10 to 20 percent in 8 years.

[19]Gregory and Smith (1988) address some of these issues using a Monte Carlo methodology.

The second source of uncertainty arises from the dependence of the implied moments on parameters that are estimated. Consequently, the vector of moments implied by the model is itself stochastic. Specifically, let μ be the vector of moments implied by the model and let θ_T be the GMM estimator of the parameter vector. In the bivariate consumption–dividends model of section 2, $\theta = (\alpha_0^c, \alpha_0^d, \alpha_1^c, \alpha_1^d, p, q, \sigma_c^2, \sigma_d^2, \sigma_{cd})$. Thus,

$$\sqrt{T}(\theta_T - \theta) \overset{\text{A}}{\sim} \text{N}(0, \Omega_\theta). \tag{24}$$

Taking a first-order Taylor series expansion of μ about θ and using (24), we obtain

$$\sqrt{T}[\mu(\theta_T; \beta, \gamma) - \mu(\theta; \beta, \gamma)] \overset{\text{A}}{\sim} \text{N}[0, \Omega_\mu(\beta, \gamma)], \tag{25}$$

where $\Omega_\mu(\beta, \gamma) = (\partial\mu/\partial\theta')\Omega_\theta(\partial\mu/\partial\theta)$. From (23) and (25), it follows that

$$\sqrt{T}\{[\psi_T - \psi] - [\mu(\theta_T; \beta, \gamma) - \mu(\theta; \beta, \gamma)]\} \overset{\text{A}}{\sim} \text{N}[0, \Omega(\beta, \gamma)], \tag{26}$$

where

$$\Omega(\beta, \gamma) = \Omega_\psi + \Omega_\mu(\beta, \gamma) - T\left\{\frac{\partial\mu}{\partial\theta'}\text{E}[(\theta_T - \theta)(\psi_T - \psi)'] + \text{E}[(\psi_T - \psi)(\theta_T - \theta)']\frac{\partial\mu}{\partial\theta}\right\}.$$

We can now test the hypothesis that the implied moments match the population moments, i.e., H_0: $\psi = \mu(\theta; \beta, \gamma)$. Consider the statistic

$$\mathscr{H} = T[\psi_T - \mu(\theta_T; \beta, \gamma)]'\Omega^{-1}(\beta, \gamma)[\psi_T - \mu(\theta_T; \beta, \gamma)]. \tag{27}$$

Note that ψ_T consistently estimates the population moments ψ both under the null hypothesis, H_0: $\psi = \mu(\theta; \beta, \gamma)$, as well as under the alternative, H_a: $\psi \neq \mu(\theta; \beta, \gamma)$. On the other hand, the implied moment vector $\mu(\theta_T; \beta, \gamma)$ consistently estimates the population moments only when H_0 is true. Thus under the null, $\mathscr{H} \sim \chi^2_{(k)}$, where k is the dimensionality of ψ. Accordingly, we can use the test static, $\mathscr{H}$, to test the model for a given (β, γ) pair.

The following sections use $\mathscr{H}$ to examine two cases of interest. First we study the standard equity premium puzzle and ask whether the model is capable of matching the means, μ^p and μ^f. We then proceed to examine the ability of the model to match both the first and the second moments of the equity premium and the risk-free rate.

We adopt a two-step testing framework to obtain an interval estimate for (β, γ) based on matching several carefully selected moments of the returns data. Our approach contrasts with that of Lee and Ingram (1991) and Burnside (1990), who estimate β and γ using unconditional moment restrictions of the data. Both of these authors propose estimation methods that can be thought of as minimizing a quantity like $\mathscr{H}$ in order to obtain estimates of β and γ. While we report the (β, γ) pair that minimizes $\mathscr{H}$ in section 5 below, we note a problem that would arise if this estimate were to be computed using standard minimization techniques and applying asymptotic theory. In particular, the asymptotic normality of the estimator from such a procedure puts positive probability on values of (β, γ) that violate the transversality condition of the model in section 2.[20] Alternatively, we could simply add the moments of returns that we want the model to match to the list used in the GMM estimation of the technological parameters reported in table 1 above and estimate β and γ jointly with θ. In addition to the technical problem resulting from the transversality condition, such an estimation procedure has a conceptual difficulty that arises from our interpretation of the exercise. We see our goal as asking whether, given the technology, there exist taste parameters capable of matching the returns data. This dictates that we proceed in two steps, first estimating the parameters of the endowment process, and then computing a confidence bound for the taste parameters – β and γ.

5. Matching the moments

5.1. First moments: The equity premium puzzle

Let ψ_T be the two-dimensional vector of the sample means of the equity premium and the risk-free rate. Label these sample moments ψ_T^{p} and ψ_T^{f}, so $\psi_T = (\psi_T^{\mathrm{p}} \psi_T^{\mathrm{f}})'$. This section examines the extent to which the model matches ψ_T. Previous work has shown that the Lucas model with iso-elastic utility, calibrated to the U.S. economy, is not capable of matching the average equity return and risk-free rate simultaneously. Generally, the model implies an equity return that appears 'too small' and a risk-free rate that is 'too large'. But the sense in which the model has failed is vague because the metric employed by the standard procedure to evaluate the model has not been well defined.

[20]In principle it is possible to reformulate the test statistic $\mathscr{H}$ to account for the transversality condition, which we can write as a set of inequality constraints on $\rho(0)$ and $\rho(1)$ in eqs. (12) and (13). Unfortunately, the statistical theory for such a test has only been developed for a specific set of circumstances, not including the one examined here. See, for example Wolak (1991) and Gourieroux, Holly, and Monfort (1982).

Table 2

First and second moments of asset prices: the data, 1892–1987 (robust standard errors in parentheses).[a]

	Mean		Std. dev.
Risk-free rate $[\psi_T^{\mathrm{f}}, \sigma^{\mathrm{f}}(T)]$	1.19 (0.81)		5.27 (0.74)
Equity premium $[\psi_T^{\mathrm{p}}, \sigma^{\mathrm{p}}(T)]$	6.63 (1.78)		19.02 (1.73)
Correlation $[\rho_{\mathrm{fp}}(T)]$		− 0.04 (0.12)	

[a]Estimates are computed using Hansen's (1982) Generalized Method of the Moments, simultaneously with the bivariate Markov-switching model for consumption and dividends reported in the first column of table 1. See the appendix for details.

The data this section attempts to match are presented in the first column of table 2.[21] Over the historical period 1892–1987, the premium on equities and relatively risk-free short-term debt has averaged 6.63 percent and 1.19 percent, respectively.[22] The uncertainty in this sample mean vector is highlighted in fig. 1, where the 95 percent confidence ellipse about ψ_T is plotted.[23] If we ignore parameter uncertainty, any vector of means $[\mu^{\mathrm{p}}(\theta; \beta, \gamma), \mu^{\mathrm{f}}(\theta; \beta, \gamma)]$ implied by the model of section 2 that falls within the ellipse will not be rejected by the data at the 5 percent level. The '×' in fig. 1 marks the sample value ψ_T. The 'Δ' marks a representative model value, $[\mu^{\mathrm{p}}(\theta_T; \beta = 0.98, \gamma = 23.0), \mu^{\mathrm{f}}(\theta_T; \beta = 0.98, \gamma = 23.0)]$.

Fig. 2 summarizes tests of the model at the 5 percent level, taking into account the uncertainty in both the sample moments and the implied moments. This figure displays the contour obtained by searching over admissible (β, γ) pairs that yield values of the $\mathscr{H}$-statistic that are less than or equal to the 5 percent critical value of the $\chi^2_{(2)}$ distribution. For a given value of γ above 11, the upper limit on β is the boundary of the admissible parameter space. That is, larger values of β result in explosive behavior of the stochastic difference equation (4).

[21]Again, see the appendix to Cecchetti, Lam, and Mark (1990) for details on the sources of the data. The real risk-free rate is computed using one-year U.S. Treasury note yields, or the equivalent, and realized CPI inflation. The equity return is computed from data supplied by Campbell and Shiller (1987).

[22]We note that over the 1871–1987 sample period, the average equity premium is only 5.92 percent, while the mean risk-free rate is 2.11 percent. Furthermore, Seigel (1992) shows that adding more nineteenth century data continues to raise the risk-free rate and lower the equity premium. Using data from 1800 to 1990, he reports an average equity premium of only 4.64 percent and an average risk-free rate of 3.13 percent. While the model would be able to match these moments more easily, the GMM procedure requires that we have data for returns, dividends, and consumption over the same sample period. The lack of consumption data prior to 1889 precludes our use of the earlier returns data.

[23]Details concerning the estimation of the relevant covariance matrices are deferred to the appendix.

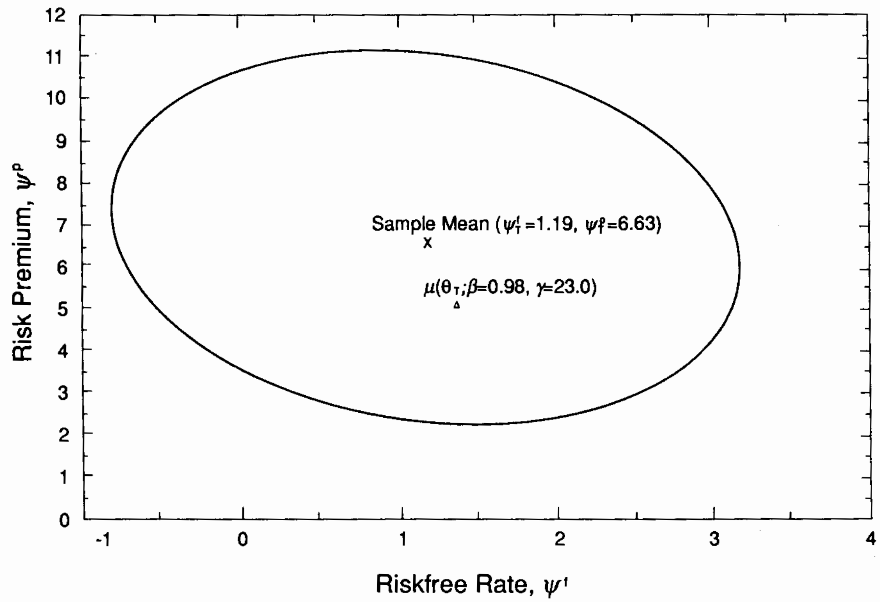

Fig. 1. 95 percent confidence ellipse for the mean equity premium and risk-free rate.

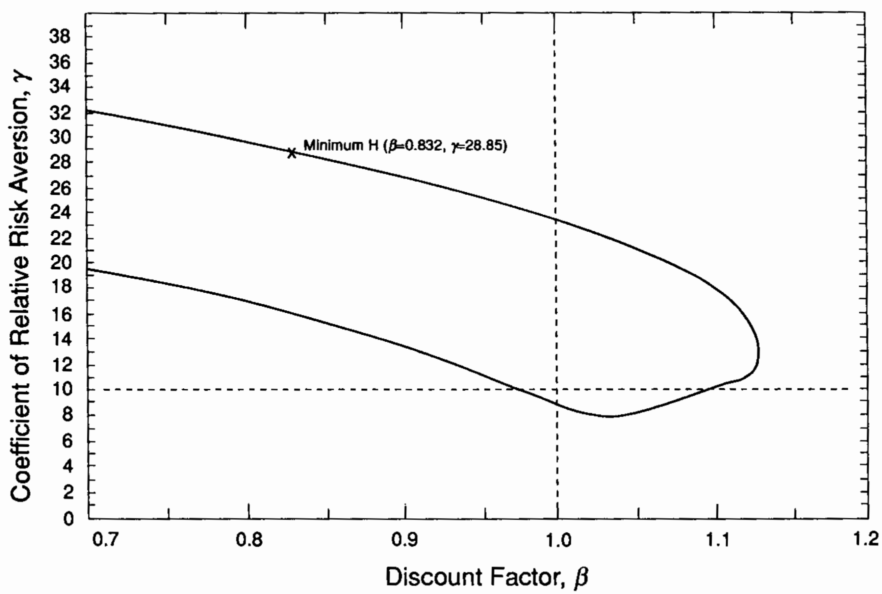

Fig. 2. Matching the means in the Markov switching model (β, γ) pairs not rejected at the 5 percent level.

The figure is partitioned into four quadrants. The lower left quadrant in which $\beta < 1.0$ and $\gamma < 10.0$ is the subset of the parameter space searched by Mehra and Prescott. As can be seen, the upper corner of this region is not rejected at the 5 percent level. In fact, we find that values such as ($\beta = 0.999$, $\gamma = 8.0$) and ($\beta = 0.968$, $\gamma = 9.9$) are within the 95 percent confidence region.

If one is willing to accept values of γ in excess of 10, there is a large class of models that is not rejected at the 5 percent level. There is a clear trade-off in that low values of γ require high values of β to match the data. The reason for this is related to the principal difficulty in matching the first moments, which has been that the implied risk-free rates is 'too high' given the implied equity return. As can be seen from eqs. (14) and (16), larger values of β work to lower the implied risk-free rate, ameliorating this problem.

The figure also plots the (β, γ) pair that minimizes the value of the test statistic $\mathscr{H}$. This minimum value, $\mathscr{H} = 0.01$, occurs at ($\beta = 0.832$, $\gamma = 28.85$) and is very close to the upper boundary of the region.[24] In the region above the upper line plotted in fig. 2 the transversality condition for the asset pricing model is violated – these values of β and γ imply a negative value for $\rho(S_t)$ in (8) and hence for prices. The location of the minimum makes clear that any estimator for β and γ computed by minimizing $\mathscr{H}$ would not be asymptotically normal.[25]

We have completed the task set forth by Mehra and Prescott: to match the sample means of the risk-free return and the equity premium with the Lucas model. We now address the ability of the model to match the covariance matrix of returns in addition to their means.

5.2. *Matching first and second moments*

Let ψ_T be the five-dimensional vector consisting of the first and second moments of returns: the two means $(\psi^{\mathrm{p}}_T \psi^{\mathrm{f}}_T)$, the standard deviations $[\sigma_{\mathrm{p}}(T), \sigma_{\mathrm{f}}(T)]$, and the correlation $\rho_{\mathrm{fp}}(T)$. This section examines the ability of the model to simultaneously match these five moments of the joint equity premium and the risk-free rate process. The sample standard deviations and correlation are presented in table 2.

[24]These results are similar in spirit to those in Kocherlakota (1990b), Hansen and Jagannathan (1991), and Kandel and Stambaugh (1990b). Kocherlakota finds that for $\beta = 1.139$ and $\gamma = 13.7$ a model calibrated to univariate consumption (with consumption set equal to dividends) is capable of generating mean equity premium and risk-free rates that are close to the sample values. Hansen and Jagannathan also find that with CRRA preferences, a high coefficient of relative risk aversion is required to generate intertemporal marginal rates of substitution with sufficient variability to be consistent with returns data. To match the mean of the risk-free rate and the equity premium, together with the variance of the equity premium, Kandel and Stambaugh require $\beta = 0.9973$ and $\gamma = 55$.

[25]In addition to the problem associated with the transversality condition, any attempt to estimate β and γ by minimizing $\mathscr{H}$ will result in very large standard errors. This is a consequence of the fact that for a given value of γ decreasing β results in very small changes in the value of $\mathscr{H}$. For example, at ($\beta = 1.0$, $\gamma = 23.0$), $\mathscr{H}$ is below 0.03.

We report the implied first and second moments from our model for a variety of parameter values in the top panel of table 3. As is apparent from the table, the Markov-switching model is not capable of matching the five first and second moments of the risk-free rate and the equity premium. Restricting β to be greater than 0.7, the minimum value for $\mathscr{H}(5)$ is 23.24 and occurs at ($\beta = 0.70, \gamma = 29.00$).

The failure of the Markov-switching model to match the first and second moments is a consequence of the correlation of the moments with each other. In fact, it is possible to find values of β and γ for which five univariate tests all fail to reject that the model and the data are the same. For example, at ($\beta = 1.04$, $\gamma = 20$) the highest t-ratio is 1.42 (in absolute value) for the standard deviation of the equity premium – the remaining four t-ratios are all below 0.10. Nevertheless, for this parameterization $\mathscr{H}(5)$ is 26.71, which has a p-value of 0.00006. This (β, γ) pair is rejected by our joint test because the deviations of all five moments from the sample values are negative, while several elements in the covariance matrix are large and negative.[26]

The Markov-switching model nests the pure Markov-endowment model of Mehra and Prescott, as well as the geometric random walk. The middle and bottom panels of table 3 display the implied moments from these models. The random walk model is clearly unacceptable since it implies a nonstochastic risk-free rate. But the Mehra–Prescott style model is capable of matching the five moments when β exceeds 1, and γ ranges from roughly 5.5 to 23. The reason for the success of this model, and the failure of the Markov-switching model is evident from the estimates in table 1. The Mehra–Prescott style model assumes that $p = q$, and then matches the first-order sample autocorrelation of consumption growth, which is *negative* in annual data. The result is an estimate of $p = q = 0.47$, and the implication that the economy crashes with probability 0.50. This makes equities very risky, raising the level and variance of the equity

[26]We conjecture that $\mathscr{H}(5)$ has poor small-sample properties, and so the evidence against the model may be overstated. For example, we have examined the distribution of $\mathscr{H}(5)$ using the following Monte Carlo experiment, with 1000 replications: (1) choose values for the taste parameters β and γ; (2) using the Markov-switching model and the parameter values in table 1, draw a sequence of consumption and dividends; (3) using the model, the sequence of consumption and dividends, and the values of β and γ, generate a sequence of returns; (4) compute estimates of the moments and the technology parameters using GMM as described in the appendix; (5) compute $\mathscr{H}(5)$. We find that for ($\beta = 0.999, \gamma = 9.9$) the small-sample p-value for $\mathscr{H}(5) = 42.71$, the value from the data, is 0.0480. We have repeated this experiment for ($\beta = 0.8, \gamma = 5.0$) and find that the small-sample distribution of $\mathscr{H}(5)$ is nearly unchanged. We note that there is a conceptual problem with our experiments as a result of the fact that they include trials with estimates of technology parameters which imply violation of the transversality condition. When this occurs, the test statistic $\mathscr{H}(5)$ is difficult to interpret. For the experiments we have performed, this problem is relatively minor. It occurs in 10.4 percent of the trials for ($\beta = 0.999, \gamma = 9.9$) and in 2.8 percent of the trials when ($\beta = 0.8, \gamma = 5.0$). Unfortunately, the frequency of this problem increases with β and γ, and will arise in roughly 50 percent of the Monte Carlo trials for (β, γ) pairs near the upper boundary of the confidence region in fig. 2. Because of this problem, we think a more thorough investigation of the sampling properties of $\mathscr{H}(5)$ will not be useful.

Table 3

First and second moments of asset returns: model values using bivariate consumption–dividends processes (model implied standard errors are in parentheses).[a]

β	γ	Risk-free rate	Equity premium	Correlation	$\mathscr{H}(2)$	$\mathscr{H}(5)$
		Markov-switching model				
0.999	9.9	11.65 (1.53)	2.42 (13.11)	−0.02	5.02	42.71
1.098	9.9	1.58 (1.39)	2.19 (11.90)	−0.02	6.33	32.32
1.056	15.0	5.02 (2.52)	3.52 (12.11)	−0.03	1.17	29.68
1.039	20.0	1.21 (3.68)	4.12 (11.31)	−0.02	0.11	26.72
0.825	29.0	0.80 (6.20)	1.28 (10.62)	0.02	0.02	24.96
		Random walk model				
0.999	9.9	12.11 (0.00)	2.25 (13.95)	0.00	5.65	—
1.098	9.9	2.00 (0.00)	2.05 (12.69)	0.00	6.56	—
1.056	15.0	6.64 (0.00)	3.26 (13.41)	0.00	2.46	—
1.039	20.0	5.23 (0.00)	4.31 (13.36)	0.00	1.07	—
0.825	29.0	15.16 (0.00)	6.89 (14.89)	0.00	0.43	—
		Mehra–Prescott model				
0.999	9.9	12.32 (2.20)	5.46 (15.51)	0.01	4.06	21.39
1.098	9.9	2.20 (2.00)	5.01 (14.23)	0.01	0.61	3.78
1.056	15.0	7.53 (3.02)	8.19 (16.18)	0.01	0.60	11.50
1.039	20.0	7.74 (3.77)	10.94 (17.33)	0.04	0.59	18.17
0.825	29.0	25.86 (5.53)	16.93 (21.37)	0.06	1.30	46.79

[a] $\mathscr{H}(2)$ tests for two means and $\mathscr{H}(5)$ tests for all five moments. The 5 percent critical value for the $\chi^2_{(2)}$ is 5.99, for the $\chi^2_{(5)}$ it is 11.07. For the random walk model, both the correlation and $\mathscr{H}(5)$ cannot be computed since R^f_t is nonstochastic.

premium. Both the historical record and the estimates of the Markov-switching model suggest that this is not a very good characterization of the consumption and dividends data. The estimates of the Markov-switching model show that p is likely to be substantially smaller than q, and so the crashes should occur much less than one-half of the time. Furthermore, the Mehra–Prescott style estimates imply that dividends should fall by more than 10 percent in roughly half the years of the sample. In fact, dividends fell by 10 percent in only 14 of the 96 years we examine.

Our results are in contrast with those of both Kandel and Stambaugh (1991) and Hansen and Jagannathan (1991). Using the standard calibration procedure, a four-state discrete Markov process for consumption growth, assuming that firms issue both debt and equity, and allowing leverage to be a free parameter, Kandel and Stambaugh are able to match both the means and standard deviations of the equity premium and the risk-free rate. Our failure to match the second moments is largely a consequence of our specification of leverage. With time-separable utility – the case that is closest to ours – Kandel and Stambaugh's success comes with a degree of leverage equal to 0.44. But, given the percentage of dividends in consumption, this is far too low. Also, we note that Kandel and Stambaugh's interest is in matching point estimates of the moments, while we concentrate on inference.

The differences between our results and those of Hansen and Jagannathan (1991) are more subtle. For a given mean intertemporal marginal rate of substitution (*IMRS*), they represent the admissible *IMRS* volatility as a function of the first and second moments of returns data. Admissibility is determined by the unconditional Euler equations, and the result is a bound on the standard deviation of the *IMRS*. Hansen and Jagannathan go on to find that a very high value of the CRRA coefficient γ is required to meet the restrictions implied by the bound. In a separate paper, Cecchetti, Lam, and Mark (1992), we develop and implement statistical tests of these lower bound restrictions. We conclude that the availability of relatively short time series of consumption data undermines the ability of tests that use the restrictions implied by the volatility bound to discriminate among different utility functions. Using the same data as that employed here, we find that there is a broad set of models that are able to meet the volatility bound. For example, with $\beta = 0.99$, values of $\gamma \geq 10$ are not rejected at the 5 percent level using the Hansen and Jagannathan framework. But just because a model meets the volatility bound does not imply it matches the moments of returns data. In fact, using a fully articulated model of asset pricing we find that the unconditional first and second moments used here are a more stringent test of a model's ability to mimic the characteristics of asset price data.

Finally, we note the similarity between our results in table 3 and those reported originally by Grossman and Shiller (1981). They conclude that the high sample variance of equity returns poses quite a challenge for asset pricing models. This accords with our main conclusion that the model of section 2 is rejected because of its inability to match simultaneously the relatively low standard deviation of the risk-free rate and the relatively high standard of the equity premium.[27]

6. Conclusion

This paper addresses two issues central to the literature in calibration and aggregate asset pricing. First, we develop a testing framework for rigorously evaluating the ability of an economic model to match specific sample moments of the historical data. Second, we examine the ability of a model in which dividends explicitly represent the flow that accrues to the owner of the equity, and they are discounted by the intertemporal marginal rate of substitution defined over consumption, to solve the equity premium puzzle.

Using a methodology that combines the features of model calibration and classical statistical inference, we conclude that the original Mehra–Prescott form of the equity premium puzzle, based solely on first moments, does not present a challenge for a simple general equilibrium Lucas asset pricing model.[28] But when the endowment is forced to conform closely to the data, as it is in the Markov-switching model, and leverage is forced to imply that the dividend flow match what we actually observe, then the model cannot match the first and second moments taken together.

Appendix

Here we describe the procedure we use to obtain the GMM estimator of the parameters of the endowment process, θ_T, the sample moments, ψ_T, and the estimate of their asymptotic covariance matrix.

Let $\{x_t\}$ be a vector-valued sequence of observations on stock and bond returns and consumption and dividend growth rates. Let $\lambda' = (\psi', \theta')$ denote the parameter vector whose true value is $\lambda_0' = (\psi_0', \theta_0')$. Finally, let $f(x_t, \lambda)$ be the vector of moment conditions used in the estimation of $\mathrm{E} f(x_t, \lambda_0) = 0$. To construct $f(x_t, \lambda)$, we stack moment conditions used to compute the sample moments of returns with those for estimating the endowment process parameters. That is,

$$f(x_t, \lambda) = \begin{bmatrix} f_1(x_t, \psi) \\ f_2(x_t, \theta) \end{bmatrix}. \tag{A.1}$$

[27]It has been suggested to us that this similarity could be a result of the fact that we both study second moments and asymptotic normality is typically a much better small-sample approximation for first moments than it is for second moments. Flavin (1983) reports evidence that tests of the Grossman and Shiller type are often severely biased toward rejection in small samples. Our Monte Carlo experiments indicate a similar bias. See footnote 26 above.

[28]As we demonstrate in our earlier paper, Cecchetti, Lam, and Mark (1990), this model is capable of matching the serial correlation in equity returns.

When both the first and second moments of returns are required,

$$f(x_t,\psi)=\begin{bmatrix} r_{p,t}-\psi^p \\ r_{f,t}-\psi^f \\ (r_{p,t}-\psi^p)^2-\sigma_p^2 \\ (r_{f,t}-\psi^f)^2-\sigma_f^2 \\ [(r_{p,t}-\psi^p)(r_{f,t}-\psi^f]/(\sigma_p\sigma_f)-\rho_{pf} \end{bmatrix}, \tag{A.2}$$

where $r_{p,t}$ and $r_{f,t}$ are the observations in the data at time t for the risk premium and the risk-free rate. When we require only the first moments of returns, we use only the first two elements of (A.2).

$f_2(x_t,\theta)$ is a vector of the deviations of the observations from their means as implied by either Hamilton's Markov-switching model, the Mehra–Prescott style Markov model, or the geometric random walk model. Let $x_{c,t}$ and $x_{d,t}$ denote the consumption and dividend growth rate at data t. We use the following moments:

Morkov-switching model – 10 moments:

$$\mathrm{E}(x_{k,t}^j), \qquad j=1,2,3, \quad k=c,d,$$

$$\mathrm{E}(x_{d,t}^2 x_{d,t-1}),$$

$$\mathrm{E}(x_{c,t}^j x_{d,t}^j), \qquad j=1,2,$$

$$\mathrm{E}(x_{c,t} x_{d,t-1}).$$

Mehra–Prescott style model – 5 moments:

$$\mathrm{E}(x_{k,t}^j), \qquad j=1,2, \quad k=c,d,$$

$$\mathrm{E}(x_{c,t} x_{c,t-1}).$$

Random walk model – 5 moments:

$$\mathrm{E}(x_{k,t}^j), \qquad j=1,2, \quad k=c,d,$$

$$\mathrm{E}(x_{c,t} x_{d,t}).$$

In order to construct the GMM estimator, let

$$g_T \quad =\frac{1}{T}\sum_{t=1}^{T} f(x_t,\lambda),$$

$$W_{T,0} = \frac{1}{T}\sum_{t=1}^{T} f(x_t, \lambda) f(x_t, \lambda)',$$

$$W_{T,j} = \frac{1}{T}\sum_{t=1}^{T-j} f(x_t, \lambda) f(x_{t+j}, \lambda)',$$

$$w_{j,m} = 1 - \frac{j}{m+1},$$

$$W_T = W_{T,0} + \sum_{j=1}^{m} w_{j,m}[W_{T,j} + W'_{T,j}],$$

$$D_T = \frac{1}{T}\sum_{t=1}^{T} \frac{\partial f(x_t, \lambda)}{\partial \lambda}.$$

The GMM estimator, λ_T, minimizes the quadratic criterion function

$$\Phi(\lambda) = g'_T W_T^{-1} g_T, \tag{A.3}$$

and the asymptotic covariance matrix of λ_T is consistently estimated by

$$\frac{1}{T}[D'_T W_T^{-1} D_T]^{-1}, \tag{A.4}$$

where W_T is the Newey and West (1987) estimator of the spectral density matrix of $f(x_t, \lambda)$ at frequency zero. We set $m = 3$, which conforms to Newey and West's $T^{1/4}$ rule. We obtain estimates of the covariance matrices Ω_θ, Ω_ψ, and $\Omega_{\psi\theta}$ from the relevant blocks of (A.4).

While the selection of any particular set of moment conditions for estimation using the GMM procedure is necessarily arbitrary, it is immaterial asymptotically under the null that the model is correctly specified. But in any finite sample, different estimates will emerge when different moment conditions are used.

The moment conditions that we included in f_2, related to the parameters of the Markov-switching model, were chosen using two criteria. First, that the number of moments conditions should be relatively small, and second, that the estimates lie close to the maximum likelihood estimates.

It is not possible to use a Hausman test to assess whether the GMM estimates lie close to the maximum likelihood estimates, since both estimators may be inconsistent under the alternative. Instead, we perform a Wald test by assuming that one set of estimates is a vector of constants. Let θ_{ml} and $\Sigma_{\theta,ml}$ be the maximum likelihood estimator of the endowment process parameter vector and its asymptotic covariance matrix. Suppose we view θ_T as a vector of constants and construct a Wald test for the hypothesis that $\theta_{ml} = \theta_T$. That is, we compute the Wald statistic $T(\theta_{ml} - \theta_T)' \Sigma^{-1}_{\theta, ml}(\theta_{ml} - \theta_T)$. Setting θ_T to the estimates obtained from the GMM procedure, the Wald statistic is 14.91 (p-value = 0.09).

Alternatively, we can use the estimate of the asymptotic covariance matrix of the GMM estimator to compute the Wald statistic and view θ_{ml} as a vector of constants. In this case, the Wald statistic is 16.92 (p-value = 0.05).

Finally, we mention that Hansen's test of the overidentifying restrictions does not indicate much evidence against the Markov-switching model. His J-statistic $[J = T\Phi(\lambda_T)]$ is 0.007 (p-value = 0.93).

References

Abel, A.B., 1988, Asset prices under heterogeneous beliefs: Implications for the equity premium puzzle, Mimeo. (Department of Finance, Wharton School, University of Pennsylvania, Philadelphia, PA).

Abel, A.B., 1990, Asset prices under habit formation and catching up with the Joneses, American Economic Review 80, 38–42.

Balke, N.S. and R.J. Gordon, 1986, Appendix B: Historical data, in: R.J. Gordon, ed., The American business cycle (University of Chicago Press for NBER, Chicago, IL).

Benninga, S. and A. Protopapadakis, 1990, Time preference and the 'equity premium puzzle', Journal of Monetary Economics 25, 49–58.

Brainard, W.C. and L.H. Summers, 1990, Leverage and the equity premium puzzle, Mimeo. (Department of Economics, Yale University, New Haven, CT).

Burnside, A.C., 1990, The comovement of asset returns with time non-separable preferences, Mimeo. (Department of Economics, Queens University, Kingston, Ont.).

Campbell, J.Y. and R.J. Shiller, 1987, Cointegration and tests of the present value model, Journal of Political Economy 95, 1062–1088.

Cecchetti, S.G., P.-s. Lam, and N.C. Mark, 1990, Mean reversion in equilibrium asset prices, American Economic Review 80, 398–418.

Cecchetti, S.G., P.-s. Lam, and N.C. Mark, 1992, Testing volatility restrictions on intertemporal marginal rates of substitution implied by Euler equations and asset returns, NBER technical working paper no. 124.

Constantinides, G.M., 1988, Habit formation: A resolution of the equity premium puzzle, Mimeo. (Graduate School of Business, University of Chicago, Chicago, IL).

Constantinides, G.M., 1990, Habit formation: A resolution of the equity premium puzzle, Journal of Political Economy 98, 519–543.

Epstein, L.G. and S.E. Zin, 1990, 'First-order' risk aversion and the equity premium puzzle, Journal of Monetary Economics 26, 387–407.

Flavin, M.A., 1983, Excess volatility in the financial markets: A reassessment of the empirical evidence, Journal of Political Economy 91, 929–956.

Gourieroux, C., A. Holly, and A. Monfort, 1982, Likelihood ratio test, Wald test, and Kuhn–Tucker test in linear models with inequality constraints on the regression parameters, Econometrica 50, 63–80.

Gregory, A.W. and G.W. Smith, 1988, Calibration as testing: Type I error in the equity premium puzzle, Mimeo. (Department of Economics, Queen's University, Kingston, Ont.).

Grossman, S.J. and R.J. Shiller, 1981, The determinants of the variability of stock market prices, American Economic Review 71, 222–227.

Grossman, S.J. and R.J. Shiller, 1982, Consumption correlatedness and risk measurement in economies with non-traded assets and heterogeneous information, Journal of Financial Economics 10, 195–210.

Hamilton, J.D., 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, Econometrica 57, 357–384.
Hansen, L.P., 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029–1054.
Hansen, L.P. and R. Jagannathan, 1991, Implications of security market data for models of dynamic economies, Journal of Political Economy 99, 225–262.
Kandel, S. and R.F. Stambaugh, 1990a, Modeling expected stock returns for long and short horizons, Mimeo. (Department of Finance, Wharton School, University of Pennsylvania, Philadelphia, PA).
Kandel, S. and R.F. Stambaugh, 1990b, Expectations and the volatility of consumption and asset returns, Review of Financial Studies 3, 207–232.
Kandel, S. and R.F. Stambaugh, 1991, Asset returns and intertemporal preferences, Journal of Monetary Economics 27, 39–72.
Kocherlakota, N.R., 1990a, A note on the 'discount' factor in growth economies, Journal of Monetary Economics 25, 43–48.
Kocherlakota, N.R., 1990b, On tests of representative consumer asset pricing models, Journal of Monetary Economics 26, 285–304.
Lee, B.-S. and B.F. Ingram, 1991, Simulation estimation of times-series models, Journal of Econometrics 47, 197–206.
Kydland, F. and E.C. Prescott, 1982, Time to build and aggregate fluctuations, Econometrica 50, 1345–1370.
Labadie, P.A., 1989, The effects of stochastic inflation on asset prices, Journal of Monetary Economics 24, 277–298.
Lucas, R.E., Jr., 1978, Asset prices in an exchange economy, Econometrica 66, 1429–1445.
Mankiw, N.G., 1986, The equity premium and the concentration of aggregate shocks, Journal of Financial Economics 17, 211–219.
Mankiw, N.G. and S.P. Zeldes, 1991, The consumption of stockholders and non-stockholders, Journal of Financial Economics 29, 97–112.
Mehra, R. and E.C. Prescott, 1985, The equity premium: A puzzle, Journal of Monetary Economics 15, 145–161.
Nason, J.N., 1988, The equity premium and time-varying risk behavior, Board of Governors of the Federal Reserve, Finance and economics discussion paper no. 11.
Newey, W.K. and K.D. West, 1987, A simple, positive definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.
Seigel, J.J., 1992, The real rate of interest from 1800–1990, Journal of Monetary Economics 29, 227–252.
Reitz, T.A., 1988, The equity premium: A solution, Journal of Monetary Economics 22, 117–133.
Weil, P., 1989, The equity premium puzzle and the riskfree rate puzzle, Journal of Monetary Economics 24, 401–422.
Wilson, J.W. and C. Jones, 1988, Inflation measure for the period, Mimeo. (Department of Economics and Business, North Carolina State University, Raleigh, NC).
Wolak, F.A., 1991, The local nature of hypothesis tests involving inequality constraints in nonlinear models, Econometrica 59, 981–995.