RISK PREMIUMS IN THE TERM STRUCTURE
Evidence from Artificial Economies*

David K. BACKUS

Federal Reserve Bank of Minneapolis, Minneapolis, MN 55480, USA

Allan W. GREGORY

Queen's University, Kingston, Ont. K7L 3N6, Canada

Stanley E. ZIN

Carnegie Mellon University, Pittsburgh, PA 15213, USA

We compare the statistical properties of prices of U.S Treasury bills to those generated by a theoretical dynamic exchange economy with complete markets. We show that the model can account for neither the sign nor the magnitude of average risk premiums in forward prices and holding-period returns. The economy is also incapable of generating enough variation in risk premiums to account for rejections of the expectations hypothesis with Treasury bill data These conclusions add to the growing list of empirical deficiencies of the representative agent model of asset pricing

1. Introduction

One of the more firmly established facts of financial economics is that the expectations hypothesis of the term structure of interest rates cannot account for observed fluctuations in multiperiod bond returns: forward rates are not simply predictors of future spot rates. Viewed as predictors, forward rates consistently overestimate future spot rates, and the 'forecast errors' are sys-

tematically related to variables that are known when the forecast is made. The consensus in the profession seems to be that forward rates contain, besides forecasts of future spot rates, risk premiums that change through time.

Risk premiums on forward contracts are not directly observable, but there are a number of pieces of evidence suggesting that they vary considerably over time. Consider the linear regression (error term omitted)

$$r_{t+1} - r_t = a + b\left(r_t^f - r_t\right),\tag{1}$$

where r is the one-period spot rate of interest and r^f is the one-period-ahead forward rate (and an error term is omitted). Under the expectations hypothesis, forward rates are market expectations of future spot rates, which implies the coefficient restrictions $a = 0$ and $b = 1$. The implicit alternative hypothesis is generally taken to be that the forward rate is the sum of the expected future spot rate and a risk premium. The evidence, presented in Fama (1984, table 4), is that neither restriction is supported by the data; related work by Startz (1982) and others is reviewed by Melino (1987) and Shiller and McCulloch (1987). Rejection of the first restriction, $a = 0$, might be explained by a nonzero but constant risk premium, but rejection of the second requires, under the alternative, a risk premium that varies through time and is correlated with the forward premium $r_t^f - r_t$.

We ask whether this interpretation of the evidence is consistent with a specific theory of the risk premium: the general equilibrium theory of asset pricing developed by Merton (1973), Lucas (1978), Breeden (1979), Brock (1982), and Cox, Ingersoll, and Ross (1985). Our approach, by focusing on the connection between prices and quantities, brings out different aspects of the theory than those emphasized in recent studies of prices alone by Brown and Dybvig (1986), Stambaugh (1986), and Gibbons and Ramaswamy (1986). The issue, in terms of the empirical literature, is whether a numerical version of the theory can account for the variation in risk premiums implicit in rejections of the expectations hypothesis. Because risk premiums are not directly observable, the theory has been difficult to confirm or reject. Most of the evidence cited in its favor, including the regression test just described, consists of rejections of the expectations hypothesis. It has the same uncertain status as any alternative hypothesis in a statistical test: rejecting the null does not confirm the alternative. We examine the alternative hypothesis directly in a monetary version of the Mehra–Prescott (1985) economy. In the artificial economy, we know exactly what risk premiums are at all times and can determine the extent of their influence by comparing regressions with and without them. We cannot say whether time-varying risk premiums are the cause of rejections of the expectations hypothesis with real-world data, but we can say whether the model is capable of generating risk premiums that produce similar results in the theoretical economy.

In the following four sections, we review the empirical record, build an artificial economy, and compare its statistical properties to those of postwar quarterly time-series data for the United States. In section 2, we survey empirical work on the term structure and report sample regressions with U.S. Treasury bill prices. We use prices, rather than the more conventional rates, to provide a closer fit with the theory, but the choice makes little difference. Both prices and rates suggest that risk premiums are nonzero on average and vary over time.

In section 3, we describe a monetary version of the Mehra–Prescott (1985) economy and derive its implications for prices of risk-free real and nominal discount bonds of different maturities. Two propositions summarize the behavior of risk premiums in this economy under fairly general conditions: the risk premium's sign is determined by the autocorrelation of two-period marginal rates of substitution, and its variability is highest at the short end of the maturity spectrum.

In section 4, we use a parametric two-state example to illustrate two empirical anomalies, features of the data that are difficult to reconcile with the theoretical model. We show, first, that to generate average risk premiums as large in absolute value as those in the Treasury bill market, the coefficient of relative risk aversion must be at least 8 or 10, values viewed by many as implausibly large. A similar anomaly has been described for equity premiums by Mehra and Prescott (1985) and for holding-period yields on a variety of assets by Grossman, Melino, and Shiller (1987). A second difficulty, which has not to our knowledge been noted elsewhere, concerns the sign of the average risk premium. In the theoretical economy, risk premiums are positive if and only if growth rates of consumption are negatively autocorrelated. Since quarterly consumption growth rates are virtually uncorrelated, the model cannot account for the positive average risk premiums apparent in the data.

We return to the variability of risk premiums in section 5. We argue there that the theoretical economy cannot generate sufficient variability to account for rejections of the expectations hypothesis without also generating implausible values for mean risk premiums. This is true even if some parameters, like the autocorrelation of consumption growth rates and the coefficient of relative risk aversion, are chosen, despite evidence to the contrary, to overcome the anomalies just noted. The analysis uses the theoretical economy to produce artificial time series of bond prices, which we use to estimate and conduct diagnostic tests of regressions reported in studies with actual data. Using the artificial data, we examine the influence of risk premiums on these results. We find that regression tests are unable to reject the expectations hypothesis with reasonable, or even some unreasonable, parameter values: the hypothesis $b = 1$ in eq. (1) is retained even in models with sizable risk premiums.

In short, the theory in its present form cannot explain the sign, the magnitude, or the variability of risk premiums observed at the short end of the

term structure of interest rates. We conclude with a brief summary and a few remarks on methodology.

2. Risk premiums in Treasury bill data

We begin by reviewing empirical work on the term structure and reporting sample statistics for quarterly U.S. Treasury bill prices. We use prices rather than rates because they fit more naturally into our theoretical environment.

We denote the spot price of a k-period Treasury bill or bond by q_{kt}. Forward prices are defined by

$$f_{kt} \equiv q_{k+1,t}/q_{kt} \quad \text{for} \quad k = 1, 2, \ldots . \tag{2}$$

These prices are related to interest rates by

$$q_{1t} \equiv 1/(1 + r_{1t}), \qquad f_{kt} \equiv 1/(1 + r_{kt}^{f}),$$

where r_{1t} is the one-period spot rate and r_{kt}^{f} is the k-period-ahead forward rate. The (gross) holding-period return – the return from holding a k-period bond for one period – is

$$h_{kt} \equiv q_{k-1,t+1}/q_{kt} \quad \text{for} \quad k = 2, 3, \ldots . \tag{3}$$

For $k = 1$ the holding-period return is just 1 plus the one-period spot rate of interest. The excess return on a k-period bond is $h_{kt} - h_{1t}$.

Both forward prices and holding-period returns can be viewed as combinations of forecasts and risk premiums. With respect to the former, we define the risk premium, fp, on a k-period forward contract by

$$fp_{kt} \equiv E_t q_{1t+k} - f_{kt}, \tag{4}$$

the difference between the forward price and the expected future spot price. The operator E_t denotes the expectation conditional on the date t information set, which we assume includes the past and present values of spot and forward prices. Note that (4) allows explicitly for the possibility that the risk premium varies with t. The sign convention means that the risk premium is positive when the forward rate exceeds the expected future spot rate, and the reverse is true for prices. Similarly, we define the risk premium, hp, implicit in the holding-period return h_{kt} as the expected excess return:

$$hp_{kt} \equiv E_t h_{kt} - h_{1t}.$$

It, too, can vary with time.

Table 1

U.S. Treasury bill prices and yields: Selected statistics.[a]

	Variable				
	Spot prices		Forward price f_{1t}	Risk premium $q_{1t+1} - f_{1t}$	Excess holding-period return $h_{2t} - h_{1t}$
Statistic	3-month q_{1t}	6-month q_{2t}			
	1959:2–1986:2				
Mean	0.98432	0.96749	0.98285	0 00142	0.00149
	(0.0021)	(0.0041)	(0.0021)	(0.0003)	(0.0003)
Standard deviation	0.00757	0.01489	0.00769	0.00357	0.00379
	(0.0014)	(0.0026)	(0.0013)	(0.0008)	(0.0009)
Autocorrelation	0.91	0.91	0.90	−0 08	−0.08
	(0 0421)	(0.0451)	(0.0486)	(0.1351)	(0.1361)
	1959:2–1972:4				
Mean	0.98953	0.97790	0.98823	0.00115	0.00117
	(0.0011)	(0.0022)	(0 0011)	(0 0003)	(0.0003)
Standard deviation	0.00334	0.00670	0.00352	0.00180	0.00185
	(0 0006)	(0.0013)	(0.0006)	(0.0003)	(0.0003)
Autocorrelation	0.89	0.88	0.84	0.11	0.10
	(0.0515)	(0.0546)	(0.0596)	(0.1545)	(0.1540)
	1973:1–1986:2				
Mean	0.97902	0.95690	0.97736	0.00171	0.00181
	(0.0021)	(0 0041)	(0.0020)	(0.0006)	(0.0006)
Standard deviation	0.00697	0.01339	0.00687	0.00475	0.00506
	(0.0011)	(0.0021)	(0.0011)	(0.0010)	(0.0011)
Autocorrelation	0.80	0.80	0.78	−0.12	−0.12
	(0.0798)	(0.0897)	(0 0997)	(0.1423)	(0.1424)

[a] Numbers in parentheses are Newey–West (1987) standard errors, computed by GMM using four lags of the autocorrelation function. Bill prices and returns are computed from monthly forward rates in the six-month Fama file of the 1987 CRSP tape.

In table 1, we report sample moments for selected U.S. Treasury bill prices and returns. The time interval is a quarter, so q_{1t} is the price of a three-month Treasury bill. Fama (1984) reports similar statistics on a monthly basis. The average price of a three-month Treasury bill over the sample period is 0.98432, which corresponds to an interest rate of 1.59 percent per quarter. The last two columns of table 1 give information about the two risk premiums. The fourth variable, $q_{1t+1} - f_{1t}$, consists of a forecast error plus a risk premium. If forecast errors have zero mean conditionally, then they do unconditionally as well, and the sample mean of this variable is an estimate of the mean risk premium over the sample period. We see that the mean risk premium is small,

but positive and statistically significant. And since forecast errors are serially uncorrelated, risk premiums must be negatively autocorrelated in the sample to account for the insignificant negative autocorrelation in the table. The estimated autocorrelation should understate (in absolute value) that of the risk premium, since it combines the time dependence of the risk premium with that of a serially uncorrelated forecast error. The last variable in the table is also the sum of a risk premium and a forecast error, and its properties are similar.

There is evidence, too, that the risk premium not only is nonzero, but also varies substantially over time. Most such work on the term structure, surveyed recently by Melino (1987) and Shiller and McCulloch (1987), has been oriented around the expectations hypothesis. Initially this hypothesis was understood to mean that risk premiums are zero, but in light of evidence like that in table 1, it was amended to mean simply that risk premiums are constant: they do not vary with the state of the economy. Even so, a large empirical literature has found the hypothesis wanting. As Shiller and McCulloch (1987, p. 61) remark: 'Empirical work has produced consensus on little more than that the...expectations [hypothesis]...can be rejected.'

Most of the evidence against the expectations hypothesis is based on regressions relating forward and spot prices to other variables. In the most popular, the difference between forward prices (or rates) and corresponding future spot prices (or rates) is examined for predictability. If the expectations hypothesis is valid, this difference is (except possibly for a nonzero mean) a pure forecast error and should not be predictable by anything in the current information set. Several tests are constructed by choosing different variables from this information set. The simplest is based on this regression (error term omitted):

$$q_{1t+1} - f_{1t} = a. \tag{5}$$

Since lagged forecast errors are part of the information set, autocorrelation of residuals from (5) is evidence against the expectations hypothesis. Other tests use regressions of the form

$$q_{1t+1} - f_{1t} = a + bx_t, \tag{6}$$

where x_t is a subset of the information available at date t. A test of the coefficient restriction $b = 0$ is a test of the expectations hypothesis. The forward premium, $f_{1t} - q_{1t}$, is a common choice for x_t. This regression often appears in the literature as

$$q_{1t+1} - q_{1t} = a + (b + 1)(f_{1t} - q_{1t}), \tag{7}$$

which is (6) with the forward premium added to both sides. Since least squares

is invariant to such linear transformations, the implied estimator of b is identical to that in (6), but it provides a useful interpretation of the parameter. If b is nonzero, we can reject the expectations hypothesis, but for all values except -1 the forward premium is still helpful in predicting future changes in spot prices. A second variant of (6) uses the lagged dependent variable as x, which under the expectations hypothesis is simply the lagged forecast error plus a constant. There is no presumption in any of these equations, even under the expectations hypothesis, that the errors are homoskedastic.

The same kinds of tests have been applied to excess holding-period returns. Examples are the regressions

$$h_{2t} - h_{1t} = a, \tag{8}$$

$$h_{2t} - h_{1t} = a + bx_t, \tag{9}$$

for some choice of x. Below we use both the forward premium and the lagged dependent variable. The expectations hypothesis requires that the residuals be serially uncorrelated and the coefficient b equal zero. Again, the errors need not be homoskedastic.

Estimates of eqs. (5)–(9) have been remarkably uniform across studies and time periods. The constants in (5) and (8) are usually small, as we saw in table 1, but significantly positive [Roll (1970, ch. 5), Fama (1976, table 1)]. Tests of the hypothesis $b = 0$ in eqs. (6) and (9) frequently reject, especially when the independent variable is the forward premium. Numerical estimates of b in this case differ across studies. In Fama's (1984, table 4) study of Treasury bill returns at monthly intervals, reported in the form of eq. (7), the estimates of $b + 1$ are between 0.11 and 0.46 for the complete sample. This suggests that, although the expectations hypothesis can be rejected, the forward premium helps to predict future changes in spot rates, at least over short forecast horizons. Shiller, Campbell, and Schoenholtz (1983, table 3, rows 1–5) estimate eq. (7) with quarterly Treasury bill returns and report similar results. The residuals in these regressions typically exhibit autoregressive conditional heteroskedasticity [Engle, Lilien, and Robins (1987)].

These features are illustrated in table 2 for quarterly U.S. Treasury bill prices. We report the six regressions mentioned earlier and subject the residuals to two kinds of analysis. The first is the Durbin–Godfrey LM test for serial correlation; we report marginal significance levels for tests of first- and fourth-order time dependence (AR1 and AR4). The second is Engle's (1982) LM test for autoregressive conditionally heteroskedastic (ARCH) errors, again for orders one and four. We also report results of Wald tests of the hypotheses $a = b = 0$, $b = 0$, and $b = -1$, all based on heteroskedasticity-consistent standard errors [White (1980)]. Since the data are far from homoskedastic, this correction is essential. None of the test statistics has precisely the distribution

Table 2

Regressions of $y = a + bx$ with U.S. Treasury bill prices and yields [a]

Variable or test	Estimated coefficient or statistic					
y	$q_{1t+1} - f_{1t}$	$q_{1t+1} - f_{1t}$	$q_{1t+1} - f_{1t}$	$h_{2t} - h_{1t}$	$h_{2t} - h_{1t}$	$h_{2t} - h_{1t}$
x	None	$f_{1t} - q_{1t}$	$q_{1t} - f_{1t-1}$	None	$f_{1t} - q_{1t}$	$h_{2t-1} - h_{1t-1}$
a	0.0014 (0.0003)	−0.0002 (0.0005)	0.0015 (0.0005)	0.0015 (0.0005)	−0.0018 (0.0004)	0.0017 (0.0005)
b	— —	−1.097 (0.222)	−0.076 (0.175)	— —	−0.148 (0.235)	−0.083 (0.174)
s	0.0036	0.0032	0.0036	0.0038	0.0034	0.0038
DW	2.15	2.50	2.02	2.16	2.52	2.02
$AR1$	0.439	0.006	0.202	0.397	0.005	0.172
$AR4$	0.041	0.000	0.045	0.032	0.000	0.040
$ARCH1$	0.002	0.002	0.003	0.003	0.004	0.005
$ARCH4$	0.022	0.022	0.025	0.031	0.033	0.033
Wald ($a = b = 0$)	—	10^{-8}	0.0002	—	10^{-8}	0.0002
Wald ($b = 0$)	—	10^{-6}	0.665	—	10^{-6}	0.636
Wald ($b = -1$)	—	—	10^{-6}	—	—	—

[a] Data are from the CRSP tape, 1959:3 to 1986:2 (108 observations). Numbers in parentheses are heteroskedasticity-consistent standard errors, s is the estimated standard error of the regression, DW is the Durbin–Watson statistic, and ARn and $ARCHn$ are marginal significance levels from tests for serial correlation and autoregressive conditional heteroskedasticity of order n. Wald is the marginal significance level for a Wald test of the indicated hypothesis based on a heteroskedasticity-consistent covariance matrix estimator

used to compute tail probabilities, even under the null, since they are based on maintained hypotheses that need not be true. The Durbin–Godfrey tests, for example, are based on independent normal errors, so they may be sensitive to autoregressive conditional heteroskedasticity as well as autocorrelation. Nevertheless, they provide useful information beyond that given by regression coefficients and standard errors.

In table 2, both the forward-spot differential and the ex post holding premium have small positive means, as we know from table 1. The value of b is significantly negative in the forward-premium regressions. The dependent variables are negatively autocorrelated, but the estimated coefficient is not significant. The residuals in all regressions indicate serial correlation, especially of the fourth order, and autoregressive conditional heteroskedasticity. The hypotheses $a = b = 0$, $b = 0$, and $b = -1$ are all rejected by the data in the forward-premium regressions. Thus, the expectations hypothesis is rejected while the ability of the forward premium to forecast future changes in spot prices is upheld. The nominal sizes of these tests are probably overstated,

given the serial correlation and heteroskedasticity in the residuals, but it is hard to avoid the conclusion that the data contradict the expectations hypothesis. Evidence from monthly data [Fama (1984)] is even stronger.

The regressions also provide direct estimates of the variability of risk premiums. Consider the forward risk premium, fp_1. Eq. (4) implies that

$$q_{1t+1} - f_{1t} = fp_{1t} + \varepsilon_{t+1},$$

where ε_{t+1} is a forecast error with conditional and unconditional mean zero. Hence,

$$\text{var}(q_{1t+1} - f_{1t}) = \text{var}(fp_{1t}) + \text{var}(\varepsilon_{t+1}).$$

Since the risk premium can be predicted from date t information but the forecast errors cannot, an estimated lower bound on the variability of the forward risk premium is the variability of the fitted values in any of the regressions explaining $q_{1t+1} - f_{1t}$. In the second regression of table 2, the implied standard deviation is 0.00157; see Startz (1982) for a similar calculation with monthly data. An analogous exercise yields 0.00164 as an estimated lower bound on the standard deviation of the holding premium. In both cases, the lower bound is of the same order as, and slightly larger than, the estimated mean risk premium, so the variability is large in that sense. These estimated lower bounds are subject to sampling variability, but they help to quantify the extent of variation in risk premiums present in the data.

3. A theoretical economy

The question we have asked is whether the behavior of bond prices in general and the variability of risk premiums in particular are consistent with the general equilibrium pricing theory of Merton (1973), Lucas (1978), Breeden (1979), Brock (1982), and Cox, Ingersoll, and Ross (1985). Our version of this theory is a monetary extension of Mehra and Prescott (1985), a stationary, recursive, pure-exchange version of an Arrow–Debreu economy. In this economy, a stochastic endowment of a single commodity is consumed by a single representative agent, whose preferences are additively separable over time and obey the expected utility axioms. With quantities exogenous, the equilibrium allocation is determined by supply. Relative prices of dated, state-contingent claims are derived from marginal rates of substitution. Prices of composite assets, like risk-free bonds, are computed by combining the relevant contingent claims. Nominal claims – and, implicitly, money – are introduced toward the end of the section. We derive prices of risk-free bonds of different maturities and use them to define forward prices, holding-period returns, and risk premiums.

The size of the endowment each period is described by a stationary Markov chain in growth rates. Output, y_t, evolves according to

$$y_{t+1} = x_{t+1} y_t, \tag{10}$$

and the distribution of its rate of growth, x_t, depends only on its previous rate of growth. The state of the economy at date t is known to the representative agent and is described completely in the real economy by the state, $s_t \equiv (y_t, x_t)$. Here, as in Mehra and Prescott (1985), growth rates take on a finite number of values, denoted λ_t. In principle this is not very restrictive, and the gain in computational simplicity is considerable. The stochastic behavior of output is characterized by a matrix Π of transition probabilities for growth rates, with typical element

$$\pi_{ij} \equiv \text{prob}(x_{t+1} = \lambda_j \mid x_t = \lambda_i).$$

Probabilities of events more than one period in the future are computed from powers of Π:

$$\text{prob}(x_{t+k} = \lambda_j \mid x_t = \lambda_i) = [\Pi^k]_{ij}.$$

These and other features of Markov chains are described in Feller (1968, ch. XV) and Isaacson and Madsen (1976).

The stochastic endowment is consumed by a representative agent whose preferences are characterized by the expected utility function

$$U = E_t \sum_{k=0}^{\infty} \beta^k u(c_{t+k}),$$

where E_t denotes the expectation conditional on the histories as of date t of all variables in the economy. With the Markov structure, this is equivalent to conditioning on s_t alone. To maintain stationarity of prices in a growing economy, we specify further that U is homothetic. With additive separability over time and states, this implies the class of power subutility functions, $u(c) = [c^{1-\alpha} - 1]/(1 - \alpha)$. The parameter $\alpha \geq 0$ governs both risk aversion and the substitutability of consumption over time.

An equilibrium in this economy is a set of state-contingent prices for which consumption equals the endowment at all dates and in all states. Relative prices are computed by equating them to marginal rates of substitution and evaluating those rates at exogenous equilibrium quantities. If the current state is (y_t, λ_i), then the relative price of one unit of the commodity next period, if

state j occurs, to one unit now is

$$\pi_{ij}\beta u'(y_{t+1})/u'(y_t) = \pi_{ij}\beta u'(\lambda_j y_t)/u'(y_t),$$

a function of the current state. With power utility, this price is $\pi_{ij}\beta\lambda_j^{-\alpha}$, which depends, through the transition probability, on the growth rate λ_j, but not on the level of output. Our interest lies in prices of risk-free assets. We examine both real bonds, which yield one unit of the commodity in some future period in all states, and nominal bonds, which yield one unit of money. Prices in both cases are computed by summing prices of pure contingent claims. Consider real bonds first. The price, q_{1t}, of a one-period, date t, risk-free discount bond, which delivers one unit of the commodity at date $t+1$ in all states, is

$$q_{1t} = \sum_j \pi_{ij}\beta u'(y_{t+1})/u'(y_t) = E_t \pi_{ij}\beta u'(y_{t+1})/u'(y_t).$$

If we define $m_{t+1} \equiv \beta u'(y_{t+1})/u'(y_t)$, we can express this more compactly as $q_{1t} = E_t m_{t+1}$. With power utility, $m_{t+1} = \beta x_{t+1}^{-\alpha}$, so that m depends on the rate of growth of output, but not on its level. The bond price is simply a function of the current growth rate, say, $q_{1t} = q_1(\lambda_t)$ for $x_t = \lambda_t$.

Prices of long-term bonds are derived the same way. A k-period real bond, yielding one unit of the commodity k periods hence in all states, is priced according to

$$q_{kt} = E_t \prod_{i=1}^{k} m_{t+1} = E_t[m_{t+1}q_{k-1,t+1}]. \tag{11}$$

Since the distribution over future m's depends only on the current growth rate, this defines pricing functions $q_k(\lambda_t)$ for $k = 2, 3, \ldots$.

Implicit in these pricing functions are analogous functions for forward prices and risk premiums. The forward pricing function is, by analogy with eq. (2),

$$f_k(\lambda_t) \equiv q_{k+1}(\lambda_t)/q_k(\lambda_t). \tag{12}$$

The risk premium, fp, defined by eq. (4), can also be written as

$$fp_{kt} = -\text{cov}_t\left[\prod_{i=1}^{k} m_{t+i}, q_{1t+k}\right]\bigg/ q_{kt},$$

where cov_t is the conditional covariance operator. Since the distribution with which the covariance is computed depends on the current state, the forward premium is a function of the state and is, in that sense, time-varying.

The behavior of bond prices, returns, and risk premiums depends on the parameters of the model: α, β, Π, and the λ's. Two useful features of this relationship are summarized in the following propositions.

Proposition 1.   If either (i) the representative consumer is risk-neutral ($\alpha = 0$) or (ii) marginal rates of substitution, m, are independent, then forward premiums fp and holding premiums hp are zero in all states.

Proof.   If (i) the representative consumer is risk-neutral, then $m = \beta$ in all states. Bond prices, forward prices, and holding-period yields are, then, constant and risk premiums are zero in all states. If, however, (ii) marginal rates of substitution are independent, as they will be in our economy if growth rates are independent, then the price of a k-period bond is simply the product of expected future spot prices:

$$q_{kt} = E_t \prod_{i=1}^{k} m_{t+i} = \prod_{i=1}^{k} E_t m_{t+i} = \prod_{i=1}^{k} E_t q_{1,t+i-1}.$$

Then forward prices equal expected future spot prices, and risk premiums on forward contracts are zero. The expected one-period yield on an n-period bond is, by the same argument,

$$E_t h_{kt} = E_t \prod_{i=1}^{k-1} m_{t+i} \Big/ E_t \prod_{i=1}^{k} m_{t+i} = 1 / E_t m_{t+k}.$$

Since the transition probabilities do not depend on the date, $E_t m_{t+1} = E_t m_{t+2} = \cdots = E_t m_{t+k}$. Holding-period returns are equal on bonds of all maturities, and holding premiums are zero.   ∎

Proposition 1 says that with serially uncorrelated marginal rates of substitution bond prices in the theoretical economy behave according to the expectations hypothesis. To generate the nonzero time-varying risk premiums that we observe in the U.S. Treasury bill market, the conditional distribution over future marginal rates of substitution must depend on the current state. Furthermore, the sign of the risk premium is determined by the form of that dependence: risk premiums on forward contracts are positive (negative) if marginal rates of substitution are negatively (positively) autocorrelated.

Beyond that, there is little we can say without placing more structure on the model. We can, however, show that forward risk premiums converge in the following sense.

Proposition 2. If the Markov chain in growth rates is ergodic, then as k approaches infinity, the forward pricing function $f_k(\lambda)$ converges to a constant, f, which does not depend on λ.

Proof. Consider the pricing functions for bonds. In state ι, the price of a one-period bond is $\sum_j \pi_{\iota j} \beta \lambda_j^{-\alpha}$. We can express this more compactly as $\sum_j b_{\iota j}$, where $b_{\iota j} \equiv \pi_{\iota j} \beta \lambda_j^{-\alpha}$ defines implicitly a matrix, B. Similarly, the price of a two-period bond is

$$\sum_j \sum_k \pi_{\iota j} \pi_{jk} \beta^2 \lambda_j^{-\alpha} \lambda_k^{\alpha},$$

or $\sum_j b_{\iota j}^{(2)}$, where $b_{\iota j}^{(2)}$ denotes the ijth element of B^2. More generally, the price of a k-period bond in state i can be expressed as $\sum_j b_{\iota j}^{(k)}$. Since the Markov chain is ergodic, the Perron–Frobenius theorem guarantees that the dominant eigenvalue of B is positive and that any positive vector operated on by powers of B will eventually approach the associated eigenvector and grow at the rate of this eigenvalue. Now consider the k-period forward pricing function, the ratio of $(k + 1)$-period to k-period bond pricing functions. As k gets large, this converges to the dominant eigenvalue of B regardless of the current state. ∎

The intuition behind this result is that in an ergodic economy forward contracts for long horizons look alike. Each such contract specifies a payment of f in, say, period $t + k$ and a receipt of 1 in the following period, regardless of what states occur in those periods. For an ergodic chain, the distribution over these states is, for large k, the long-run equilibrium distribution of the Markov chain and does not depend on the current state.

An immediate corollary of Proposition 2 is that forward risk premiums also converge: since the Markov chain is ergodic, the expected future spot price, $E_t q_{1t+k}$, converges to the unconditional mean of q_1, and the forward premium, $fp_k(\lambda_\iota)$, converges to the difference. Of course, if forward prices converge, then the yield to maturity on a long-term bond, the geometric mean of forward prices, also converges. It is not true, however, that holding premiums converge in a similar manner. Instead, the distribution over holding premiums converges, so that holding premiums on two long-term bonds of different maturities are similar functions of the current state.

Nominal bond pricing can be handled in much the same way as real bond pricing. The primary issue is how to get money – and, therefore, nominal bonds – into the economy. To be concrete about this, we imagine following Lucas (1980) in using cash-in-advance constraints. If monetary growth rates are described by a Markov chain, then the equilibrium will be characterized by a Markov chain in consumption growth rates and inflation. The state then consists of three variables – output, its growth rate, and inflation – which we

can describe by a Markov chain in the latter two. The only restriction on the process is that nominal bond prices, to be defined shortly, cannot exceed one in any state (that is, nominal interest rates must be nonnegative). This theoretical structure is described more completely in Sargent (1987, ch. 5) and Backus and Kehoe (1987, sec. 5).

Consider the currency price of a one-period nominal bond, a claim to one unit of currency next period in all states. The value in date t marginal utility units of one dollar in period $t + 1$ is $\beta E_t[u'(c_{t+1})/p_{t+1}]$, where p_{t+1} is the (state contingent) currency price of one unit of the commodity. Similarly, the value of one unit of currency at date t is $u'(c_t)/p_t$. The equilibrium price of a one-period nominal bond is the ratio

$$q_{1t} = E_t\{[\beta u'(y_{t+1})/u'(y_t)][p_t/p_{t+1}]\} \equiv E_t n_{t+1},$$

with the obvious definition of n_{t+1}, the nominal marginal rate of substitution. Prices of multiperiod nominal bonds are, by analogy with eq. (11),

$$q_{kt} = E_t \prod_{i=1}^{k} n_{t+i} = E_t n_{t+1} q_{k-1, t+1}.$$

From these we compute forward prices and holding-period yields as before.

From this point on, the mathematical treatments of real and nominal bonds are almost identical. As in Proposition 1, risk premiums on nominal bonds are driven by persistence in the marginal rate of substitution, given here by n. If there is no autocorrelation in n, risk premiums are zero and the expectations hypothesis holds. Risk neutrality, however, is no longer sufficient to eliminate risk premiums: with $\alpha = 0$, m is constant as before, but fluctuations in the rate of inflation may lead to systematic differences between forward and expected future spot prices. The same effect has been noted in the pricing of forward foreign exchange contracts, which are also expressed in nominal terms.

Proposition 2 changes in the natural way: forward prices converge if the Markov chain for consumption growth rates and inflation together is ergodic. The theory is the same, but as a practical matter, the assumption of ergodicity may be less plausible with nominal variables than with real ones.

4. A two-state example

In this section, we specialize the model and compare its properties to those of U.S. time series. Using a two-state example, we point out two features of U.S. data that are not consistent with quantitative properties of the theory.

We assume, first, that the transition matrix has elements

$$\pi_{i,j} = (1 - \theta)\pi_j + \theta\delta_{i,j},\tag{13}$$

where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise; $\pi_j \geq 0$; and $\sum_j \pi_j = 1$. Barton, David, and Fix (1962) refer to this parameterization as *simple persistence*. Clearly, all the transition probabilities must be nonnegative, which places restrictions on the range of θ. In the two-state case, (13) places no restrictions on the form of the transition matrix, but it does lead to a useful interpretation of the parameters. When all transition probabilities are positive, conditional probabilities converge to a unique invariant, or long-run equilibrium, distribution, with probability π_j of being in state j in the distant future. The parameter θ governs persistence: if θ is zero, states are independent over time; positive and negative values indicate positive and negative autocorrelation of states. From Proposition 1, we know that nonzero values of θ are a prerequisite for nonzero risk premiums.

The preference parameters, α and β, enter the calculations through the marginal rate of substitution. To make this as simple as possible, we begin by considering an economy in which the price level, p, is constant. Then, if consumption growth rates take on two values, λ_1 and λ_2, say, the marginal rate of substitution, n, assumes the values μ_1 and μ_2 defined by $\mu_j = \beta\lambda_j^{-\alpha}$. Larger values of α, indicating greater risk aversion, lead in general to greater variation in n. The discount factor, β, is simply a scale factor.

Now consider bond prices when the equilibrium distribution is symmetric, so that the long-run probability of each state is one-half. We can express the values of n in the two states as $a - b$ and $a + b$, where a and b are, respectively, the mean and standard deviation of n. Bond prices, forward prices, and holding-period returns can then be computed using eqs. (11), (12), and (3). Their values, which can be found by tedious but straightforward calculation, are summarized in table 3. Note, in particular, that the one-period forward premium in both states is approximately

$$-\theta(1 - \theta^2)b^2,$$

since three-month bond prices are close to 1. (The range in our sample is from 0.962 to 0.995.) Thus, the size of the risk premium in our example depends in a nonlinear way on θ and is proportional to the variance of n. As we noted in Proposition 1, the risk premium is zero when $\theta = 0$.

Our objective is to compare the theory – particularly, the two-state example – with U.S. data on consumption growth rates and Treasury bill prices. In table 4 we report quarterly growth rates of consumption and consumption price deflators for the postwar period. From the mean, standard deviation, and autocorrelation of the consumption growth rate, we can get an

Table 3

Pricing formulas for the two-state example.

Random variable	Value of variable		
	State 1	State 2	Mean
m	$a - b$	$a + b$	a
q_1	$a - \theta b$	$a + \theta b$	a
q_2	$a^2 - \theta(1+\theta)ab + \theta b^2$	$a^2 + \theta(1+\theta)ab + b^2$	$a^2 + \theta b^2$
E,q_{1r+1}	$a - \theta^2 b$	$a + \theta^2 b$	a
f_1	q_2/q_1	q_2/q_1	$-^{[a]}$
fp_1	$-\theta(1-\theta^2)b^2/(a-\theta b)$	$-\theta(1-\theta^2)b^2/(a+\theta b)$	$a[a^2 - \theta(1 - \theta - \theta^2)b^2]/(a^2 - \theta^2 b^2)$
h_1	$1/(a-\theta b)$	$1/(a+\theta b)$	$1/(a^2 - \theta^2 b^2)$
h_2	$(a-\theta^2 b)/[a^2 - \theta(1+\theta)ab + \theta b^2]$	$(a+\theta^2 b)/[a^2 + \theta(1+\theta)ab + \theta b^2]$	$-^{[a]}$
hp_2	$\dfrac{-\theta(1-\theta^2)^2 b}{a^3 - \theta(2+\theta)a^2 b + \theta(1+\theta+\theta^2)ab^2 - \theta^2 b^3}$	$\dfrac{-\theta(1-\theta^2)b^2}{a^3 + \theta(2+\theta)a^2 b + \theta(1+\theta+\theta^2)ab^2 + \theta^2 b^3}$	$-^{[a]}$

[a]These expressions are averages of the variable's value in the two states.

Table 4

Per capita consumption and prices: Selected statistics.[a]

Statistic	Value for selected statistics			
	Total	Services	Nondurables	Durables
Per capita consumption growth (x_t)				
1950:1–1986:2				
Mean	1.0053	1.0065	1.0032	1.0087
	(0 0008)	(0.0005)	(0.0007)	(0.0024)
Standard deviation	0.0083	0.0053	0.0084	0.0426
	(0.0010)	(0.0004)	(0.0008)	(0.0068)
Autocorrelation	0.07	0.06	0.08	−0.09
	(0.1441)	(0.0659)	(0.1569)	(0.0909)
1959:2–1972:4				
Mean	1.0068	1.0080	1.0043	1.0120
	(0.0011)	(0.0006)	(0.0012)	(0.0043)
Standard deviation	0.0064	0.0043	0.0072	0.0311
	(0.0005)	(0.0005)	(0.0007)	(0.0037)
Autocorrelation	0.13	−0.01	0.19	−0.12
	(0.0877)	(0.1431)	(0.0956)	(0.1445)
1973:1–1986:2				
Mean	1.0043	1.0052	1 0022	1.0079
	(0.0016)	(0.0008)	(0.0016)	(0.0057)
Standard deviation	0 0080	0.0049	0.0083	0.0372
	(0.0009)	(0.0005)	(0.0010)	(0.0042)
Autocorrelation	0.22	−0.08	0.31	−0.07
	(0.0546)	(0.0714)	(0.0791)	(0.1058)
Price level growth (p_t/p_{t-1}) (continued)				
1950:1–1986:2				
Mean	1.0103	1.0122	1 0094	1.0071
	(0.0017)	(0.0016)	(0.0021)	(0.0016)
Standard deviation	0.0080	0.0065	0 0108	0.0123
	(0.0009)	(0.0008)	(0 0014)	(0.0019)
Autocorrelation	0.60	0.83	0.65	0.08
	(0.1956)	(0.0484)	(0.0800)	(0.1596)
1959:2–1972:4				
Mean	1.0071	1.0084	1.0066	1.0042
	(0.0011)	(0.0012)	(0.0011)	(0.0010)
Standard deviation	0.0038	0.0041	0 0046	0.0070
	(0.0002)	(0.0004)	(0.0003)	(0.0008)
Autocorrelation	0.65	0.66	0.34	−0.04
	(0.0972)	(0.0797)	(0.1185)	(0.1567)

Table 4 (continued)

Statistic	Value for selected statistics			
	Total	Services	Nondurables	Durables

Price level growth (p_t/p_{t-1}) (continued)

1973:1–1986:2

Mean	1.0164	1.0183	1.0154	1.0119
	(0.0021)	(0.0015)	(0.0035)	(0.0082)
Standard deviation	0.0072	0.0050	0.0129	0.0083
	(0.0008)	(0 0007)	(0.0018)	(0.0012)
Autocorrelation	0 75	0.82	0.65	0.80
	(0.0621)	(0 0542)	(0.0521)	(0.0760)

	Value for selected levels of risk aversion (α)			
	2	5	10	20

Nominal marginal rates of substitution (n_t)

1950:1–1986:2

Mean	0.9900	0.9900	0.9900	0.9900
	(0.0019)	(0.0037)	(0.0075)	(0.0160)
Standard deviation	0 0177	0.0415	0.0837	0.1745
	(0 0029)	(0.0059)	(0.0113)	(0.0224)
Autocorrelation	−0.06	−0.02	0 01	0.02
	(0.2023)	(0.1649)	(0.1536)	(0.1452)

1959:2–1972:4

Mean	0.9900	0 9900	0.9900	0.9900
	(0 0025)	(0.0058)	(0.0113)	(0.0225)
Standard deviation	0.0131	0.0317	0.0630	0.1260
	(0.0013)	(0.0029)	(0.0056)	(0.0117)
Autocorrelation	0.21	0.15	0.13	0.12
	(0.1137)	(0.0481)	(0.0947)	(0 0952)

1973:1–1986:2

Mean	0.9900	0.9900	0.9900	0.9900
	(0.0021)	(0.0067)	(0.0150)	(0.0327)
Standard deviation	0.0139	0.0375	0 0797	0.1716
	(0.0016)	(0.0047)	(0.0106)	(0.0261)
Autocorrelation	0.00	0.11	0.16	0 17
	(0.0723)	(0.0544)	(0.0560)	(0.0561)

[a] Numbers in parentheses are Newey–West (1987) standard errors. Consumption, price level, and population data are from the 1987 CITIBASE tape, series GC82, GCS82, GCN82, GCD82, GDC, GDCS, GDCN, GDCD, and POPRES. Nominal marginal rates of substitution are computed from GC82 and GDC.

Table 5

How the standard deviation of the marginal rate of substitution varies with risk aversion.

Variable	Value for selected levels of risk aversion (α)				
	1	2	5	10	40
μ_1	0.9802	0.9703	0.9408	0.8918	0.6155
μ_2	0.9998	1.0097	1.0392	1.0882	1.3645
b	0.0099	0.0197	0.0492	0.0982	0.3745

idea of what bond prices would be like in a noninflationary environment. For illustration, we use the two-state distribution, $x = 1.005 \pm 0.010$, which has a slightly larger standard deviation than we observe in U.S. nondurables consumption.

We noted that the risk premium is approximately $-\theta(1 - \theta^2)b^2$. The first two terms have a maximum (in absolute value) of 0.385 at $|\theta| = 1/\sqrt{3} \approx 0.58$. For the model to generate a risk premium of 0.0012, we need a standard deviation, b, of the marginal rate of substitution of at least 0.056. If there is no inflation, the observed variance of x tells us what α must be for the model to match the data. As we vary α from 1 to 40, choosing β so that the mean of n in each case is 0.99, we change b, the standard deviation of n, as shown in table 5. To get risk premiums as large as those observed in the U.S. Treasury bill market, α must be at least 7 or 8 and larger for different values of θ. Inflation modifies the calculations slightly, and in principle, volatility of the price level could account for some of the required volatility of n. The last section of table 4 makes clear, however, that a large value of α is still necessary to generate a sufficiently large variance in the marginal rate of substitution. Nominal marginal rates of substitution are not much more variable than real ones in postwar U.S. data.

We conclude that the model is inconsistent with consumption data for modest values of the risk aversion parameter. Mehra and Prescott (1985), who studied the risk premium on equity, interpret a similar result as evidence against the model. Although this issue has been debated, most econometric evidence points to smaller values of α than are required to reconcile the model with the data. Hansen and Singleton (1982, 1983), for example, using a theoretical framework similar to ours, estimate that α lies between 1 and 2.

A second anomaly concerns the persistence parameter, θ. We noted that the model generates risk premiums and autocorrelated prices only if there is autocorrelation in marginal rates of substitution. Further, the sign of θ determines the sign of the risk premium, so that a positive risk premium can only be generated in the artificial economy when θ is negative – that is, when marginal rates of substitution are negatively autocorrelated. In our economy, marginal rates of substitution are simple functions of consumption growth

rates. The evidence for such autocorrelation in quarterly U.S. consumption growth rates is just not there. The estimates in table 4 are typically positive and not significantly different from zero. This property is not altered by accounting for the nonlinearity introduced by risk aversion, the additional randomness due to inflation, or even the particular sample period chosen.

These two discrepancies between theory and data appear to be robust to both small modifications of the theory and reinterpretations of the evidence. With regard to risk aversion, Epstein and Zin (1989a) suggest that an alternative class of preferences with risk aversion and intertemporal substitution (both determined by α in our economy) independent may generate more plausible asset prices with modest degrees of risk aversion. Subsequent empirical work by Hansen and Jagannathan (1988) and Epstein and Zin (1989b) indicates, however, that this extension of the class of preferences is not sufficient to account for the dynamic behavior of bond and equity prices.

With regard to serial correlation in marginal rates of substitution, two considerations arise. One is time aggregation. The consumption and price level data used in table 4 are averages over a quarter, while the Treasury bill prices used in tables 1 and 2 are measured at points in time. In the theory, both are measured at the same point in time, but since point-sampled consumption is not available, we used time-averaged quarterly data. It is difficult to think of a situation in which time aggregation is innocuous with respect to the serial correlation properties of the data, and Ermini (1988) and Heaton (1988) show that at monthly intervals consumption growth rates are negatively autocorrelated. This might imply negative autocorrelation in quarterly consumption growth rates for point-sampled data, but a more plausible explanation is measurement error in the levels of the monthly series.

A different direction of attack on the lack of autocorrelation in the data is to posit modifications to the economy, like durable goods or habit formation, that generate serially correlated marginal rates of substitution from serially uncorrelated growth rates of expenditures. Dunn and Singleton (1986) estimate a consumption-based model for the term structure with durable goods, but find that this produces positive, not negative, autocorrelation of marginal rates of substitution. Constantinides' (1988) model with habit persistence, however, can produce negative autocorrelation in marginal rates of substitution even when growth rates of consumer expenditures are nearly independent.

5. Risk premiums in artificial data

In section 4, we examined the model's ability to match some of the sample moments of Treasury bill data, an exercise in the spirit of Mehra and Prescott's (1985) study of the equity premium. Here we return to the issue noted at the start of our study: Can the model account for rejections of the expectations hypothesis with U.S. Treasury bill data? This requires not only a

Table 6

Test results from experiment 1: A symmetric equilibrium distribution (number of rejections by 5 percent tests in 1,000 replications).[a]

$\theta = -0.3$ $\pi = (0.5, 0.5)$ $n = (0.9063, 1\ 0737)$

| Variable or test | Variable or number of rejections in each regression | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variable	None		$f_{1t} - q_{1t}$		$q_{1t} - f_{1t-1}$	
AR1	54	56	44	44	22	23
AR4	49	46	40	40	42	42
ARCH1	50	49	53	53	53	52
ARCH4	42	42	42	42	37	37
Wald ($a = b = 0$)	—	—	179	58	177	59
Wald ($b = 0$)	—	—	56	55	57	55
Wald ($b = -1$)	—	—	1,000	1,000	—	—

[a] The odd-numbered equations have $q_{1t+1} - f_{1t}$ as their dependent variable; the even ones, the same minus the forward risk premium, fp_{1t}. The tests are explained in the notes to table 2.

nonzero risk premium, but one that varies over time, so the issue is whether this variation is consistent with the theory and the parameter values suggested in the last section. We also examine the model's ability to reproduce the autoregressive conditional heteroskedasticity (ARCH) found in regressions with historical data.

We begin by reestimating the Treasury bill regressions of table 2 with data from our artificial economy. Since the forward-price and holding-period return regressions are almost identical, we report only the former. In each experiment we report summary statistics from 1,000 replications of 200 observations each. The initial state is selected at random (using NAG pseudorandom number generator G05CAF) from the equilibrium distribution, and succeeding states are drawn from the distribution described by the appropriate row of the transition matrix. We use the simple persistence parameterization, eq. (13), unless noted otherwise. Values for nominal marginal rates of substitution in each state are entered directly.

In *experiment 1*, reported in table 6, the mean and standard deviation of n are taken from table 4 with $\alpha = 10$. We set $\theta = -0.3$, which provides the required negative serial correlation, and set $\pi_1 = \pi_2 = 0.5$ (a symmetric equilibrium distribution). The nonzero value of θ has little empirical support, but without it the model is doomed to failure. We view the exercise as telling us whether modifications of the economy designed to induce negative serial correlation in marginal rates of substitution are likely to successfully account for variations in risk premiums as well as nonzero means. The mean risk premium in this economy is 0.00193, which is about 30 percent larger than that we saw in the Treasury bill data.

Table 7

Sample regressions from experiment 1: A symmetric equilibrium distribution ($y = a + bx$).[a]

$\theta = -0.3$ $\pi = (0.5, 0.5)$ $n = (0.9063, 1.0730)$

Variable or test	Variable or estimated value in each regression					
	(1)	(2)	(3)	(4)	(5)	(6)
x	None		$f_{1t} - q_{1t}$		$q_{1t} - f_{1t-1}$	
a	0 0029	0.0010	0.0029	0.0009	0 0029	0.0009
	(0 0017)	(0.0017)	(0.0017)	(0.0017)	(0 0017)	(0.0017)
b	—	—	−0.013	−0.015	0.014	0.016
	—	—	(0.053)	(0.052)	(0.071)	(0.071)
s	0.0242	0 0242	0.0242	0.0242	0 0242	0.0242
DW	1 98	1.97	2.01	2.01	2.01	2.01
$AR1$	0.934	0.913	0.601	0.601	0.198	0.196
$AR4$	0 904	0.901	0.818	0.818	0.597	0.595
$ARCH1$	0 555	0.556	0.543	0.543	0 542	0.542
$ARCH4$	0 394	0.395	0.386	0 386	0.399	0 401
Wald ($a = b = 0$)	—	—	0.227	0.818	0.230	0 829
Wald ($b = 0$)	—	—	0.801	0.781	0 840	0.823
Wald ($b = -1$)	—	—	10^{-10}	10^{-10}	—	—

[a] See table 6.

In fact, the model economy with these parameter values does not have nearly enough variability in the risk premium to account for rejections of the expectations hypothesis. There is ample evidence of nonzero risk premiums, but Wald tests of the expectations hypothesis $b = 0$ and LM tests for autoregressive and ARCH errors have little power. The numbers of rejections for 5 percent tests are not much different from 50, which is what we expect from chance alone. The slight differences may, in fact, be due to sampling variability in our Monte Carlo experiment: a 95 percent confidence interval around 50 is approximately [36, 64]. The hypothesis $b = -1$, that the forward premium has predictive power for changes in spot prices, is rejected overwhelmingly in all 1,000 replications, a feature that recurs in all our experiments. Table 7 describes a single draw from the experiment in the same format as table 2.

The assumption in experiment 1 of a symmetric equilibrium distribution turns out to be extremely restrictive. With this specification, we virtually guarantee that the risk premium is constant and that forecast errors are homoskedastic. In the notation of section 4, the risk premium on forward contracts is $-\theta(1 - \theta^2)b^2/q_1$, where b is the standard deviation of the marginal rate of substitution. Unless the one-period bond price has a huge variance, which it does not, the variance of the risk premium is too small to cause difficulties for the expectations hypothesis. In experiment 1, its standard

Table 8

Test results from experiment 2: An asymmetric equilibrium distribution (number of rejections by 5 percent tests in 1,000 replications).[a]

$\theta = -0.3$ $\pi = (0.6, 0.4)$ $n = (0.9216, 1.0925)$

Variable or test	Variable or number of rejections in each regression					
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variable	None		$f_{1t} - q_{1t}$		$q_{1t} - f_{1t-1}$	
AR1	40	45	54	54	43	43
AR4	48	49	53	53	46	45
ARCH1	478	495	479	479	474	476
ARCH4	299	320	280	280	273	278
Wald ($a = b = 0$)	—	—	195	66	197	64
Wald ($b = 0$)	—	—	49	50	50	58
Wald ($b = -1$)	—	—	1,000	1,000	—	—

[a]See table 6.

deviation is 0.00005, which is 2 percent of its mean and just 3 percent of the estimated lower bound of the standard deviation in U.S. Treasury bill data. The symmetric distribution also rules out ARCH errors. The conditional variance of the one-period forecast error, $q_{1t+1} - E_t q_{1t+1}$, depends in general on the state, but when the equilibrium distribution is symmetric, it is constant at $\theta^2(1 - \theta)^2 b^2$. Similar propositions hold approximately for holding premiums.

The clear implication is that we cannot generate realistic data with a symmetric equilibrium distribution in the two-state version of the economy. Since this choice is arbitrary, we experiment with alternatives. In the first alternative, labeled *experiment 2*, we change the equilibrium distribution slightly to $(\pi_1, \pi_2) = (0.6, 0.4)$, with the same mean and standard deviation for n as experiment 1. The mean risk premium is still 0.00193, but its standard deviation rises by a factor of six to 0.00029. The results are reported in tables 8 and 9, and the outcome is somewhat closer to what we see in Treasury bill data. Regressions with the risk premium (the odd-numbered regressions) reject the joint hypothesis $a = b = 0$ more often than those without do. Tests for $b = 0$, however, reject about 5 percent of the time whether or not the risk premium is included, which is consistent with a time-invariant risk premium. We also see very little evidence of serial correlation. The test for first-order ARCH, however, rejects close to 50 percent of the time whether or not the risk premium is included. The same features are apparent in sample regressions from the experiment, shown in table 9.

These two experiments together suggest that risk premiums in the artificial economy do not vary enough or in the right manner to explain econometric rejections of the expectations hypothesis. Our two experiments have this

Table 9

Sample regressions from experiment 2: An asymmetric equilibrium distribution ($y = a + bx$).[a]

$\theta = -0.3$ $\pi = (0.6, 0.4)$ $n = (0.8875, 1.1438)$

Variable or test	Variable or estimated value in each regression					
	(1)	(2)	(3)	(4)	(5)	(6)
x	None		$f_{1t} - q_{1t}$		$q_{1t} - f_{1t-1}$	
a	0.0050	0.0030	0.0050	0.0031	0.0051	0.0032
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
b	—	—	0.016	0.024	−0.026	−0.038
	—	—	(0.046)	(0.046)	(0.064)	(0.064)
s	0.0232	0.0232	0.0233	0.0233	0.0232	0.0233
DW	2.06	2.08	2.02	2.02	2.01	2.01
$AR1$	0.652	0.539	0.616	0.616	0.268	0.270
$AR4$	0.972	0.957	0.969	0.969	0.839	0.840
$ARCH1$	0.0001	0.00004	0.0002	0.0002	0.0003	0.0002
$ARCH4$	0.0002	0.0001	0.0004	0.0004	0.0004	0.0004
Wald ($a = b = 0$)	—	—	0.004	0.089	0.004	0.086
Wald ($b = 0$)	—	—	0.791	0.596	0.732	0.552
Wald ($b = -1$)	—	—	10^{-10}	10^{-10}	—	—

[a]See table 6.

property, and reasonable changes in parameter values are unable to reverse it. More highly skewed equilibrium distributions, for example, increase the variability of the risk premium, but not enough to produce frequent rejections of $b = 0$. Alternatively, if we increase α, lowering (the absolute value of) θ at the same time to maintain the mean value of the risk premium, or the converse, the results change very little from those of table 8. We need a large value of α to produce large risk premiums and a large value of θ to make them predictable, as the regression results suggest they are. If we do both, raising α to 20 and θ to −0.5, we can reject the expectations hypothesis easily, but with a mean risk premium of 0.01080, almost eight times larger than we see in U.S. Treasury bill data.

Another example of an extreme experiment, which may be loosely associated with the peso problem, is described in tables 10 and 11. We call this *experiment 3*. In it, we have specified a three-state economy in which two similar states alternate with a third. There is strong positive persistence between the two similar states, as a group, and the unusual third state, but negative persistence between the two similar states themselves. We think of this experiment as illustrating the possibility of a change in regime – from, say, low to high inflation. The change has a small probability conditional on

Table 10

Test results from experiment 3: Unusual states (number of rejections by 5 percent tests in 1,000 replications).[a]

$n = (0.5, 1.3, 1.4)$

Variable or test	Variable or number of rejections in each regression					
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variable	None		$f_{1t} - q_{1t}$		$q_{1t} - f_{1t-1}$	
ARI	447	60	34	39	247	65
AR4	447	59	53	50	229	69
ARCH1	70	56	62	62	72	52
ARCH4	71	56	62	62	72	52
Wald ($a = b = 0$)	—	—	801	77	851	183
Wald ($b = 0$)	—	—	519	47	426	165
Wald ($b = -1$)	—	—	1,000	1,000	—	—

[a] The transition matrix is

$$\begin{bmatrix} 0.90 & 0.05 & 0.05 \\ 0.05 & 0.25 & 0.70 \\ 0.05 & 0.70 & 0.25 \end{bmatrix}$$

with equilibrium distribution $= [0.33, 0.33, 0.33]$. Also, see table 6.

Table 11

Sample regressions from experiment 3: Unusual states ($y = a + bx$).[a]

$n = (0.5, 1.3, 1.4)$

Variable or test	Variable or estimated value in each regression					
	(1)	(2)	(3)	(4)	(5)	(6)
x	None		$f_{1t} - q_{1t}$		$q_{1t} - f_{1t-1}$	
a	-0.071	-0.015	-0.062	-0.031	-0.062	0.014
	(0.016)	(0.016)	(0.019)	(0.019)	(0.0159)	(0.016)
b	—	—	-0.138	0.230	0.134	0.026
	—	—	(0.167)	(0.168)	(0.030)	(0.032)
s	0.2245	0.2249	0.2246	0.2243	0.2231	0.2253
DW	1.73	1.90	1.80	1.81	1.98	1.95
ARI	0.066	0.492	0.077	0.086	0.600	0.040
AR4	0.354	0.226	0.432	0.393	0.892	0.225
ARCH1	0.104	0.101	0.102	0.104	0.064	0.092
ARCH4	0.551	0.544	0.549	0.548	0.429	0.520
Wald ($a = b = 0$)	—	—	0.00003	0.236	10^{-9}	0.447
Wald ($b = 0$)	—	—	0.411	0.170	10^{-6}	0.427
Wald ($b = -1$)	—	—	10^{-6}	10^{-7}	—	—

[a] See tables 6 and 10.

Table 12

Means and standard deviations of forward risk premiums.

Economy	Mean	Standard deviation
United States (estimated)	0.0014	0.0016[a]
Experiment 1 (symmetric equilibrium distributions)	0.0019	0.00005
Experiment 2 (asymmetric equilibrium distributions)	0.0019	0.0003
Experiment 3 (unusual states)	−0.0459	0.0342
Experiment S1 (extreme values)	0.0108	0.0041
Experiment S2 [Rietz (1988)]	0.0141	0.0008

[a] This is the estimated lower bound.

Table 13

Test results from experiment S1: Extreme values for α and θ (number of rejections by 5 percent tests in 1,000 replications).[a]

$$\theta = -0.5 \qquad \pi = (0.6, 0.4) \qquad n = (0.8504, 1.2002)$$

Variable or test	Variable or number of rejections in each regression					
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variable	None		$f_{1t} - q_{1t}$		$q_{1t} - f_{1t-1}$	
AR1	66	47	64	64	39	36
AR4	77	56	65	65	56	54
ARCH1	167	303	310	310	236	228
ARCH4	89	162	188	188	141	163
Wald ($a = b = 0$)	—	—	457	100	457	73
Wald ($b = 0$)	—	—	103	59	80	56
Wald ($b = -1$)	—	—	1,000	1,000	—	—

[a] See table 6.

being in the other regime, but is very persistent once the change occurs. In this case, there is strong evidence of serial correlation, and the coefficient restrictions are frequently rejected. Both are caused by the risk premium: the even-numbered regressions reject roughly 5 percent of the time.

The failure of experiments 1 and 2, and modifications of them, to account for rejections of the expectations hypothesis can be traced directly to lack of variability in the risk premium. In table 12, we report the mean and standard deviation of the forward risk premium for each of the three experiments and for two supplementary experiments. The first supplementary experiment, presented in table 13, is the one just described, with extreme values of α and θ. The other, presented in table 14, is the three-state 'crash' economy of Rietz (1988), who proposed it as a solution to the equity premium puzzle. Only the experiment with extreme parameter values has as much variation in the risk

Table 14

Test results from experiment S2: The Rietz economy (number of rejections by 5 percent tests in 1,000 replications).[a]

$\theta = -0.5 \qquad n = (1.1091, 0.7012, 116.18)$

Variable or test	Number of rejections in each regression					
	(1)	(2)	(3)	(4)	(5)	(6)
Independent variable	None		$f_{1t} - q_{1t}$		$q_{1t} - f_{1t-1}$	
AR1	79	51	55	58	23	24
AR4	72	61	52	53	52	51
ARCH1	88	74	74	71	76	71
ARCH4	55	50	48	46	48	44
Wald ($a = b = 0$)	—	—	1,000	63	1,000	57
Wald ($b = 0$)	—	—	85	54	84	54
Wald ($b = -1$)	—	—	1,000	1,000	—	—

[a] See table 6

premium as we estimate in the data, and that has a mean risk premium almost eight times as large as that we estimate. Even the Rietz experiment, with its extreme state, has less than half the estimated variability of the forward risk premium and ten times the mean.

6. Final remarks

We have examined the term structure of interest rates from the perspective of general equilibrium asset pricing and found that the representative agent model with additively separable preferences fails to account for the sign or the magnitude of risk premiums on forward contracts or holding-period returns on multiperiod bonds. This failure is robust to small modifications of the theory and to reinterpretations of the data suggested by sampling variability and aggregation over time. These discrepancies between theory and data have the same flavor as work by Mehra and Prescott (1985) on the average excess return of equity over bonds. Both emphasize differences in the mean values of risk premiums between observed asset prices and prices derived from a widely used theoretical economy.

We also show that the theoretical economy cannot account for the variability of risk premiums. It fails, in particular, to account for the time variation in risk premiums implied by rejections of the expectations hypothesis with U.S. Treasury bill data: regressions with artificial data, produced by simulating our theoretical economy, invariably accept the expectations hypothesis. This happens even when we relax the strict mapping between the data and the theoretical economy and use parameter values that are inconsistent with other

evidence. A large body of empirical work suggests that risk premiums on a variety of assets vary in systematic ways over time. Our work contributes to the increasingly strong evidence that such variation is not explicable in this class of theoretical models. Both sets of results, on the mean and the variability of risk premiums, add to the growing list of empirical deficiencies of the representative agent model of asset pricing.

Finally, we think the artificial economy approach, which has the same flavor as Tauchen's (1986) Monte Carlo study, is a useful one for bringing theory and econometric work together. It highlights the empirical content of the theory and provides a useful laboratory for examining econometric estimators and test statistics.

References

Backus, D.K. and P.J Kehoe, 1987, Trade and exchange-rate dynamics in a dynamic competitive economy, Research Department working paper 348 (Federal Reserve Bank of Minneapolis, Minneapolis, MN)

Barton, E., F. David, and E. Fix, 1962, Persistence in a chain of multiple events when there is simple persistence, Biometrika 49, 351–357.

Breeden, D.T., 1979, An intertemporal asset pricing model with stochastic consumption and investment opportunities, Journal of Financial Economics 7, 265–296.

Brock, W.A., 1982, Asset prices in a production economy, in: J.J. McCall, ed., The economics of information and uncertainty (University of Chicago Press, Chicago, IL) 1–43.

Brown, S J. and P.H. Dybvig, 1986, The empirical implications of the Cox, Ingersoll, Ross theory of the term structure of interest rates, Journal of Finance 41, 617–630.

Constantinides, G.M., 1988, Habit formation: A resolution of the equity premium puzzle, Working paper (University of Chicago, Chicago, IL).

Cox, J C., J.E Ingersoll, Jr., and S.A. Ross, 1985, A theory of the term structure of interest rates, Econometrica 53, 385–407.

Dunn, K.B. and K.J. Singleton, 1986, Modeling the term structure of interest rates under non-separable utility and durability of goods, Journal of Financial Economics 17, 27–55

Engle, R.F , 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987–1007.

Engel, R.F., D.M. Lilien, and R.P. Robins, 1987, Estimating time varying risk premia in the term structure: The ARCH-M model, Econometrica 55, 391–407.

Epstein, L. and S. Zin, 1989a, Substitution, risk aversion and the temporal behavior of consumption and asset returns: A theoretical framework, Working paper (University of Toronto, Toronto, Ont); forthcoming in Econometrica.

Epstein, L. and S Zin, 1989b, Substitution, risk aversion and the temporal behavior of consumption and asset returns: An empirical analysis, Working paper (University of Toronto, Toronto, Ont.).

Ermini, L., 1988, Some new evidence on the timing of consumption decisions and on their generating process, Working paper (University of Sydney, Sydney, Australia); forthcoming in Review of Economics and Statistics.

Fama, E.F., 1976, Forward rates as predictors of future spot rates, Journal of Financial Economics 3, 361–377.

Fama, E F., 1984, The information in the term structure, Journal of Financial Economics 13, 509–528.

Feller, W., 1968, An introduction to probability theory and its applications, Vol. 1, 3rd ed. (Wiley, New York, NY).

Gibbons, M. and K. Ramaswamy, 1986, The term structure of interest rates: Empirical evidence, Working paper (Stanford University, Stanford, CA).

Grossman, S.J., A. Melino, and R.J. Shiller, 1987, Estimating the continuous-time consumption-based asset-pricing model, Journal of Business and Economic Statistics 5, 315–327.

Hansen, L P and R Jagannathan, 1988, Using asset market data to restrict the volatility of intertemporal marginal rates of substitution, Working paper (University of Chicago, Chicago, IL).

Hansen, L P and K.J Singleton, 1982, Generalized instrumental variables estimation of nonlinear rational expectations models, Econometrica 50, 1269–1286.

Hansen, L.P and K.J Singleton, 1983, Stochastic consumption, risk aversion, and the temporal behavior of asset returns, Journal of Political Economy 91, 249–265.

Heaton, J , 1988, The interaction between time-nonseparable preferences and time aggregation, Working paper (University of Chicago, Chicago, IL).

Isaacson, D. and R. Madsen, 1976, Markov chains: Theory and applications (Wiley, New York, NY).

Lucas, R.E , Jr., 1978, Asset prices in an exchange economy, Econometrica 46, 1429–1445.

Lucas, R E., Jr., 1980, Equilibrium in a pure currency economy, in: J.H. Kareken and N. Wallace, eds., Models of monetary economies (Federal Reserve Bank of Minneapolis, Minneapolis, MN) 131–145

Merton, R.C , 1973, An intertemporal capital asset pricing model, Econometrica 41, 867–887.

Mehra, R. and E C. Prescott, 1985, The equity premium: A puzzle, Journal of Monetary Economics 15, 145–161

Melino, A , 1987, The term structure of interest rates: Evidence and theory, Working paper (University of Toronto, Toronto, Ont.); forthcoming in Journal of Economic Surveys.

Newey, W.K and K.D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, Econometrica 55, 703–708.

Rietz, T., 1988, The equity risk premium: A solution, Journal of Monetary Economics 22, 117–131

Roll, R , 1970, The behavior of interest rates: An application of the efficient market model to U.S. Treasury bills (Basic Books, New York, NY).

Sargent, T.J., 1987, Dynamic macroeconomic theory (Harvard University Press, Cambridge, MA).

Shiller, R.J. and J H McCulloch, 1987, The term structure of interest rates, Working paper 2341 (National Bureau of Economic Research, Washington, DC).

Shiller, R J., J.Y. Campbell, and K L. Schoenholtz, 1983, Forward rates and future policy: Interpreting the term structure of interest rates, Brookings Papers on Economic Activity 1, 173–217.

Stambaugh, R., 1986, The information in forward rates: Implications for models of the term structure, Working paper (University of Chicago, Chicago, IL).

Startz, R , 1982, Do forecast errors or term premia really make the difference between long and short rates?, Journal of Financial Economics 10, 323–329

Tauchen, G , 1986, Statistical properties of generalized method-of-moments estimators of structural parameters obtained from financial market data, Journal of Business and Economic Statistics 4, 397–416.

White, H , 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, Econometrica 48, 817–838.