Nonlinearities in the Relation Between the Equity Risk Premium and the Term Structure

Jacob Boudoukh • Matthew Richardson • Robert F. Whitelaw
Stern School of Business, New York University, 44 W. 4th Street, New York, New York 10012

This paper investigates the relation between the conditional expected equity risk premium and the slope of the term structure of interest rates. Theoretically, these variables are linked, the relation may be nonlinear, and negative risk premiums are consistent with equilibrium. Given these implications, we employ a nonparametric estimation technique to document the empirical relation between the risk premium and the slope of the term structure using almost two hundred years of data. Of particular interest, the risk premium is increasing in the term structure slope; however, for either small or negative slopes, the risk premium is much more sensitive to changes in interest rates. In addition, the empirical results imply negative expected equity risk premiums for some inverted term structures. Finally, variations in the risk premium do not appear to be related to variations in the variance of equity returns. We illustrate these features in a stylized consumption-based model, and provide the economic intuition behind the results.
(*Equity Risk Premium; Predictability; Nonlinearity; Nonparametric*)

1. Introduction

Over the past 10 years, there has been a surge in research documenting time-variation in expected returns using a variety of instrumental variables. Yields on government bonds are natural instruments as they should theoretically depend on many of the unidentified state variables that drive expected stock returns. Thus, future expected returns should be linked to information contained in the term structure of interest rates.[1] The empirical evidence indicates that expected stock returns tend to be positively related to interest rate maturity spreads (e.g., Campbell 1987, Fama and French 1989, and Chen 1991). Furthermore, upward versus downward sloping term structures are important for determining this relation (e.g., Fama 1986, and Harvey 1988). Finally, most of the evidence covers post World War II data, which includes very few inverted term structures. The conclusion from this literature is that there is a significant relation between stock return premiums and term structure variables. In this paper, we investigate further the relation between the conditional expected risk premium and the slope of the term structure of interest rates.

We provide several contributions to the existing literature. First, we employ data over the past two centuries on both stock returns and short- and long-term interest rates. This leads to enough *independent* observations on upward sloping and inverted term structures to establish a link with the expected risk premium. Moreover, using a new sample, we avoid some of the "data snooping" biases associated with

[1] For earlier examples of research on time-varying expected returns, see Hansen and Singleton (1983), Gibbons and Ferson (1985), and Keim and Stambaugh (1986). For extensive reviews of the literature on expected returns and the term structure, see survey articles by Fama (1991) and Hawawini and Keim (1995).

0025-1909/97/4303/0371$05.00

repeated attempts to find statistically significant relations within a single dataset.[2] Second, given that theory does not dictate a linear relation between the risk premium and the term structure, we provide tests of the nonlinearity in this relation.[3] We find that, when the slope of the term structure is either small or negative, the risk premium is much more sensitive to changes in the spread between long- and short-term interest rates. This evidence is confirmed using both nonparametric and parametric methods. Third, while existing evidence suggests that the magnitude of the expected risk premium depends upon the slope of the term structure of interest rates, we report evidence that the sign of the risk premium does also. In particular, when the term structure is downward sloping, the expected risk premium can be negative. Fourth, the magnitude of the risk premium does not appear to be related to the variance of the risk premium. This empirical observation is inconsistent with the standard dynamic capital asset pricing model which implies a positive relation between expected return and variance. However, it is consistent with a consumption-based asset pricing model in which these two moments are not so closely tied. In fact, all of the empirical observations above are illustrated in a stylized model that provides the economic intuition behind the results.

The paper is organized as follows. In §2, we review the existing evidence and briefly illustrate the theoretical linkage between the risk premium and the term structure in a general setting. Section 3 provides a description of the data and some preliminary analysis. In §4, we test for nonlinearity in a parametric setting. Section 5 employs a nonparametric analysis to better describe the empirical relation between the risk premium and the term structure spread. Section 6 provides the economic intuition behind the empirical results in the context of a stylized numerical example. In §7, we make some concluding remarks.

[2] For example, Foster and Smith (1997) argue that there is a tendency to substantially overfit time-series models and provide some examples from the current empirical finance literature.

[3] Nonlinear models have been used previously to predict the equity risk premium. See, for example, Kairys (1993), who uses changes in commercial paper rates to predict the sign of the risk premium.

2. Theory and Existing Evidence

There is substantial empirical evidence regarding the relation between stock returns and interest rates. In general, researchers have found a positive relation between the risk premium and the slope of the term structure of interest rates.[4] Fama and French (1989) argue that the risk premium moves countercyclically. That is, expected risk premiums during recessions are large relative to premiums during expansions. There is also evidence linking the term structure to the business cycle (e.g., Kessel 1965, Fama 1986, Harvey 1988 and Estrella and Hardouvelis 1991). For example, Harvey (1988) finds that the term structure is upward sloping during recessions (i.e., at the trough of the cycle), while inverted term structures generally occur towards the end of expansions (i.e., at the peak of the cycle). These characteristics of the data then produce a link between the equity risk premium and the term structure.

Outside of the intuition outlined above, there has been little theoretical work on the relation between the term structure and the expected risk premium.[5] Although it is difficult to relate these variables in closed-form, it is possible to make several observations in the context of a general model. In the absence of arbitrage, assets can be priced as the expected product of their payoff with a pricing operator $M_{t,t+1}$

$$Q_t = E_t[M_{t,t+1}(Q_{t+1} + D_{t+1})], \tag{1}$$

where Q is the nominal price of the asset, D is the nominal dividend, and $E_t[\cdot]$ is the expectation conditional on information at time t (see Harrison and Kreps 1979). In many contexts, this pricing operator can be thought of as the nominal marginal rate of substitution (MRS) between time t and time $t + 1$ of a representative agent, which is defined as a function of the ratio of the marginal utilities of consumption in the two periods (see §6

[4] Campbell (1987) uses the short-end of the yield curve to predict excess monthly returns on stocks and bonds over the 1959 to 1983 period. He finds that the spreads of the two-month and six-month bills over the one-month bill have predictive power. Using longer interest rate spreads, Fama and French (1989) and Chen (1991) find similar results in a post World War II sample.

[5] An exception is Campbell (1986) who relates bond prices to a claim on the market's dividend in a given period.

for an example). Equation (1) can also be rewritten in terms of asset returns

$$E_t[M_{t,t+1}R_{t,t+1}] = 1, \tag{2}$$

where $R_{t,t+1} = (Q_{t+1} + D_{t+1})/Q_t$.

Using Eq. (1), the price of a τ-period, riskless bond which pays \$1 in the future is

$$P_{t,\tau} = E_t[M_{t,t+\tau}] \tag{3}$$

where $M_{t,t+\tau}$ is the MRS between time t and time $t + \tau$, which is just the product of the single period marginal rates of substitution, i.e., $M_{t,t+\tau} = M_{t,t+1}M_{t+1,t+2} \cdots M_{t+\tau-1,t+\tau}$. The slope of the term structure (denoted $\Delta r_{lf,t}$) is simply the yield on a long-term bond (denoted R_{lt}) minus the yield on a 1-period bond (the risk-free rate, denoted R_{ft}), where

$$R_{lt} \equiv E_t[M_{t,t+\tau}]^{-1/\tau}, \tag{4}$$

$$R_{ft} \equiv E_t[M_{t,t+1}]^{-1}, \tag{5}$$

$$\Delta r_{lf,t} \equiv R_{lt} - R_{ft}, \tag{6}$$

and the long-term bond has τ periods to maturity.

Equations (2) and (5) imply the following expression for the expected risk premium (i.e., the expected return on the asset minus the risk-free rate)

$$E_t[R_{t,t+1} - R_{ft}] = -R_{ft}\,\mathrm{Cov}_t[R_{t,t+1}, M_{t,t+1}], \tag{7}$$

where $\mathrm{Cov}_t[\cdot,\cdot]$ is the conditional covariance at time t. This relation holds true for all assets, including the stock market, so the expected equity risk premium (denoted rp_{t+1}) can also be written as a function of the covariance of the market return with the MRS

$$\begin{aligned} E_t[rp_{t+1}] &= E_t[R_{mt+1} - R_{ft}] \\ &= -R_{ft}\,\mathrm{Cov}_t[R_{mt+1}, M_{t,t+1}] \end{aligned} \tag{8}$$

where R_{mt+1} is the 1-period return on the market. In addition, Eq. (1) also implies that the price of the market next period, which is one important component of the return, is itself the sum of future dividends discounted back at the appropriate marginal rates of substitution, i.e.,

$$Q_{mt+1} = E_{t+1}\left[\sum_{\iota=1}^{\infty} M_{t+1,t+1+\iota}D_{mt+1+\iota}\right], \tag{9}$$

where Q_m and D_m are the market price and dividend respectively.

Given this level of generality, there are four preliminary conclusions that can be drawn. First, the expected risk premium and the term structure are linked via their dependence on expectations of future marginal rates of substitution (see Eqs. (4)–(6), (8) and (9)). Note that this dependence coincides with explanations in the literature linking both variables to business cycle fluctuations through such factors as economic growth.

Second, the relation between the risk premium and the slope of the term structure is likely to be nonlinear given the nonlinear functions of the marginal rates of substitution in Eqs. (6) and (8).

Third, negative risk premiums are not precluded in this environment. Equation (8) implies that the expected return on equity will be less than the risk-free rate if, and only if, this return covaries positively with the MRS.

Fourth, while the expected risk premium will vary depending on the covariance of the market return with the MRS, this variation will not necessarily coincide with variation in the variance of the risk premium. In particular, the covariance in Eq. (8) will be proportional to the variance of equity returns only when the market return is perfectly correlated with the MRS.

3. Data and Preliminary Analysis

We use annual data on short- and long-term yields and stock returns over the past two centuries for the U.S. (1802–1990). The data are described in detail in Siegel (1992) and Schwert (1990), so we provide only a brief synopsis. The stock return series is constructed to most closely match a broad index. With respect to the yield data, there was an active market for long-term U.S. government bonds over most of the sample period. Although the maturities differ, Siegel (1992) chooses the bonds closest to twenty years. In the earlier period, Siegel (1992) must construct the U.S. risk-free rate using U.S. commercial paper rates, U.K. short-term rates (under the gold standard), and available U.S. government rates. He finds that his constructed series matches actual available rates during this period.[6]

[6] Throughout the paper we use simple, not continuously compounded, yields. The results are essentially invariant to this choice.

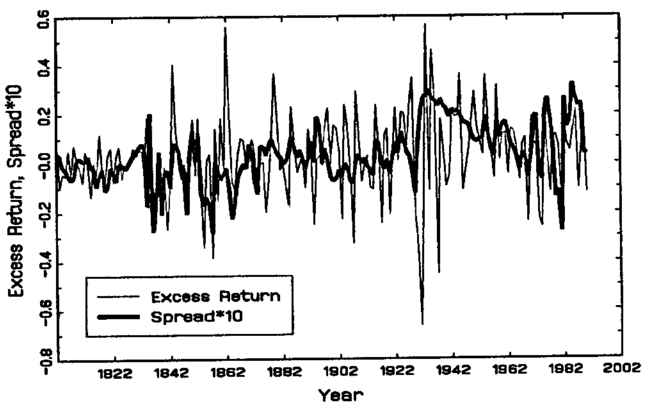
In spite of potential measurement errors in the dataset, using such a long sample has several advantages relative to the post World War II, monthly data used in previous empirical investigations. First, short sample periods provide only limited information due to (i) the high autocorrelation in monthly term structure spreads, and (ii) the limited number of business cycles contained in a small sample. Second, new data is less subject to data snooping concerns. Third, the robustness of the results over various subsamples can be studied.

As an introduction to the data, Figure 1 plots the excess stock market return and the yield spread (times ten) over the sample period 1802–1990. The risk premium is defined as the annual *ex post* stock market return in excess of the risk-free rate ($rp_{t+1} \equiv R_{mt+1} - R_{ft}$), and the spread ($\Delta r_{lf,t}$) is the long-term yield minus the short-term yield. While the excess return shows substantially more variation, it is apparent that downward (upward) sloping yield curves tend to be associated with low (high) returns.

Over the sample period 1802–1990 and the subperiods 1802–1896 and 1897–1990, Table 1 provides the correlation between the excess equity return and the term structure spread. The standard errors on the correlation estimates are adjusted for both autocorrelation and heteroskedasticity. The correlation between the excess return and the term structure spread in the full sample is 0.316, with a standard error of 0.0826; therefore, the estimate is 3.8 standard errors from zero. For the first subperiod the estimate is 0.380 (4.3 standard errors from zero), while for the second subperiod the correlation is 0.233 (2.4 standard errors from zero).[7]

Figure 1 **The Excess Stock Market Return and the Spread**

The Figure Plots the Annual Excess Stock Market Return and the Yield Spread Between Long-term and Short-term Bonds (Times 10) versus Time.



4. Testing for Nonlinearity

We first apply some common nonlinear methods to gauge both the statistical and economic significance of the nonlinearities in the relation between the risk premium and the spread. We use two descriptive methods for approximating the form of the relation: (i) a piecewise linear regression of the risk premium on the slope of the term structure, and (ii) a Taylor expansion of the risk premium as a function of the term structure spread. While the Taylor expansion method has some critics (e.g., Gallant 1982), it has a convenient economic interpretation in terms of the distribution of the underlying variables.

As a first approximation of nonlinearities in the data, we perform a piecewise linear regression of the risk premium on the predetermined spread between the long- and short-rate of interest. We use a spread of 0% (i.e., the point at which the term structure is flat) as a breakpoint. Specifically, we run the regression

$$rp_{t+1} = \alpha + \beta_1(\Delta r_{lf,t}) + \beta_2 \max[0, \Delta r_{lf,t}] + \epsilon_{t,t+1},$$

where $\epsilon_{t,t+1}$ is the disturbance term.

The results are presented in Table 2. The estimate of the coefficient β_2 is consistently negative for the overall sample and the subperiods. Moreover, in most of the regressions, the estimator is significantly different from zero. This suggests nonlinearity in the relation between the risk premium and the term structure. For illustration purposes, consider the regression for the full sample period. Up to the breakpoint (i.e., a spread of 0%), the slope of the regression of the risk premium on the term structure spread is approximately 7.32. Hence, for every 0.1% change in $\Delta r_{lf,t}$, there is a 0.7% change in $E_t[rp_{t+1}]$. In contrast, for positive spreads, where the relevant

[7] Although the correlation coefficients are similar in magnitude across subperiods, this is not true of the mean excess returns. Siegel (1992) finds that the average excess return is much lower in the nineteenth than the twentieth century. Siegel's explanation is based on there being low levels of interest rates in the latter period.

Table 1 **The Excess Stock Market Return and the Term Structure Spread**

Period	$E[rp_{t+1}]$	$E[\Delta r_{lf,t}]$	$sd[rp_{t+1}]$	$sd[\Delta r_{lf,t}]$	$corr(rp_{t+1}, \Delta r_{lf,t})$
1802–1990	0.0311	0.00356	0.167	0.0117	0.316
	(0.0111)	(0.00212)	(0.0203)	(0.00131)	(0.0826)
1802–1896	0.00745	−0.00135	0.135	0.00915	0.380
	(0.00901)	(0.00154)	(0.0195)	(0.00116)	(0.0888)
1897–1990	0.0550	0.00853	0.192	0.0119	0.233
	(0.0166)	(0.00281)	(0.0304)	(0.00145)	(0.0986)

Note: Table 1 provides summary statistics (means, standard deviations, and correlations) for the excess stock market return (rp_{t+1}) and the term structure spread ($\Delta r_{lf,t}$) using annual data on stock returns, one-year rates and long-rates (maturity close to 20 years). Standard errors, in parentheses, are heteroskedasticity and serial correlation adjusted using Hansen's (1982) generalized method of moments (GMM).

slope coefficient is $\beta_1 + \beta_2 = 7.3 - 4.4 = 2.9$, there is only a 0.3% change. This result suggests that the nonlinear relation is concave. That is, the expected risk premium is more sensitive to changes in the spread when the term structure spread is negative.

A second way of approximating a functional form is to apply a Taylor series expansion. Specifically, assuming $E_t[rp_{t+1}] = f(\Delta r_{lf,t}, \theta)$, where θ represents parameters describing the relation, the Taylor expansion of $f(\cdot)$ around zero leads to

$$rp_{t+1} = \alpha + \beta_1(\Delta r_{lf,t}) + \beta_2(\Delta r_{lf,t})^2 + \cdots + \epsilon_{t,t+1}, \quad (10)$$

where $\epsilon_{t,t+1}$ captures the unanticipated return and also the remaining terms of the expansion not used in estimation. While the Taylor expansion holds only locally, multicollinearity issues aside, it allows the researcher to interpret movements in the risk premium in terms of the moments of the distribution of the slope of the term structure.

Table 3 provides estimates of the relation for Taylor expansions of various orders. While the results are not as stable across subperiods as those described previously, it is possible to form several conclusions. First, the relation between the risk premium and the term structure is, for the most part, nonlinear. For example, in the 1802–1990 sample period, Table 3 shows that the second-order term, $(\Delta r_{lf,t})^2$, has additional explanatory power for movements in the expected risk premium. Moreover, the coefficients change very little as Taylor expansion terms are added.

Second, the second-order Taylor term always appears with a negative coefficient. For a given change in the spread, $\Delta r_{lf,t}$, there is a $\beta_1 + 2\beta_2(\Delta r_{lf,t})$ change in $E_t[rp_{t+1}]$. Consider the coefficients of the second-order Taylor expansion, $\hat{\beta}_1 = 5.26$ and $\hat{\beta}_2 = -90.74$ (adjusted for the fact that the spread is multiplied by 10 for the regressions in Table 3), and their implication for the

Table 2 **Piecewise Linear Estimation of the Risk Premium—Yield Spread Relation**

Period	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	R^2
1802–1990	0.033*	7.32*	−4.39	0.109
	(0.016)	(1.70)	(2.77)	
1802–1896	0.014	5.46*	0.41	0.144
	(0.024)	(2.12)	(4.67)	
1897–1990	0.056*	11.20*	−9.58*	0.076
	(0.025)	(4.65)	(5.53)	

Note: Table 2 estimates the relation between the risk premium and the term structure spread using a piecewise linear regression. Specifically, the coefficient estimates are from the regression

$$rp_{t+1} = \alpha + \beta_1(\Delta r_{lf,t}) + \beta_2 \max[0, \Delta r_{lf,t}] + \epsilon_{t,t+1},$$

where $\epsilon_{t,t+1}$ is the disturbance term, rp_{t+1} is the market risk premium and $\Delta r_{lf,t}$ is the spread between the long- and short-rate of interest. Standard errors, in parentheses, are heteroskedasticity and serial correlation adjusted using Hansen's (1982) GMM. Coefficients significant at the 10% level are marked with an asterisk. Note that the coefficient β_2 represents the difference in slopes around the breakpoint. Negative (positive) values suggest concavity (convexity). Annual data on stock returns, one-year rates and long-rates (maturity close to 20 years) are used.

Table 3 Taylor Series Expansion of the Risk Premium—Yield Spread Relation

Period	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	R^2	χ^2_{linear}
1802–1990	0.027*	0.587*	−1.294	−1.363	5.589	0.137	6.869
	(0.014)	(0.206)	(1.664)	(3.836)	(22.358)		(0.877)
1802–1990	0.026*	0.570*	−0.921*	−1.022		0.113	6.561
	(0.012)	(0.216)	(0.532)	(3.800)			(0.945)
1802–1990	0.027*	0.526*	−0.974*			0.113	5.183
	(0.011)	(0.091)	(0.428)				(0.968)
1802–1990	0.015*	0.452*				0.100	
	(0.008)	(0.103)					
1802–1896	0.019	0.419	−0.420	3.095		0.151	0.622
	(0.015)	(0.282)	(0.663)	(7.721)			(0.088)
1802–1896	0.021	0.496*	−0.759			0.150	0.512
	(0.018)	(0.166)	(1.061)				(0.157)
1802–1896	0.015	0.560*				0.144	
	(0.013)	(0.148)					
1897–1990	0.031	0.656*	−1.178	−1.805		0.072	3.94
	(0.020)	(0.347)	(0.718)	(5.023)			(0.727)
1897–1990	0.034*	0.588*	−1.370*			0.070	3.211
	(0.017)	(0.202)	(0.765)				(0.830)
1897–1990	0.023	0.376*				0.054	
	(0.016)	(0.132)					

Note: Table 3 estimates the relation between the risk premium and the term structure spread using a Taylor Series expansion. The estimates are from the regression:[1]

$$rp_{t+1} = \alpha + \sum_{i=1}^{k} \beta_i(\Delta r_{lt,t})^i + \epsilon_{t,t+1},$$

where $\epsilon_{t,t+1}$ captures the unexpected return and the remaining terms of the Taylor expansion, rp_{t+1} is the risk premium, and $\Delta r_{lt,t}$ is the spread between the long- and short-rate of interest. Standard errors, in parentheses, are heteroskedasticity and serial correlation adjusted using Hansen's (1982) GMM. Coefficients significant at the 10% level are marked with an asterisk. The statistic χ^2_{linear} provides a test of whether the relation is linear or not, i.e., $\beta_2 = \beta_3 = \cdots = 0$, with the corresponding *P*-value in parentheses. Annual data on stock returns, one-year rates and long-rates (maturity close to 20 years) are used.

[1] Note that the yield spread has been multiplied by 10 in the regression for ease in reading the table.

expected risk premium at spreads of −1%, 0% and 1%. At these term structure spreads, the change in the expected risk premium (for a given 0.1% change in the spread) will be 0.71%, 0.53% and 0.35% respectively. Thus, the estimated relation between the risk premium and the term structure spread is concave.

Finally, the coefficients on the higher order terms in the expansion are statistically insignificant, and these terms add little explanatory power. Thus, as an approximation, the relation can be described by a second-order expansion with the coefficient on the second term being negative. It should be noted that the point estimates of the other terms, though not significant, imply a slight flattening of the relation at extreme negative spreads.

While both the Taylor approximation and the piecewise linear regression provide statistically significant evidence of nonlinearity in the relation between the risk premium and the slope of the term structure, this evidence should not be interpreted as suggesting that either of these functional forms provides a precise model for predicting expected equity returns. In the following section we employ a nonparametric analysis to gain a better understanding of the true functional relation.

5. Nonparametric Analysis

5.1. Kernel Estimation

In this section, we employ a kernel estimation procedure for estimating the relation between the expected risk premium and the slope of the term structure.[8] Kernel estimation is a nonparametric method for estimating the joint density of a set of random variables. Specifically, given m-dimensional vectors $z_1, z_2, \ldots, z_T$ from an unknown density $f(z)$, then a kernel estimator of this density is

$$\hat{f}(z) = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{h^m} K\left(\frac{z - z_t}{h}\right), \tag{11}$$

where $K(\cdot)$ is a suitable kernel function and h is the window width or smoothing parameter. In practice, a multivariate normal probability density function is frequently used as the kernel function,[9] and the window width is chosen based on the dispersion of the observations. This fixed window width estimator is often called the Parzen estimator. The density at any point is estimated as the average of densities centered at the actual data points. The further a data point is away from the estimation point z, the less it contributes to the estimated density. Consequently, the estimated density is highest near high concentrations of data points and lowest when observations are sparse.

The asymptotic properties of these Parzen estimators, including consistency and rates of convergence, have been studied extensively.[10] Nevertheless, from an implementation standpoint, they are not totally satisfactory. In particular, in finite samples, the Parzen estimator responds poorly to variations in the true density. When the density is low and consequently observations are sparse, the estimated density tends to have spurious peaks at the actual data points. In contrast, when the density is high and consequently observations are dense, the estimated density tends to spread the influence of these data points over too large a region. A fixed window width does not permit both smoothing in the tails and high resolution at the peaks.

One solution to this problem is to consider a variable kernel estimator (VKE) in which the window width varies across the data points. In order to avoid the data snooping concerns that can arise when a window width is selected to achieve a desired degree of smoothing, we choose an objective criterion for the VKE. Specifically, Breiman, Meisel and Purcell (1977) propose replacing the fixed window width in Eq. (11) with

$$h_t = \alpha_k d_{t,k} \tag{12}$$

where $d_{t,k}$ is the distance from the point z_t to its kth nearest neighbor, and α_k is a constant multiplicative factor. The concept behind the VKE is that $d_{t,k}$ will be small when the data points are dense, and large when they are sparse, hence overcoming the limitations of the Parzen estimator. In fact, Breiman, Meisel and Purcell (1977) confirm that, in their simulations, the VKE significantly outperforms the Parzen estimator. Of equal importance, they develop a procedure for estimating the best values of k and α_k, and this procedure successfully locates the region of parameter values that gives the best fit to the actual density. Surprisingly, they find that locating these parameters is actually easier than locating the optimal fixed window width h in Eq. (11).

The suggested procedure, and the one employed in this analysis, is to first select an initial k equal to 10% of the sample size. A search is then conducted for the α_k which minimizes the goodness-of-fit measure

$$\hat{S} = \sum_{t=1}^{T} \left(\hat{w}_{(t)} - \frac{t}{T}\right)^2, \tag{13}$$

where $\hat{w}_{(t)}$ is the ordered permutation of the variables

$$\hat{w}_t = \exp[-T\hat{f}(z_t)V(d_{t,1})], \tag{14}$$

$V(r)$ is the volume of an m-dimensional sphere of radius r, and $d_{t,1}$ is the distance from point z_t to its nearest

[8] For examples of nonlinear methods for approximating functional forms in the empirical asset pricing literature, see, for example, Harvey (1991) and Pagan and Hong (1991) who apply the kernel method in order to capture the nonlinear relation between the risk premium and volatility and Boudoukh, Richardson, Stanton, and Whitelaw (1997) who apply it to the pricing and hedging of mortgage-backed securities.

[9] Epanechnikov (1969) demonstrates that many reasonable kernel functions give almost optimal results in terms of asymptotic rates of convergence, and Scott (1992) provides conditions under which the normal density function is optimal. The issue of the performance of different kernel functions is beyond the scope of this paper.

[10] For a detailed discussion of kernel estimation and its statistical properties, see Ullah (1988) and the references therein.

neighbor.[11] The minimizing value of α_k is then used to compute the ratio

$$\lambda = \frac{\alpha_k(\overline{d_k})^2}{\sigma(d_k)}, \qquad (15)$$

where $\overline{d_k}$ is the mean of the kth nearest neighbor distances, and $\sigma(d_k)$ is their standard deviation. The final step is to search for the value of k that minimizes $\hat{S}$ while selecting α_k so as to keep the ratio λ constant.

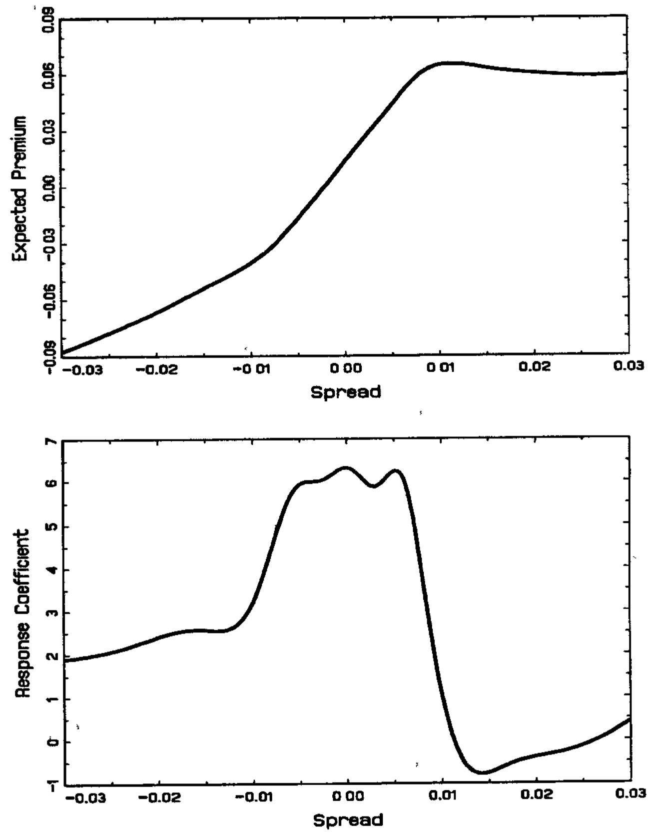
For our specific application, we are less interested in the joint density of the risk premium and the slope of the term structure, than in the *conditional* mean and variance of the risk premium, conditional on the term structure slope. However, these functionals can be computed relatively easily given the density (see Ullah 1988, for details). The obvious benefit of kernel estimation is that it is a nonparametric way to look at the relation between the expected risk premium and the slope of the term structure. The drawback is that its rate of convergence is relatively slow; therefore, kernel estimation can provide a noisy functional form in small samples.[12] Because our null does not specify a functional form, the researcher needs to be cautious when interpreting the results.

The Expected Risk Premium. Figure 2 graphs the kernel estimation of $E[rp_{t+1} \mid \Delta r_{lf,t}]$ versus $\Delta r_{lf,t}$ (top) and the derivative of the conditional mean with respect to the term structure slope (i.e., the response coefficient), which measures the sensitivity of the risk premium to changes in the term structure slope (bottom). For simplicity, a multivariate normal density function is used as the kernel function for estimation of the density. Several features of the estimated relation between the expected risk premium and the slope of the term structure are especially interesting. First, the expected risk premium is, for the most part, increasing in the term structure spread, i.e., the response coefficient is predominantly positive. There is some evidence that, for large positive spreads, this relation is reversed; however, the decline in the premium between spreads of 1% and 3% is small enough to be economically insignificant. For this range of spreads, the response coefficient is close to zero.

Second, the relation between the expected risk premium and the slope of the term structure is nonlinear. In particular, there appear to be three important components of this relation, and the response coefficient suggests an almost piecewise linear relation (with two breakpoints). At especially low spreads of less than −1%, an increase of 0.25% in the spread implies an increase in the risk premium of approximately 0.50% (a

Figure 2 **The Risk Premium and the Spread: Kernel Estimation**

The Top Figure Plots the Expected Risk Premium as a Function of the Spread, Estimated Using Variable Kernel Estimation. The Bottom Figure Plots the First Derivative of the Above Relation.



[11] The goodness-of-fit measure $\hat{S}$ is developed in Cover (1972) based on the fact that the variables w_t, evaluated at the true density, have a density that is approximately uniform.

[12] The rate of convergence is approximately $T^{-2/(p+4)}$, where T is the number of observations and p is the number of independent variables. Consequently, kernel estimation with one explanatory variable and 189 observations is somewhat comparable to OLS with 66 observations.

derivative of 2). Note that although these negative spreads are not common, over 10% of the sample still falls into this region. The majority of the sample (approximately 65%), however, covers term structure spreads between −1% and 1%. Over this range of values, the risk premium is more sensitive to movements in the slope of the term structure. For a 0.25% change in the spread, the risk premium is expected to change by approximately 1.5% (a derivative of 6). For large spreads (i.e., over 1%), and over approximately 25% of the sample, the expected risk premium is relatively constant, irrespective of changes in the term structure spread, and the response coefficient is small.

Could measurement errors be affecting the results? One issue is that measurement errors in the risk-free rate, which appears on both the left and right hand sides of the regression, may be driving the positive relation between the risk premium and the slope of the term structure. A priori, this is unlikely, and a replication of the analysis with total returns as the dependent variable yields similar results. A second issue is that measurement error in the spread may be introducing spurious nonlinearity. For example, if the variance of the spread is smaller for large absolute spreads (i.e., large negative or large positive spreads), then the effects of measurement error will be exacerbated in these regions. As a result, the coefficient will be downward biased in these regions, generating a pattern similar to the one observed in Figure 2. Of course, there is little reason to believe that the spread volatility takes this form, and, in fact, the reverse effect, wherein the variance of the spread is higher for higher absolute spreads, seems more likely.

Measurement error aside, the kernel estimation suggests a nonlinear relation between the risk premium and the term structure spread. However, it is difficult to gauge its statistical significance. Given the lack of a specific null hypothesis regarding the functional form of the relation between the risk premium and the term structure slope, the most appropriate statistical test of the kernel estimation results is a general test for the presence of nonlinearity. With this goal in mind, a nonparametric bootstrapping experiment is designed as follows:

1. excess returns are regressed on the slope of the term structure to obtain linear regression coefficients,
2. these coefficients and the term structure data are then used to construct 1000 artificial sequences of 189 excess returns by sampling with replacement from the regression errors,
3. the conditional mean of these returns is then estimated using kernel estimation,
4. the estimated mean is then projected linearly on the term structure spread data and the mean squared error (MSE) is calculated, and finally,
5. the MSEs from the 1000 replications are compared to the MSE from projecting the kernel estimation mean from the actual data on the slope of the term structure.

Note that the simulated returns are constructed to be a linear function of the slope of the term structure. Therefore, the linear projection of the kernel estimation mean should produce smaller MSEs in the simulated data than in the actual data, if, in fact, the actual data are generated by a nonlinear function. The MSE from the actual data falls at the 94.1 percentile of the simulated MSEs, giving strong support for a hypothesis of nonlinearity.

One caveat to the above bootstrapping exercise is that the resampling technique destroys any autocorrelation or heteroskedasticity patterns that may exist in the errors. In fact, there is evidence of a small, yet statistically significant, degree of autocorrelation and autoregressive heteroskedasticity in the residuals from the linear regression. One method for addressing this concern is to perform bootstrapping with errors that are constructed to have these properties. The drawback of this procedure is that it requires a detailed parametric specification for the structure of the residuals, which introduces potential misspecification problems. Nevertheless, we perform a simple parametric bootstrapping exercise that captures the salient features of the estimated residuals. Specifically, in step 2 of the procedure above, the sequences of artificial returns are generated using the model

$$rp_{t+1} = \beta_0 + \beta_1 \Delta r_{lf,t} + \epsilon_{t+1}, \tag{16}$$

$$\epsilon_{t+1} - \theta\epsilon_t = u_{t+1} \quad u_{t+1} \sim N(0, h_t^2), \tag{17}$$

$$h_t^2 = \gamma_0 + \gamma_1 u_t^2, \tag{18}$$

where the coefficients are estimated from the data. In this case, the MSE from the actual data falls at the 96.2 percentile of the simulated MSEs, which is even more supportive of nonlinearity than the nonparametric bootstrapping exercise above.

The final interesting feature of the kernel estimation is that the expected risk premium is often negative. Moreover, given the sample distribution of the term structure spread, these negative expected risk premiums are not that rare. For example, over one-third of the sample implies term structure spreads which are consistent with negative risk premiums. Note that the negative (positive) premiums are almost exclusively associated with downward (upward) sloping term structures.

The Conditional Variance of the Risk Premium. Figure 3 graphs the variance of the risk premium conditional on the term structure spread, i.e., $\sigma^2[rp_{t+1} \mid \Delta r_{lf,t}]$ versus $\Delta r_{lf,t}$. The results are in remarkable contrast to those presented above for the conditional expected risk premium. In particular, at large values of the spread (e.g., above 1%), the variance is increasing in the slope of the term structure. For example, the variance increases from 3.0% to 5.4% over the range of spreads above 1%. However, over this same range, the risk premium does not change. Apparently, there is little link between the expected risk premium and its variance when conditioning on the term structure.[13] As additional evidence of the behavior of the risk premium's variance, note that, when the risk premium is increasing in the spread (i.e., at $\Delta r_{lf,t} < 1\%$), the variance is, for the most part, constant around 2.3%. This, again, contradicts the well known intuition implied by the standard dynamic capital asset pricing model that the risk premium is increasing in its variance (e.g., see Merton 1980).

5.2. The Sign of the Expected Risk Premium

Results from the kernel estimation imply that the distribution of the risk premium varies with different shapes of the term structure of interest rates. To better understand the statistical significance of these results, consider conditioning the risk premium on downward versus upward sloping term structures. In a generalized method of moments (GMM) framework, the moment conditions are:

$$E\begin{pmatrix} (rp_{t+1} - \mu_{up}) \times I_{t,up} \\ (rp_{t+1} - \mu_{down}) \times I_{t,down} \\ \{(rp_{t+1} - \mu_{up})^2 - \sigma^2_{up}\} \times I_{t,up} \\ \{(rp_{t+1} - \mu_{down})^2 - \sigma^2_{down}\} \times I_{t,down} \end{pmatrix} = 0,$$

where $I_{t,up} = 1$ if $\Delta r_{lf,t} \geq 0$ and zero otherwise; $I_{t,down} = 1$ if $\Delta r_{lf,t} < 0$ and zero otherwise. The moments we estimate, μ_{up} and μ_{down}, correspond to the mean risk premiums conditional on an upward or downward sloping term structure respectively. Similarly, σ^2_{up} and σ^2_{down} are the variances of the risk premium in the two states.

These moment conditions allow us to test a variety of restrictions. First, is $\mu_{up} \geq \mu_{down}$?[14] Second, does the risk premium change sign depending on the slope of the term structure as the results in §5.1 suggest? Table 4 provides estimates of μ_{up}, μ_{down}, σ^2_{up}, and σ^2_{down}, and the corresponding test statistics. Note that the framework allows for both autocorrelation and heteroskedasticity in the risk premium when calculating the variance-covariance matrix of the estimates. Further, the cross-correlation between the estimates is taken into account in deriving the test statistics.

The average risk premium, over the entire sample period, conditional on the term structure being upward sloping, is 6.86%. In contrast, the average premium is −2.85% when the term structure is downward sloping. The difference in the expected risk premium for upward versus downward sloping yield curves is significant at conventional levels, i.e., $\mu_{up} \geq \mu_{down}$. Specifically, the test statistic, which is asymptotically normal, is 3.789 with a corresponding p-value of 1.000. This result appears to be robust over the two subperiods. In the 1802–1896 period $\hat{\mu}_{up} = 3.67\%$ and $\hat{\mu}_{down} = -2.12\%$, while in the 1897–1990 period $\hat{\mu}_{up} = 9.02\%$ and $\hat{\mu}_{down} = -4.25\%$. These differences are again significant at usual levels (although the evidence is stronger for the latter subperiod).

Given that the conditional means are so different for upward versus downward sloping term structures, it is

[13] This finding is generally consistent with evidence presented in Pagan and Hong (1989) and Whitelaw (1994). The results in French, Schwert and Stambaugh (1987) and Harvey (1991) are more mixed.

[14] Boudoukh, Richardson and Smith (1993) provide a test of multiple inequality restrictions in the context of conditional asset pricing tests. However, since the restrictions in this paper are univariate, we avoid the complications that arise with multivariate one-sided tests.

Table 4 **The Risk Premium and the Term Structure Spread: Conditional Moments**

Period	P_{up}	$\hat{\mu}_{up}$	$\hat{\mu}_{down}$	$Z_{\mu_{up}=\mu_{down}}$	$\hat{\sigma}^2_{up}$	$\hat{\sigma}^2_{down}$	$Z_{\sigma^2_{up}=\sigma^2_{down}}$
1802–1990	0.614	0.0686	−0.0285	3.789	0.0267	0.0241	0.2491
		(0.0135)	(0.0184)	(1.000)	(0.0098)	(0.0063)	(0.598)
1802–1896	0.495	0.0367	−0.0212	1.814	0.0142	0.0205	−0.6630
		(0.0160)	(0.0216)	(0.965)	(0.0035)	(0.0090)	(0.2537)
1897–1990	0.734	0.0902	−0.0425	3.629	0.0341	0.0309	0.2461
		(0.0164)	(0.0326)	(1.000)	(0.0161)	(0.0057)	(0.5972)

Note: Table 4 provides the mean and variance of the risk premium conditional on whether the term structure is upward or downward sloping. In particular, the following moments are estimated:

$$E\begin{pmatrix} (rp_{t+1} - \mu_{up}) \times I_{t,up} \\ (rp_{t+1} - \mu_{down}) \times I_{t,down} \\ \{(rp_{t+1} - \sigma_{up})^2 - \sigma^2_{up}\} \times I_{t,up} \\ \{(rp_{t+1} - \mu_{down})^2 - \sigma^2_{down}\} \times I_{t,down} \end{pmatrix} = 0,$$

where $I_{t,up} = 1$ if $\Delta r_{ff,t} \geq 0$ and zero otherwise; $I_{t,down} = 1$ if $\Delta r_{ff,t} < 0$ and zero otherwise. Standard errors, in parentheses, are heteroskedasticity and serial correlation adjusted using Hansen's (1982) GMM. The statistics $Z_{\mu_{up}=\mu_{down}}$ and $Z_{\sigma^2_{up}=\sigma^2_{down}}$ test whether the conditional means and conditional variances are equal. Both test statistics are asymptotically normal, with *P*-values reported in parentheses. P_{up} is the fraction of years during the period for which the term structure is upward sloping. Annual data on stock returns, one-year rates and long-rates (maturity close to 20 years) are used.

of some interest to estimate their corresponding conditional variances. Table 4 presents estimates of these variances. Remarkably, there is almost no difference in the variance estimates during periods with upward versus downward sloping term structures. Over the entire sample period, the variance of the risk premium conditional on an upward sloping term structure is 2.67%, versus 2.41% conditional on a downward sloping term structure. At first glance, these results seem to contradict the variance estimates calculated from the kernel estimation performed in §5.1. Note, however, that the variance of the risk premium is increasing in the spread only at very high levels of the spread. In fact, Figure 3 shows that, for the majority of spreads between 0% and 0.5%, the variance of the risk premium is actually lower than the variance corresponding to negative spreads.

The overall conclusion from this analysis is that there is statistically reliable evidence that the expected risk premium has varied with the slope of the term structure over the past two centuries. This result appears robust across subsamples. Furthermore, the evidence suggests that the risk premium changes sign depending on whether the term structure is upward or downward sloping. Of some importance, the evidence implies that these changes in the risk premium do not seem to be due to changes in variance per se. While these results are consistent with equilibrium models of asset pricing (as demonstrated in the following section), they illustrate how important it is in practice to take account of time-varying risk.

6. A Stylized Example

To get a better understanding of the economic intuition behind the empirical results documented in §§3–5, we return to the theoretical framework of §2. At the level of generality of Eq. (1), little can be said about the relation between the risk premium and the slope of the term structure. Consequently, we make two simplifying assumptions:

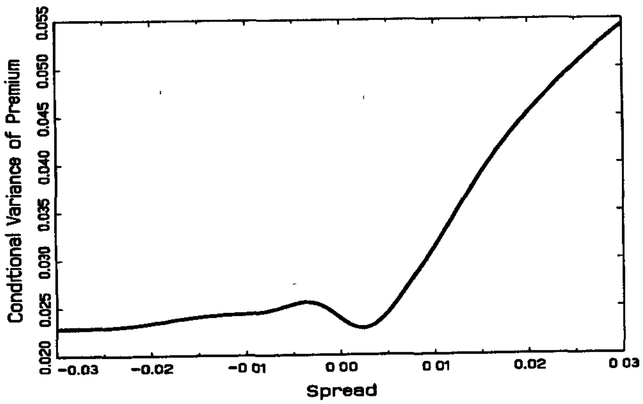
1. The representative agent has constant relative risk averse preferences with risk aversion parameter γ and time discount factor β, giving a MRS of $M_{t,t+1} = \beta(C_{t+1}/C_t)^{-\gamma}$, where C_t is consumption at time t.[15]

2. The dividend stream of the market portfolio is equal to the consumption stream of the representative agent, i.e., $D_{mt} = C_t$.

[15] This model is a specialization of the Lucas (1978) consumption-based model.

Figure 3 **The Variance of the Risk Premium and the Spread: Kernel Estimation**

The Figure Plots the Variance of the Risk Premium as a Function of the Spread, Estimated Using Variable Kernel Estimation.



This second assumption tends to work against two of the empirical results that we are attempting to replicate. Equating consumption and dividends links the MRS and the return on the market, and therefore makes negative risk premiums more difficult to attain, and it also strengthens the link between the risk premium and the volatility of returns. Nevertheless, this model is still sufficiently rich, yet tractable, to illustrate the relation between the risk premium and the term structure.

Time variation in both the equity risk premium and the term structure are determined by the dynamics of the consumption/dividend growth process. In order to solve for returns and interest rates, this process must be parameterized. For simplicity we consider a stylized, discrete state space, Markovian economy.[16] In particular, consumption/dividend growth can take on only one of four possible values each period. There are four possible states of the world, and the dynamic evolution of the economy is completely described by a four-by-four transition matrix (denoted Ω), which gives the probability of moving from any state to any other state (e.g., the (2,3) element of the matrix, $\omega_{2,3}$, gives the probability of going from state 2 at time t to state 3 at time $t + 1$). For our particular numerical example, we use the parameter values

$$\Omega = \begin{bmatrix} 0.50 & 0.25 & 0.25 & 0.00 \\ 0.25 & 0.50 & 0.25 & 0.00 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0.00 & 0.00 & 0.75 & 0.25 \end{bmatrix},$$

$$g^c = \begin{bmatrix} 1.04 \\ 1.02 \\ 1.00 \\ 0.92 \end{bmatrix}, \quad \beta = 0.93, \quad \gamma = 2.00, \qquad (19)$$

where g^c denotes consumption/dividend growth (i.e., $g^c = C_{t+1}/C_t = D_{mt+1}/D_{mt}$).

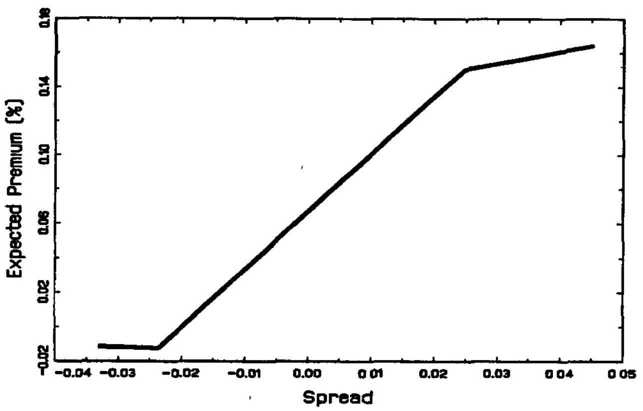
Note that these parameter values are not chosen to match any particular features of consumption data, but rather to broadly replicate the results of the empirical analysis. The intention is to illustrate the intuition underlying the relation between the equity risk premium and the slope of the term structure in as simple a setting as possible. From this perspective, states of the world with high consumption growth represent our stylized expansions, and those with low consumption growth represent contractions. The structure of the transition matrix governs the probability of moving between these states. For example, states 1 and 2 are expansionary states and state 4 is the contractionary state. However, $\omega_{1,4} = \omega_{2,4} = \omega_{4,1} = \omega_{4,2} = 0$, therefore there are no direct transitions from expansions to contractions or vice versa. Instead, the economy must go through the no-growth state (state 3) between phases.

Given this simple structure, it is straightforward to calculate conditional expectations of consumption/dividend growth, conditional on being in a particular state of the world. For example, $E_t[g^c_{t,t+1}] = \Omega g^c$, where the result is a four-by-one vector giving the expectation in each state of the world. By extension the price of a 1-period riskless bond in each state is simply the expectation of the MRS, i.e., $P_1 = \Omega\beta(g^c)^{-\gamma}$. The price of a 2-period bond is the expected discounted value of a 1-period bond, i.e., $P_2 = \Omega[\beta(g^c)^{-\gamma} .* P_1]$ (where $.*$ is element by element multiplication), and longer maturity bonds are defined analogously.

[16] Backus and Gregory (1993) use a similar framework for investigating the relation between risk premiums and conditional variances. Tauchen and Hussey (1991) provide a more general discussion of using discrete state space models to approximate continuous state space processes. Cecchetti, Lam, and Mark (1990) follow a related strategy of parameterizing a Markov switching model to investigate autocorrelations in equity returns.

Figure 4 **The Risk Premium and the Spread in a Stylized Model**

The Figure Plots the Expected Risk Premium versus the Yield Spread for the Discrete State Space, Markovian Economy Described in Section 6. The Parameter Values Are Given in Equation (19).



To calculate the risk premium, recall that equity returns are

$$R_{mt,t+1} = \frac{Q_{mt+1} + D_{mt+1}}{Q_{mt}} \quad (20)$$

$$= \frac{Q_{mt+1} + C_{t+1}}{Q_{mt}} \quad (21)$$

$$= \left(\frac{C_{t+1}}{C_t}\right)\frac{Q_{mt+1}/C_{t+1} + 1}{Q_{mt}/C_t}. \quad (22)$$

State-by-state consumption/dividend growth is given, hence to calculate returns it is sufficient to calculate the price-dividend ratio in each state of the world. Rewriting Eq. (1) in terms of the price-dividend ratio and substituting for the MRS

$$Q_t/C_t = E_t[\beta(g^c_{t,t+1})^{-\gamma}(Q_{t+1}/C_t + C_{t+1}/C_t)] \quad (23)$$

$$= E_t[\beta(g^c_{t,t+1})^{1-\gamma}(Q_{t+1}/C_{t+1} + 1)] \quad (24)$$

$$= E_t[\beta(g^c_{t,t+1})^{1-\gamma}] + E_t[\beta(g^c_{t,t+1})^{1-\gamma}Q_{t+1}/C_{t+1}]. \quad (25)$$

Solving for the price-dividend ratio (denoted Q/C).

$$Q/C = [I - (\mathbf{1}(\beta(g^c)^{1-\gamma})^T)\cdot*\Omega]^{-1}[\Omega\beta(g^c)^{1-\gamma}], \quad (26)$$

where I is the identity matrix, $\mathbf{1}$ is a vector of ones, and superscript T denotes transpose. Combining this price-dividend vector with the consumption/dividend growth vector gives realized equity returns as the economy moves from one state to any other state. It is then straightforward to compute the conditional expectations and variances of these returns using the transition matrix.

For the parameter values above, Figure 4 graphs the expected equity risk premium against the slope of the term structure for each of the four states of the world. The state-by-state values are also provided in Table 5. The slope of the term structure is defined as the yield spread between a 10-period risk-free bond and a 1-period risk-free bond, although similar patterns arise for any spread between long-term and short-term risk-free yields. The figure illustrates a number of interesting features. First, the spread and the equity risk premium are generally positively related, although this relation is nonlinear. Note that the graph resembles quite closely the estimated empirical relation shown in Figure 2.

Table 5 **The Risk Premium and the Yield Spread in a Stylized Model**

State	g^c	Q/C	R_f	R_l	Δr_{lf}	$E_t[rp_{t+1}]$	$Var_t[rp_{t+1}]$
1	1.04	11.62	12.88%	9.58%	−3.30%	−0.012%	0.96
2	1.02	11.69	11.81%	9.45%	−2.36%	−0.013%	1.33
3	1.00	12.26	5.90%	8.23%	2.48%	0.150%	3.84
4	0.92	12.72	3.02%	7.38%	4.52%	0.165%	4.79

Note: Table 5 provides state-by-state values of consumption growth, the price-dividend ratio, the 1-period risk-free rate, the 10-period risk-free rate, the yield spread, the expected risk premium, and the variance of the risk premium for the discrete state space, Markovian, consumption-based model described in Section 6. Parameter values are given in Equation (19).

During expansions (states 1 and 2), short-term expected consumption growth is high, and therefore short-term interest rates are also high. In the long-term, consumption growth is expected to revert to its long-run mean, therefore long-term rates are lower. Together these conditions imply inverted term structures (see Table 5). The expected equity risk premium is also low because the return on equity does not covary negatively with the MRS—a point which is discussed in more detail below. The opposite is true in contractions (state 4)—short-term rates are low, the term structure is upward sloping, and the equity risk premium is higher. The countercyclical behavior of both the expected risk premium and the yield spread has been noted previously in the literature, as discussed in §2.

A second interesting observation from Figure 4 is that negative risk premiums occur in states 1 and 2 (expansions). Recall that the risk premium can be negative if, and only if, the return on equity covaries positively with the MRS (see Eq. (8)). In this simple world, this condition reduces to

$$\mathrm{Cov}_t\left[g^c_{t,t+1}\,\frac{1+Q_{mt+1}/C_{t+1}}{Q_{mt}/C_t},\,(g^c_{t,t+1})^{-2}\right] > 0. \qquad (27)$$

At first glance, this seems unlikely, because consumption growth will clearly vary inversely with the inverse of consumption growth. The covariance can only be positive if the second element of market returns dominates the dividend growth effect, i.e., if the price-dividend ratio decreases when consumption growth is high. In fact, this is exactly what happens, as can be seen from the price-dividend ratios reported in Table 5. High consumption growth implies high dividends, but it also generates low price-dividend ratios. The intuition behind this result is relatively straightforward. Recall from Eq. (9) that the price is the expected discounted value of future dividends. Consumption/dividend growth is persistent; therefore, high growth today implies high growth in the future. However, for a risk aversion parameter greater than 1, the effect of growth on the discount rate dominates the effect on future dividends. High growth actually hurts the value of the equity. Thus, the effect of current dividends on the equity return is more than offset by the price effect, i.e., the capital gain or loss. As a result, equity is actually a hedge against aggregate consumption shocks. Consequently, investors are willing to accept a lower expected return on equity than on the risk-free bond.

Finally, the variance of the risk premium is not monotonically related to the expected risk premium. In this particular numerical example, the variance is monotonically related to the spread, but the expected risk premium is not (see Table 5). This result is simply a function of the fact that in this model risk is measured by covariance with the MRS. The variance of the risk premium is not necessarily a good proxy for priced risk; therefore there is no reason to expect the variance to be related to the expected risk premium.

7. Conclusion

We find that there is a *statistically* significant nonlinear relation between the equity risk premium and slope of the term structure. From a practical viewpoint, our results imply that the linear approximations implicit in prior work may brush aside useful information regarding the predictive power of the term structure for movements in the expected risk premium. For example, we show that changes in the term structure spread have vastly different implications for expected stock return premiums, depending on the level of the spread itself. In addition, the nonlinearity, the sign change in the risk premium, and the apparent importance of upward versus downward sloping term structures provide an interesting set of stylized facts to be explained by equilibrium asset pricing models.[17]

[17] We would like to thank Jeremy Siegel for use of the data and the editor, Rob Heinkel, two anonymous referees, Jay Shanken, and seminar participants at Duke University and the 1993 WFA meetings for helpful comments.

References

Backus, D. and A. Gregory, "Theoretical Relations Between Risk Premiums and Conditional Variances," *J. Business and Economic Statistics*, 11 (1993), 177–185.

Boudoukh, J., M. Richardson and T. Smith, "Is the Ex Ante Risk Premium Always Positive? A New Approach to Testing Conditional Asset Pricing Models," *J. Financial Economics*, 34 (1993), 387–408.

——, ——, R. Stanton, and R. F. Whitelaw, "Pricing Mortgage-Backed Securities in a Multi-Factor Interest Rate Environment: A Multivariate Density Estimation Approach," *Rev. Financial Studies* (1997), forthcoming.

Breiman, L., W. Meisel, and E. Purcell, "Variable Kernel Estimates of Multivariate Densities," *Technometrics*, 19 (1977), 135–144.

Campbell, J. Y., "Bond and Stock Returns in a Simple Exchange Model," *Quarterly J. Economics*, 100 (1986), 785–802.
——, "Stock Returns and the Term Structure," *J. Financial Economics*, 18 (1987), 373–399.
Cecchetti, S. G., P.-S. Lam, and N. C. Mark, "Mean Reversion in Equilibrium Asset Prices," *American Economic Rev.*, 80 (1990), 398–418.
Chen, N., "Financial Investment Opportunities and the Macroeconomy," *J. Finance*, 46 (1991), 529–554.
Cover, T., "A Hierarchy of Probability Density Function Estimates," *Frontiers of Pattern Recognition*, Academic Press, New York, 1972.
Epanechnikov, V., "Nonparametric Estimates of Multivariate Probability Density," *Theory of Probability and Applications*, 14 (1969), 153–158.
Estrella, A. and G. Hardouvelis, "The Term Structure as a Predictor of Real Economic Activity," *J. Finance*, 46 (1991), 555–576.
Fama, E., "Term Premiums and Default Premiums in Money Markets," *J. Financial Economics*, 17 (1986), 175–196.
——, "Efficient Capital Markets: II," *J. Finance*, 46 (1991), 1575–1617.
—— and K. French, "Forecasting Returns on Stocks and Bonds," *J. Financial Economics*, 25 (1989), 23–49.
Ferson, W. and C. Harvey, "The Variation of Economic Risk Premiums," *J. Political Economy*, 99 (1991), 385–415.
Foster, D. and T. Smith, "Assessing Goodness-of-Fit of Asset Pricing Models: The Distribution of the Maximal R^2," *J. Finance* (1997), forthcoming.
French, K., W. Schwert, and R. Stambaugh, "Expected Stock Returns and Volatility," *J. Financial Economics*, 19 (1987), 3–29.
Gallant, R., "Unbiased Determination of Production Technologies," *J. Econometrics*, 20 (1982), 285–323.
Gibbons, M. and W. Ferson, "Tests of Asset Pricing Models with Changing Expectations and an Unobservable Market Portfolio," *J. Financial Economics*, 14 (1985), 217–236.
Hansen, L., "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50 (1982), 1029–1054.
—— and K. Singleton, "Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns," *J. Political Economy*, 91 (1983), 249–265.
Harrison, M. and D. Kreps, "Martingales and Arbitrage in a Multiperiod Security Markets," *J. Economic Theory*, 20 (1979), 381–408.
Harvey, C., "The Real Term Structure and Consumption Growth," *J. Financial Economics*, 22 (1988), 305–334.
——, "The Specification of Conditional Expectations," Working Paper, Duke University, Durham, NC, 1991.
Hawawini, G. and D. Keim, "On the Predictability of Common Stock Returns: Worldwide Evidence," in R. A. Jarrow, V. Maksimovic, and W. T. Ziemba (Eds.), *The Handbook of Finance*, North-Holland, Amsterdam, 1995.
Kairys, J. P., "Predicting Sign Changes in the Equity Risk Premium Using Commercial Paper Rates," *J. Portfolio Management*, Fall (1993), 41–51.
Keim, D. and R. Stambaugh, "Predicting Returns in the Stock and Bond Markets," *J. Financial Economics*, 17 (1986), 357–390.
Kessel, R., "The Cyclical Behavior of the Term Structure of Interest Rates," Occasional Paper 91, National Bureau of Economic Research, New York, 1965.
Lucas, R., "Asset Pricing in an Exchange Economy," *Econometrica*, 16 (1978), 1429–1446.
Merton, R., "On Estimating the Expected Return on the Market: An Exploratory Investigation," *J. Financial Economics*, 8 (1980), 323–361.
Pagan, A. and Y. S. Hong, "Non-parametric Estimation and the Risk Premium," in W. Barnett, J. Powell and G. Tauchen (Eds.), *Semiparametric and Nonparametric Methods in Econometrics and Statistics*, Cambridge University Press, Cambridge, 1991.
Schwert, W., "Indexes of Stock Prices from 1802 to 1987," *J. Business*, 63 (1990), 399–426.
Scott, D. W., *Multivariate Density Estimation: Theory, Practice and Visualization*, John Wiley & Sons, Inc., New York, 1992.
Siegel, J., "The Real Rate of Interest from 1800–1990: A Study of the U.S. and U.K.," *J. Monetary Economics*, 29 (1992), 227–252.
Tauchen, G. and R. Hussey, "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59 (1991), 371–395.
Ullah, A., "Non-Parametric Estimation of Economic Functionals," *Canadian J. Economics*, 21 (1988), 625–658.
Whitelaw, R. F., "Time Variations and Covariations in the Expectation and Volatility of Stock Market Returns," *J. Finance*, 49 (1994), 515–541.

Accepted by Robert Heinkel; received March 30, 1994. This paper has been with the authors 15 months for 3 revisions.