THE MONOTONICITY OF THE TERM PREMIUM: ANOTHER LOOK

by

Matthew Richardson
Paul Richardson
Tom Smith

3-92

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

The Monotonicity of the Term Premium: Another Look

Matthew Richardson, Paul Richardson and Tom Smith*

January 13, 1992

Abstract

This paper reexamines evidence on the monotonicity of the term premium. Using a recently developed approach for testing inequality constraints, we propose and conduct tests for whether the term premium is monotonic and reach different conclusions from those implied by individual t-statistics on term premiums, even under a Bonferroni-type adjustment. Our results generally support McCulloch's (1987). view that the liquidity preference hypothesis remains unrefuted.

*Assistant Professor (The Wharton School, University of Pennsylvania), Vice President (Lehman Brothers), and Assistant Professor (Fuqua School of Business, Duke University), respectively. We would like to thank Kobi Boudoukh, Michael Gibbons, Rob Stambaugh, William Schwert (the editor), and Eugene Fama (the referee) for helpful comments and suggestions. Research support from the Geewax-Terker Research Fund is gratefully acknowledged. The views expressed here are those of the authors and do not necessarily reflect the views of Lehman Brothers.

1 Introduction

The liquidity preference hypothesis implies that the expected return on Treasury securities increases monotonically with remaining time to maturity. Fama (1984) concludes from evidence on term premiums for T-bills that expected returns are not monotonic. In a reply, McCulloch (1987) questions Fama's conclusions and argues that his results are due to anomalous behavior in bid-ask mean prices of T-bills at certain maturities.

The disagreement centers around the bid-ask spread of nine-month T-bills during the 1964 — 1972 period, when the Treasury conducted nine-month bill auctions. McCulloch finds that the spread during this period is about half the spread in the surrounding months. Using a transaction-cost analysis, he argues that the drop in the term premium from nine to ten months may be spurious. But even though the nine-month bid-ask spread argument is compelling, it does not necessarily lead to acceptance of the monotonicity of the term premium.

The reason is that monotonicity mplies not only that the ten-month premium exceeds the nine-month premium, but also that the eleven-month premium exceeds the nine-month premium (and so on). For example, Fama (1984) finds that the premium at nine, ten and eleven months is .089%, .057%, and .064%, respectively. Even if .057% is spurious because of the anomalous bid-ask behavior, this does not explain the drop from .089% to .064%. Moreover, contrary to McCulloch's argument, this drop also occurs in the 1973-1982 period as well as the overall period Fama examines, 1964-1982. At least for individual t-statistics, therefore, evidence suggests expected returns are in fact not monotonically increasing.

But there is reason to be cautious in interpreting Fama's results. In particular, consider results from the overall sample period. Only one of eleven j-month term premiums (either ten-month or eleven-month, depending on McCulloch's argument) lies significantly below a $j-1$-month premium. At what level should we judge this significance? We would expect, for example, that even if expected returns were monotonic, there would be some drops in the term premium due simply to sampling error. Fama's results become even more difficult to interpret when we recognize the high correlation across the premiums for different maturities.

The difficulty in taking account of this correlation and jointly testing the monotonicity of the term premium is that monotonicity implies a series of inequality constraints. It does not suffice merely to ntest the equality of the premiums, which would be the standard approach to testing general linear restrictions. Recent work by Gourieroux, Holly and Monfort (1982), Kodde and Palm (1986) and Wolak (1989), among others, however, provides a way of testing inequality constraints directly within this type of framework. Therefore, we can in fact test the null hypothesis that expected returns are monotonic and specifically address the question of how much statistical weight should be given to a particular maturity.

Using this inequality testing method, we find that our results generally support McCulloch's view. Rejection in this joint framework occurs only when the anomalous ten-month minus nine-month premium (over 1964-1972) is included in the sample. That is, even though the individual t-statistics of the eleven-month minus nine-month premium imply nonmonotonocity, this evidence is consistent with sampling error in the data. As suggested by McCulloch (1987), apparently too much weight has been placed on the maturities in question. Moreover, this conclusion is consistent across all the subperiods as well as the overall sample.

This paper is organized as follows. Section 2 describes the inequality testing method for the null hypothesis that expected returns are monotonically increasing. Section 3 analyzes this hypothesis empirically, with and without McCulloch's anomalous month. Section 4 concludes.

2 Statistical Method

In this section, we describe the test for whether expected returns are monotonic. Our discussion of inequality constraints follows Wolak (1989). We provide only a basic outline. For more details, see Wolak (1989) or Gourieroux, Holly, and Monfort (1982).

2.1 Liquidity Preference Hypothesis Set-up

Following Fama (1984), define $H_{\tau_{t+1}}$ as the continuously compounded return from t to $t+1$ on a bill with maturity τ at date t. Also, let $H_{l_{t+1}}$ be the continuously compounded return calculated from the price of a bill with one month to maturity at t. The term premium in the return on a bill with maturity τ is defined as

$$P_{\tau_{t+1}} \equiv H_{\tau_{t+1}} - H_{l_{t+1}}.$$

The liquidity preference hypothesis implies that the τ-month premium ($P_{\tau_{t+1}}$) is at least as large as the $(\tau - i)$-month premium ($P_{(\tau-i)_{t+1}}$) for all i. In particular, we wish to test the hypothesis that

$$P_{(N+1)_{t+1}} \geq P_{N_{t+1}} \geq \cdots \geq P_{2_{t+1}} \quad \tau = 2, \ldots, N+1.$$

2.2 Test methodology

Define P as the $(NT \times 1)$ vector of T observations on the N term premiums. Let α be the N vector of means of P. Let ϵ be a $(NT \times 1)$ random vector that is distributed $N(0, \Sigma \otimes I_T)$, where Σ is the $N \times N$ variance-covariance matrix across the term premiums. Let $\underline{1}$ be a $T \times 1$ matrix of ones. If we define the $([N-1] \times N)$ matrix of constants R to be

$$R = \begin{pmatrix} 1 & -1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & -1 & 0 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & 1 & -1 \end{pmatrix},$$

then the testing framework for monotonicity in term premiums can be written as

$$P = (I_N \otimes \underline{1})\alpha + \epsilon \tag{1}$$

$$H_0 : R\alpha \geq 0 \quad \text{versus} \quad H_A : R\alpha \in R^{N-1}.$$

As an illustration, we derive the analog to the Wald statistic for testing inequality constraints in (1).

Step 1

Calculate the unrestricted estimate $\hat{\alpha}$ from the minimization:

$$\min_{\alpha}(P - (I_N \otimes \underline{1})\alpha)'(\hat{\Sigma}^{-1} \otimes I_T)(P - (I_N \otimes \underline{1})\alpha),$$

where $\hat{\Sigma}^{-1}$ is a consistent estimate of Σ.

Step 2

Calculate the restricted estimate $\bar{\alpha}$ from the minimization:

$$\min_{\alpha}(P - (I_N \otimes \underline{1})\alpha)'(\hat{\Sigma}^{-1} \otimes I_T)(P - (I_N \otimes \underline{1})\alpha)$$

subject to $R\alpha \geq 0$.

Step 3

Using results in Wolak (1989), it is possible to derive the Wald statistic for testing the null hypothesis that term premiums are monotonically increasing. Specifically, we get

$$\begin{aligned} W &\equiv [R\bar{\alpha} - R\hat{\alpha}]'[R(\hat{\Sigma}^{-1} \otimes T)^{-1})R']^{-1}[R\bar{\alpha} - R\hat{\alpha}] \\ &\equiv T[R\bar{\alpha} - R\hat{\alpha}]'[R\hat{\Sigma}R']^{-1}[R\bar{\alpha} - R\hat{\alpha}]. \end{aligned}$$

Step 4

Evaluate the statistic W at the appropriate level of significance. Wolak (1989) and others show that the Wald statistic here, unlike the Wald statistic under equality constraints, no longer has an asymptotic chi-squared distribution. Instead, the statistic is now distributed as a weighted sum of chi-squared variables with different degrees of freedom. Specifically, the asymptotic distribution of W is given by

$$\sum_{k=0}^{N-1} Pr[\chi_k^2 \geq c]\; w\left(N-1, N-1-k, \frac{R\hat{\Sigma}R'}{T}\right),$$

where $c \in R^+$ and the weight $w\left(N-1, N-1-k, \frac{R\hat{\Sigma}R'}{T}\right)$ is the probability that $\bar{\alpha}$ has exactly $N-1-k$ positive elements.

Although the weights have a closed-form solution for a small number of restrictions ($N \leq 5$), the monotonicity test requires ten restrictions across the T-bill maturities. Kodde and Palm (1986, table 1, page 1246), however, provide upper-bound and lower-bound critical values that do not require calculation of the weights. For a given level of significance, the econometrician rejects monotonicity if the Wald statistic W exceeds the upper bound. Similarly, he cannot reject this hypothesis if W is less than the lower bound. Only for values within these bounds does he have to calculate the weights.

Although the weights are not tractable for a large number of restrictions, Wolak (1989) describes an approximate method for calculating the weights based on a Monte Carlo simulation. Specifically, in our case, the econometrician should simulate a multivariate normal distribution with mean zero and covariance matrix $\left(\frac{R\hat{\Sigma}R'}{T}\right)$ (denote these values α^*). The next step is to find the solution $\hat{\tilde{\alpha}}$ to the minimization problem

$$\min_{\tilde{\alpha}}(\alpha^* - \tilde{\alpha})' \left(\frac{R\hat{\Sigma}R'}{T}\right)^{-1} (\alpha^* - \tilde{\alpha})$$

$$\text{subject to } \tilde{\alpha} \geq 0.$$

For each replication, count the number of elements in the $N-1$ vector $\hat{\tilde{\alpha}}$ that exceed zero. As Wolak points out, the approximate weight $\hat{w}\left(N-1, N-1-k, \frac{R\hat{\Sigma}R'}{T}\right)$ will be the fraction of replications in which $\hat{\tilde{\alpha}}$ has exactly $N-1-k$ elements exceeding zero.

3 Empirical Results

3.1 Replication of Fama's results

Fama (1984) performs individual t-tests for first-lag differences in the term premium over various subperiods and the overall sample period 1964-1982. Table 1 replicates these results and extends the sample period to 1990. There are minor differences between the two data sets. In particular, we do not use the twelve-month bill. The Fama files define a twelve-month bill as the longest bill with more than eleven months and ten days to maturity. We consider this definition too unreliable. As a result of dropping the twelve-month bill, we have more observations than Fama. Eleven-month bills are not available for two months of the 1973-1981 subperiod, and a ten-month bill is missing for one month in the same subperiod. Following Fama, we delete these months for all maturities.

Other than in the latest sample period (1982-1990), our results seem to coincide with Fama's conclusions. The individual results in table 1 imply that the term premium reaches a peak around nine months. For example, the average premium drops from .082% per month to .027% and from .044% per month to .024% over the sample periods 1964-1972 and 1973-1981. Moreover, these drops are both significant at the 10% level for t-statistics.[1] The strength of these results carries through to the overall sample period 1964-1990 and Fama's overall period, 1964-1982.

[1] Note, however, that the 1973-1981 period is not significant using the Bonferroni-adjusted levels. (See Fama (1984), table 1).

As McCulloch (1987) points out, however, these results may be spurious because of anomalous behavior in the bid-ask spread of the nine-month bill during the 1964-1972 period. We therefore report results for lagged differences in the term premium without this anomalous bill. Specifically, we report individual t-statistics for the difference between the nine- and ten-month term premiums (i.e., $H_{11} - H_9$). Over the 1964-1972 and 1973-1981 subperiods, Fama's 1964-1982 period, and the overall period 1964-1990, the t-statistic for $H_{11} - H_9$ is significant at the 10% level.[2] At least on the basis of the individual t-statistics, McCulloch's compelling argument on monontonicity appears to be a strawman.

But this is not necessarily the case. McCulloch's general point is that without an a priori theory, we seem to be placing too much weight on particular maturities (e.g., in his analysis, the ten-month premium). The statistical significance of the individual t-statistics, therefore, may be due to sampling error rather than true nonmonotonicity in term premiums. For example, we would expect to find some spurious declines in the term premiums over a small sample even if expected returns are monotonic.

For these reasons, it is difficult to interpret the individual t-statistics in table 1. For example, given results for the other lagged differences in the term premium, can the appropriate level of significance for $H_{11} - H_9$ really be taken from a standardized normal distribution? Fama and McCulloch recognize this problem and use a Bonferroni-type approach to test for the monotonicity of the term premium. As Wolak describes, however, the Bonferroni-type approach to testing multiple inequality constraints suffers from major shortcomings when the parameter estimators are highly correlated. But we can form a joint test that the term premiums are monotonic across the maturities. Unfortunately, the Hotelling T^2 tests equality across the maturities and is therefore not a test for monotonicity. The appropriate test needs to take account of the inequality constraints implied by monotonicity. In the next subsection, we provide results for these inequality tests.

3.2 Tests for monotonicity

Table 1 reports results for the Wald test statistic derived in section 2. Among the subperiods, only 1964-1972 rejects the monotonicity of the term premium. The value of the statistic is 33.06, which is well above the 5% cut-off value. In contrast, the p-values for the 1973-1981 and 1982-1990 periods are .47 and .71. For the overall period and Fama's pe-

[2]As with the results that include the ten-month premium, the 1973-1981 period is again not significant using a Bonferroni-type adjustment.

riod, the strength of the earlier subperiod leads to rejection of monotonicity. Nevertheless, these periods include McCulloch's anomalous premium.

When we take out the anomalous month, the Wald tests no longer reject monotonicity. This is true of all three subperiods, the overall period, and Fama's period. Of particular interest, this occurs even though some of the individual t-statistics suggest nonmonotonicity. For example, consider the sample period 1964-1972. The t-statistic for $H_{11} - H_9$ is -2.38, which is significant at the 10% level using Bonferroni-adjusted table values. However, the joint test of the inequality constraints is 4.00, which has a p-value equal to .39.

The evidence is consistent with McCulloch's argument that the humped shape in term premiums is due to bid-ask spreads in nine-month bills during the 1964-1972 period. Other evidence of nonmonotonicity (as in the "significantly" negative eleven-month minus nine-month premium) is more likely due to sampling error. That is, the joint tests of monotonicity do not reject in any of the other subperiods and in periods without the anomalous T-bill. At the very least, these results illustrate the different conclusions one can reach from individual t-statistics and tests for inequality constraints.

4 Conclusion

A recent paper by McCulloch (1987) argues that Fama's (1984) evidence of nonmonotonicity in the term premium is spurious. Although there is reason to question the argument with respect to the individual t-statistics, McCulloch's point holds up well under further scrutiny. In particular, we need to be cautious when interpretating Fama's and McCulloch's results across the T-bill maturities. This is because a test for monotonicity implies a series of inequality constraints and is not necessarily equivalent to the equality of means test or Bonferroni-type tests performed by Fama. We reexamine the term premium evidence, incorporating the inequality constraints implied by the monotonicity hypothesis. In accordance with McCulloch, our results do not support Fama's conclusion that expected returns do not increase with remaining time to maturity. Moreover, this evidence appears to hold across various subperiods as well as for the overall sample period.

As an aside, the general approach to testing models with inequality constraints seems well suited to financial economics applications. In a number of financial models, theory tells us the sign (but not the magnitude) of the parameter restrictions. In addition, orderings across parameters are implied by models of expected return and risk and also by recent models describing the short-horizon and long-horizon behavior of stock returns. The papers by Gourieroux, Holly, and Monfort (1982) and Wolak (1989) provide a good starting point for further investigations.

TABLE 1

Average premiums, t statistics, and Wald inequality tests for T-bills with up to eleven months to maturity

The first section of table 1 provides the average premium P_τ, where P_τ equals the the log monthly return on a T-bill maturing in τ months minus the log monthly return on a T-bill maturing in one month. t-statistics for the difference in premiums is provided in the second section. The third and fourth sections provide Wald test statistics of the hypothesis that the premiums are monotonically increasing. The tests are conducted over the 1964-1972, 1973-1981, and 1982-1990 subperiods and the overall period 1964-1990 (T is the number of observations in each sample).

Bill	T=313 8/64-11/90	T=101 8/64-12/72	T=105 1/73-12/81	T=107 1/82-11/90	T=218 8/64-12/82
		Average Premiums (P_τ)			
P2	.030	.028	.024	.039	.026
P3	.056	.045	.049	.074	.054
P4	.055	.045	.046	.074	.055
P5	.070	.061	.047	.102	.068
P6	.073	.066	.038	.114	.067
P7	.070	.064	.035	.108	.065
P8	.088	.082	.045	.135	.082
P9	.093	.082	.044	.151	.084
P10	.071	.027	.024	.158	.051
P11	.077	.061	.009	.159	.060
	t statistics for average values of premium differences ($P_\tau - P_{\tau-1}$)				
$P_3 - P_2$	6.78	5.58	2.97	4.81	5.30
$P_4 - P_3$	-0.12	-0.00	-0.23	0.15	0.22
$P_5 - P_4$	3.77	3.83	0.07	4.37	2.37
$P_6 - P_5$	0.80	0.87	-1.12	2.60	-0.20
$P_7 - P_6$	-0.80	-0.37	-0.28	-1.28	-0.31
$P_8 - P_7$	4.55	3.37	1.08	4.53	3.32
$P_9 - P_8$	1.13	-0.01	-0.10	2.98	0.28
$P_{10} - P_9$	-3.97	-5.75	-1.75	1.20	-4.49
$P_{11} - P_{10}$	1.36	5.35	-1.39	0.21	1.43
$P_{11} - P_9$	-2.00	-2.38	-1.79	0.91	-2.28
	Wald test of nine inequality constraints ($P_\tau - P_{\tau-1} \geq 0\ \forall\tau = 3,\ldots,11$)				
Wald	15.79	33.06	3.34	1.64	20.18
P-level	.0131	.0000	.4700	.7077	.0023
5% Cut-off	11.90	11.46	11.76	11.73	11.73
	Wald test of eight inequality constraints ($P_\tau - P_{\tau-1} \geq 0\ \forall\tau = 3,\ldots,9$; $P_{11} - P_9 \geq 0$)				
Wald	4.00	5.65	3.21	1.64	5.19
P-level	.3853	.2470	.4410	.6480	.2755
5% Cut-off	10.94	10.53	10.71	10.47	10.58

References

Fama, Eugene, 1984, Term premiums in bond returns, *Journal of Financial Economics* 13, 529–546.

Gourieroux, Christian, Alberto Holly, and Alain Monfort, 1982, Likelihood ratio test, wald test, and kuhn-tucker test in linear models with inequality constraints on the regression parameters, *Econometrica* 50, 63–80.

Kodde, David A. and Franz C. Palm, 1986, Wald criteria for jointly testing equality and ineaquality restrictions, *Econometrica* 54, 1243–1248.

McCulloch, J. Huston, 1987, The monotonicity of the term premium, *Journal of Financial Economics* 18, 185–192.

Wolak, Frank A., 1989, Testing inequality constraints in linear econometric models, *Journal of Econometrics* 41, 205–235.