15

Journal of Financial Economics 18 (1987) 000–000. North-Holland

THE MONOTONICITY OF THE TERM PREMIUM
A Closer Look

J. Huston McCULLOCH*

Ohio State University, Columbus, OH 43210, USA

Received September 1985, final version received February 1986

Fama's evidence that the term premium on Treasury securities is not monotonically increasing is found to depend entirely on the behavior of bid–asked mean returns on 9- and 10-month bills, and only during the subperiod 8/64–12/72. When transactions costs, as reflected in the bid–asked spread, are taken into account, there is found to be no way to exploit this non–monotonicity. The anomalous behavior of the quotations is attributed to the Treasury's auctions of 9-month bills during the period 9/66–10/72. The hypothesis that the term premium is a monotonically increasing function of maturity remains unrefuted.

1. Introduction

In a recent paper in this Journal, Fama (1984b) reports evidence that appears to reject the hypothesis that the expected return on Treasury securities is a monotonically increasing function of remaining time to maturity. This is contrary to the hypothesis of Kessel (1965) that the term premium is a monotonically increasing 'liquidity' premium, and to my own empirical findings [McCulloch (1975)] that monotonicity cannot be rejected.

Table 1 shows Fama's estimates of the term premium for the Treasury bill range. This premium is defined as the expected excess return on τ-month bills over 1-month bills, and is measured as the ex post average of this excess return. The first column shows the estimated value of this premium for Fama's longest period, 8/64–12/82, with t-statistics in parentheses for the null hypothesis that the premium is zero. The second column shows the first difference of this series, with Fama's estimated t-statistics for the null hypothesis that the first difference is zero. The third and fourth columns show the same values for the subperiod 8/64–12/72. Clearly the premium is positive and rising to at least month 3.

The heart of Fama's evidence against monotonicity is the sharp decline in the estimated premium between 9 and 10 months. The accompanying t-statis-

*The author is grateful to Min-Teh Yu for assistance with data collection, and Stephen Buser, Beni Lauterbach, Anthony Sanders, and an anonymous referee for helpful comments and discussions. Any errors or omissions remain the responsibility of the author.

Table 1

Fama's estimates of term premiums in the period 8/64–12/82. Estimated premium is the mean excess logarithmic return on τ-month bills over 1-month bills. Premium estimates are percents per month. t-statistics are given in parentheses.[a]

	8/64–12/82		8/64–12/72	
Maturity τ (months)	Premium (%)	First difference (%)	Premium (%)	First difference (%)
2	0.032 (6.40)		0.028 (6.97)	
3	0.057 (6.40)	0.025 (4.68)	0.045 (7.17)	0.017 (4.40)
4	0.063 (4.70)	0.006 (1.05)	0.046 (5.18)	0.001 (0.08)
5	0.074 (4.14)	0.011 (1.93)	0.061 (5.12)	0.015 (3.01)
6	0.073 (3.34)	−0.001 (−0.27)	0.066 (4.35)	0.005 (0.74)
7	0.069 (2.75)	−0.004 (−0.72)	0.071 (3.68)	0.005 (0.81)
8	0.088 (3.04)	0.019 (3.20)	0.084 (3.86)	0.013 (2.32)
9	0.089 (2.59)	0.001 (0.10)	0.086 (3.27)	0.002 (0.19)
10	0.057 (1.49)	−0.032 (−4.15)	0.025 (0.83)	−0.061 (−5.66)
11	0.064 (1.54)	0.007 (1.06)	0.066 (2.09)	0.041 (5.00)
12	0.074 (1.61)	0.010 (1.28)	0.103 (2.87)	0.037 (4.27)

[a]*Source*: Fama (1984b).

tics for the first difference are −4.15 for the long period and −5.66 for the subperiod. The differences for longer maturities out to 10 years (that we do not tabulate here) are primarily negative, suggesting a declining term premium. However, only a handful of these negative differences exceed 1.96 in absolute value and none of these exceeds the Bonferroni adjusted levels Fama gives in his table 1 for multiple t-tests. The two t-statistics noted above are the only ones that are negative and still exceed these adjusted critical values at even the 0.90 level on a two-tailed test.

Furthermore, the strong negative value of the first difference between the 9- and 10-month maturities virtually vanishes in the later subperiods Fama reports. The point estimates are still negative, but the t-statistic never exceeds

Table 2

Summary statistics for Treasury bill quotations in the period 7/31/64–12/29/72. Measured by average month-end banker's discount rates, expressed as a percent per annum.[a]

Maturity (months)	Mean bid discount	Mean asked discount	Mean bid–asked spread
8	5.171	5.074	0.097
9	5.191	5.133	0.058
10	5.185	5.078	0.107

[a]*Data source*: *Wall Street Journal*, except as indicated in footnote 2.

0.50 in absolute value. Fama's evidence against monotonicity is therefore specific to these two maturities *and* to the subperiod 8/64–12/72.[1] A pattern of primarily declining premium point estimates does persist after nine months, but to reiterate, Fama's paper presents no evidence that this decline is anything more than sampling error.

2. Replication of Fama's results

In order to check Fama's results with a somewhat independent data source, I collected month-end bid and asked banker's discount quotations on 8-, 9-, and 10-month Treasury bills from the *Wall Street Journal* for Fama's critical period 7/31/64–12/29/72. (The extra month is necessary to calculate the returns for the month ending on 8/31/64.) Suspicious quotations were checked against the *New York Times* for accuracy.[2]

Table 2 shows the mean bid discount rate, the mean asked discount rate, and the mean bid–asked spread for the three critical maturities. We see that the bid–asked spread is slightly smaller, on average, for 8-month bills than for 10-month bills in terms of yield, and therefore also in terms of price. The bid–asked spread for 9-month bills, however, is barely half that for the other two maturities, in terms of either price or yield. Apparently there is something

[1]Fama also provides tests of the compound hypothesis that the term premium is flat beyond 2 months. This hypothesis is a strawperson, however, since much of the strength of his 'T^2' statistics comes from the very significant *rise* in the premium between 2 and 3 months, plus the generally continuing, but not so significant, rise in the point estimates out to the ninth month. A fairer test would have been whether the premium is flat beyond approximately 5 months. See section 6 below.

[2]For 11/30/67, the *WSJ* reported impossible bid and asked discounts of 5.35 and 5.43, respectively, for the note maturing 9/30/68. The *NYT* quotes of 5.52/5.42 were substituted. Also, for 10/31/72, the *WSJ* reported equal bid and asked discounts of 5.10 for the 8-month bill, so the more plausible *NYT* quotations, 5.18 and 5.10, were substituted. No other substitutions were made. Like Fama, we take each month to be 365/12 days long. This approximation is more than adequate for the purpose at hand.

Table 3

Mean 1-month return on 9- and 10-month Treasury bills in percent per month. Values in square brackets are normalized von Neumann ratios which have been adjusted to have mean zero and unit standard deviation under the null hypothesis of zero population serial correlation, and to have the same sign as the sample serial correlation coefficient. (See footnote 6).

	8/64–12/72	
Maturity	Bid–asked Mean	Buy asked, Sell bid
9 months	0.469	0.412
10 months	0.415	0.486
First difference	−0.054	0.074
t-statistic	(−5.58)	(7.56)
NvNR	[6.06]	[6.78]

special about 9-month bills during this period that makes their transactions costs much smaller than for the other two maturities.

Fama does not report whether he is using bid prices, asked prices, or bid–asked mean prices, but only that they come from the CRSP data bank. He has privately informed me that his results were based on bid–asked mean prices.

The first column of table 3 reports the mean logarithmic 1-month return on 9- and 10-month bills during 8/66–12/72, using bid–asked mean prices derived from the discount quotations. The return on 10-month bills is indeed 0.054% lower than that on 9-month bills, which is very close to the 0.061% derived in table 1 from Fama's figures, despite the slight differences in data sources. Furthermore, the *t*-statistic I obtain for the difference is −5.58, which is very close to the −5.66 reported by Fama. Fama's results are therefore not due to some peculiarity of the CRSP data bank.

3. Operational significance of Fama's findings

A term premium means that investors can obtain a higher expected return in one maturity than another. That they nevertheless hold all the outstanding securities of all maturities indicates that at least the marginal investors have some sort of preference for the maturities with the lower expected returns that compensates them indirectly for the lower average observed return.

In the present context, Fama's evidence suggests that investors have some sort of preference for 10-month bills relative to 9-month bills that prevents them from replacing their 10-month bills with 9-month bills, in spite of the higher expected return on the latter. In order to perform this substitution, the owner of a bill who purchased it new when it had 12 months to run would have to sell it when it had 10 months to go and use the proceeds to buy a

9-month bill. At the end of the month, in order to be in the same position as an investor who did not perform this substitution, the investor would have to sell the 9-month bill (which now has 8 months to go) and replace it with the original 10-month bill (which now has only 9 months to go).

If all these transactions could be performed at the bid–asked mean price, Fama's calculation would be the appropriate one, and would demonstrate, subject to technical qualifications I will return to in section 5, that investors for some reason like 10-month bills so much that they do not make this substitution, in spite of the higher expected return it would give them.

In fact, however, investors who attempted to churn their portfolios in this manner would have to buy at the high asked price and sell at the low bid price. The relevant question, then, is whether an investor who sold 10-month bills at the bid price, purchased 9-month bills with the proceeds at the asked price, sold these after one month at the bid price, and then got back into the original bill at the asked price, would expect to end up with more than another holder of 10-month bills who did nothing.

The second column of table 3 shows the average 1-month return on 9-month bills, assuming one buys at the asked price and sells at the bid price, along with the average cost in terms of return of financing this investment by selling 10-month bills at the bid price and replacing them one month later at the asked price. When transactions costs are taken into account, the expected returns of the first column are almost reversed. The investor would actually lose 0.074% on average, rather than gaining 0.054%, by making this substitution. The t-statistic for the difference, 7.56, is even larger in absolute value than that which either Fama or I obtain using bid–asked mean prices.

Investors who do not make this substitution are therefore losing nothing in terms of expected return. There is therefore no conclusive evidence here that they have any particular preference for 10-month securities over 9-month securities, or that there is a non-monotonicity of the term premium at these maturities in any operational sense. The problem is that the value of a security is ambiguous to within its bid–asked spread. It is therefore risky to focus, as Fama does, on one particular value within this range.[3]

Conceivably a holding period other than one month could be found for which replacing a 10-month bill with a 9-month bill increased the expected return, but this is unlikely given that the transactions costs involved in the full round-trip analyzed above are more than double the apparent premium based on bid–asked means. Even if the circuit could be closed with no transactions costs at all at some other pair of maturities, the remaining half of the full round-trip transactions costs could well completely eat up the potential

[3] Published quotations often show much larger bid–asked spreads than do the 'inside market' quotations traders actually face. Ideally researchers should use these more accurate quotations, though they are more difficult to obtain. The CRSP data Fama used is based on published dealer quotations similar to those reported in the press.

premium.[4]

The pattern of asked discounts shown in table 2 would indicate that investors have a preference for the 10-month bill over the 9-month bill, if we knew that investors actually purchased 10-months bills at these rates. However, all an asking offer indicates is that dealers stand ready to sell on these terms, not that anyone ever takes them up on it. All we know is that the price implicit in the asked discount is an upper bound on the value of the security to investors, and not that it actually equals that value to any of them. If dealers are not attempting to make an active selling market in 10-month T-bills, they can shut buyers out simply by setting the asked discount sufficiently low. We know that investors are actively holding 10-month bills, since someone must be holding the outstanding volume, but there is no reason to assume that investors are actively buying 10-month bills at the published asking rates.

The magnitude of the transaction cost correction in table 3 is so great that it calls into question other results based on bid–asked mean prices, in particular the rise in the premium at the shortest maturities. In McCulloch (1975, p. 109), however, it is demonstrated that the rising term premium at shorter maturities *is* unambiguous even after allowing for transactions costs, using similar data for 1951–66.[5]

4. Why 9 months?

It is of course interesting to investigate why the bid and asked prices behave as peculiarly as they do around the 9-month maturity, and in particular during the period 8/64–12/72. Fama's results are possible because, as already shown in table 1, the bid–asked spread is much smaller during this period for 9-month bills than for adjacent maturities, and because it is the asked discount that happens to rise to meet the bid discount. If the bid and the asked discounts both moved together symmetrically, there would be little if any abnormal behavior of the bid–asked mean, despite the lower transactions costs at 9 months.

Why, then, are there such small transactions costs for 9-month bills relative to 8- and 10-month bills? Roll (1970, pp. 52–61) notes that 13- and 26-week bills similarly have much lower bid–asked spreads than adjacent maturities, and attributes this to the fact that the Treasury injects large quantities of bills at these particular maturities. These maturities therefore necessarily have large

[4] Note that our calculations do not use the 8- or 10-month asking quotations. They are based on the 8- and 10-month bid discounts, which are nearly in line with the 9-month bid discount, along with the 9-month asked discount. Thus it is the *small* spread at 9 months that eats up the profit, not the much larger spreads at 8 and 10 months.

[5] Although my primary estimates of the term premium in that paper were based on bid–asked mean prices, they would not have picked up Fama's anomaly even if it were present in the 1947–1966 period I studied, since they were based on a quadratic spline curve-fitting method rather than directly on the individual security prices. However, one of the virtues of curve-fitting is precisely that it eliminates such security-specific anomalies. Another advantage is that it enables one to bring an exhaustive data base to bear on preset maturities.

trading volumes and lower transactions costs if, as it is reasonable to assume, dealers have positive economies of scale. Today, the Treasury conducts 13-, 26-, and 52-week auctions, but not 39-week auctions. Why then should anything be special about the ninth month?

The answer is that during the bulk of the only period in which Fama found his anomaly, namely 8/64 to 12/72, the Treasury *did* conduct 9-month bill auctions. Beginning in September 1966, the Treasury auctioned additional quantities of the 1-year bills it had issued three months earlier, for the express purpose of increasing the marketability of the 1-year issues, which matured on the last day of the month and therefore did not, except by accident, eventually become perfect substitutes for the 13- and 26-week bills. The supplemental issues at 9 months were approximately 50% of the initial issue at one year, and were therefore substantial. (See *Treasury Bulletin*, October 1966 and subsequent issues.) Starting with the bill maturing August 28, 1973, however, it instead shifted the maturity of its new 1-year bills to the same Tuesday cycle as its 13- and 26-week bills, and therefore discontinued the 9-month auctions after October 1972. (See *Treasury Bulletin*, September 1972.)[6]

5. Serial correlation

The numbers in square brackets in table 3 are normalized von Neumann ratios which, under the null hypothesis of no serial correlation in the residuals, have zero mean, unit standard deviation, and are asymptotically normal.[7] They indicate overwhelmingly significant positive serial correlation for both methods of calculating the difference in returns. Since the excess returns are positively correlated, the t-statistics reported are biased toward rejecting the null hypothesis that the unconditional difference is zero. [See, e.g., Kmenta (1971, pp. 278–282).] Even apart from transactions costs, therefore, Fama's evidence of nonmonotonicity is severely flawed.

Each month's excess return observation can be interpreted as a combination of a term premium, a transactions cost, and a pure forecasting error. With zero transactions costs and a constant term premium, such serial correlation would indicate that the market's forecasting errors are serially correlated and therefore inefficient. However, we have demonstrated that transactions costs are

[6] Roll (1970, p. 99) likewise rejects monotonicity, but in the vicinity of the 13-week bill, using bid–asked mean prices and calculations similar to Fama's. In McCulloch (1973, pp. 52–53), I demonstrated that transactions costs similarly deprive this non-monotonicity of any operational significance.

[7] The normalized von Neumann ratio used here,

$$NvNR = \left[2 - \frac{\Sigma(\varepsilon_i - \varepsilon_{i-1})^2}{\Sigma(\varepsilon_i)^2}\right]\left[\frac{(n-1)(n+1)}{n-2}\right]^{1/2},$$

is based on the traditional von Neumann ratio tabulated by Hart (1942). By Hart's tabulation, the normal (0,1) approximation is already excellent when $n = 60$. The sign has been reversed so that a positive value indicates positive serial correlation.

very important in this calculation. Furthermore, Fama, in his companion article (1984a), provides valid evidence that the term premium exhibits substantial time variation. It is therefore not necessary to conclude that the market's forecasts are inefficient. Why such time variation should occur is an interesting issue in its own right, but one which is beyond the scope of this comment.

6. Conclusion

The evidence Fama reports on the term premium, taken together with the caveats raised in this note, is consistent with the hypothesis that the term premium is a monotonically increasing function of term to maturity. The apparent reversal at 9 months is just an artifact of the behavior of transactions costs during the period of 9-month bill auctions. His estimated premia near 5 months are almost equal to the average of his 9- and 10-month premia, and are therefore the values that would best fit the critical 9-month bill quotations. The monotonic premium that would best fit his data would therefore rise to approximately 5 months and have no important further change. His point estimates generally show a peak at 9 months and a decline thereafter, reaching negative values beyond 48 months, but neither the rise beyond approximately 5 months nor the decline and ultimately negative values are significant. The premium may continue to rise slightly beyond 5 months, but the noise at these maturities is too great to demonstrate the presence of such a rise, if it is indeed present.[8]

References

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

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

Hart, B.I., 1942, Significance levels for the ratio of the mean square successive difference to the variance, Annals of Mathematical Statistics 13, 445–447.

Kessel, Reuben A., 1965, The cyclical behavior of the term structure of interest rates, National Bureau of Economic Research occasional paper no. 91.

Kmenta, Jan, 1971, Elements of econometrics (Macmillan, New York).

McCulloch, J. Huston, 1973, An estimate of the liquidity premium, Unpublished Ph.D. dissertation (Department of Economics, University of Chicago, Chicago, IL).

McCulloch, J. Huston, 1975, An estimate of the liquidity premium, Journal of Political Economy 83, 95–119.

McCulloch, J. Huston, Interest-risk sensitive deposit insurance premia: Stable ACH estimates, Journal of Banking and Finance 9, 137–156.

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

[8] In McCulloch (1985, p. 146), I do find negative and marginally significant (at conventional levels with no Bonferroni adjustment) term premia for 5- and 10-year maturities using a more efficient Paretian stable adaptive conditional heteroskedastic (ACH) procedure and data spanning 1951–1982, fit with a cubic spline to reduce idiosyncrasies arising from transactions costs. Nevertheless, I am willing to believe that this period simply had an unusual run of unanticipated increases in interest rates and that investors did not really anticipate their low average returns.