NBER WORKING PAPER SERIES

MONETARY POLICY AND THE TERM STRUCTURE OF INTEREST RATES

Bennett T. McCallum

Working Paper No. 4938

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1994

For helpful comments on an earlier draft, the author is indebted to John Campbell, Tim Cook, Spencer Dale, Eugene Fama, Kenneth Froot, Marvin Goodfriend, Charles Goodheart, Greg Mankiw, Allan Meltzer, Danny Quah, Tony Smith and Julian Wiseman. This paper is part of NBER's programs in Economic Fluctuations and Monetary Economics. Any opinions expressed are those of the author and not of the National Bureau of Economic Research.

© 1994 by Bennett T. McCallum. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

NBER Working Paper #4938
November 1994

MONETARY POLICY AND THE TERM STRUCTURE OF INTEREST RATES

ABSTRACT

This paper addresses a prominent empirical failure of the expectations theory of the term structure of interest rates under the assumption of rational expectations. This failure concerns the magnitude of slope coefficients in regressions of short rate (or long-rate) changes on long-short spreads. It is shown that the anomalous empirical findings can be rationalized with the expectations theory by recognition of an exogenous random (but possibly autoregressive) term premium plus the assumption that monetary policy involves smoothing of an interest rate instrument -- the short rate -- together with the responses to the prevailing level of the spread.

Bennett T. McCallum
Graduate School of Industrial Administration
Carnegie Mellon University
Pittsburgh, PA 15213-3890
and NBER

I. Introduction

In a recent paper, McCallum (1994) argues that a prominent empirical puzzle involving exchange rate behavior--namely, the drastic apparent failure of uncovered interest parity--can be rationalized as a consequence of systematic monetary policy behavior that has been ignored in most previous studies.[1] Here it will be argued that a similar result is applicable in the context of an apparent failure of the expectations theory of the term structure of interest rates. In particular, the failure of short-rate (and long-rate) changes to be related as predicted to prevailing long-short spreads is shown to be a plausible consequence of monetary policy behavior that features interest rate smoothing in combination with policy responses to movements in the long-short spread.[2] This explanation is entirely consistent with, but more general and more fully developed than, the one proposed in a notable study by Mankiw and Miron (1986).[3]

The paper's organization is as follows. In Section II, the term-structure puzzle is reviewed and the paper's rationalization is developed for the simplest two-period case. Then in Section III, the analysis is extended to long rates of greater maturity. Additional evidence is developed in Section IV and concluding remarks appear in Section V.

II. Two-Period Case

In this section we consider the issue and our proposed explanation for the two-period case, i.e., for the relationship between yields on one-period and two-period bonds, denoted r_t and R_t respectively. Assuming that the securities in question are discount bonds, the expectations theory of the term structure posits that the "long" rate R_t is related to r_t and the expected future short rate $E_t r_{t+1}$ as follows:[4]

(1) $R_t = 0.5(r_t + E_t r_{t+1}) + \xi_t$.

Here $E_t r_{t+1} = E(r_{t+1}|\Omega_t)$ with $\Omega_t = \{r_t, r_{t+1}, \ldots, R_t, R_{t-1}, \ldots\}$ so we are assuming rational expectations. The term ξ_t is a "term premium" that is often assumed constant. Defining the expectational error $\varepsilon_{t+1} = r_{t+1} - E_t r_{t+1}$, equation (1) implies

(2) $\quad 1/2\ (r_{t+1} - r_t) = (R_t - r_t) - \xi_t + 1/2\ \varepsilon_{t+1}.$

Then if ξ_t is assumed constant, $\xi_t = \xi$, the orthogonality of ε_{t+1} with R_t and r_t implies that the slope coefficient β in a regression of the form

(3) $\quad 1/2\ (r_t - r_{t-1}) = \alpha + \beta(R_{t-1} - r_{t-1}) +$ disturbance,

should have a probability limit of 1.0. An estimated value significantly different from 1.0 is inconsistent either with the expectations theory or one of the maintained hypotheses.

In fact, it has been documented by many researchers that slope coefficients tend to be well below 1.0 in post-1914 data for the United States, often significantly so in terms of estimated standard errors. Point estimates obtained in a number of studies are reported in Table 1. There we see that the slope coefficient values are all well below 1.0, with the exception of Mankiw and Miron's value for 1890-1914 and Campbell and Shiller's final value.[5] The former, which pertains to observations taken before the founding of the Federal Reserve, will be discussed in Section IV. The latter is accompanied by a rather large standard error and pertains to an exceedingly long short rate.[6]

One possible explanation for these findings is, of course, that the expectations theory is simply untrue--but the quantitative extent of the discrepancy seems surprisingly large. Another possibility is invalidity of the rational expectations (RE) hypothesis,[7] but it seems unlikely that the same general type of systematic expectational error would prevail over different sample periods. In any event, our proposed explanation is that ξ_t is not constant--i.e., that there is a variable term premium--and that monetary policy is conducted in a manner to be explained momentarily. The

Table 1

Empirical Results, Two-Period Case

Study	Sample Period	Short Rate	Slope Coefficient
Mankiw & Miron (1986)	1959 - 1979	3 mo.	0.23
" " " "	1951 - 1958	"	-0.33
" " " "	1934 - 1951	"	-0.25
" " " "	1915 - 1933	"	0.42
" " " "	1890 - 1914	"	0.76
Evans & Lewis (1994)	1964 - 1988	1 mo.	0.42
Campbell & Shiller (1991)	1952 - 1987	1 mo.	0.50
" " "	" "	2 mo.	0.19
" " "	" "	3 mo.	-0.15
" " "	" "	6 mo.	0.04
" " "	" "	12 mo.	-0.02
" " "	" "	24 mo.	0.14
" " "	" "	60 mo.	2.79
Fama (1984)	1959 - 1982	1 mo.	0.46
Roberds, Runkle & Whiteman (1993)	1984 - 1991	3 mo.	-0.01
"	1979 - 1982	3 mo.	0.19
"	1975 - 1979	3 mo.	0.43

process generating ξ_t is assumed to be covariance stationary but not necessarily white noise. For specificity, the ξ_t process will be taken to be autoregressive of order one [AR (1)]:

(4) $\xi_t = \rho\xi_{t-1} + u_t$

Here u_t is white noise and $|\rho|<1.0$. To this writer it seems implausible that there would not be <u>some</u> period-to-period variability in the discrepancy term ξ_t in (1), a term that reflects changes in tastes regarding the need for financial flexibility and a myriad of other disturbing influences, none major enough to justify separate recognition. In any event, it is not the case that the inclusion of a random ξ_t disturbance in (1) converts the expectations theory into a tautology. That would be the case if ξ_t were related to r_t, $E_t r_{t+1}$, and R_t as in (1) without restriction. But instead the present assumption is that ξ_t is exogenous with respect to r_t and R_t. This reflects the idea that the expected one-period holding yields on one-period and two-period bonds are equal up to a constant plus a random disturbance term; that these yields differ from that constant only randomly. This is, for the case at hand, the essence of the expectations theory.

Regarding monetary policy, our hypothesis begins with the observation that actual policy behavior in the U.S. (and many other nations) involves manipulation of a short-term interest rate "instrument" or "operating variable." Specifically, we assume that[8]

(5) $r_t = \sigma r_{t-1} + \lambda(R_t - r_t) + \zeta_t$

where $\sigma \geq 0$ is presumed to be close to 1.0 and $\lambda \geq 0$ to be smaller than 2.[9] Thus there is a considerable element of interest rate "smoothing" -- keeping r_t close to r_{t-1} -- and also a tendency to tighten policy (by raising r_t) whenever the spread $R_t - r_t$ is higher than normal. Whether this reaction to $R_t - r_t$ occurs because the central bank views it as a good predictor of future output growth or as a good indicator of recent policy laxity does not matter for current purposes. The final term ζ_t reflects other components of policy behavior. It would not impair our analysis to let ζ_t be autocorrelated, but it would not help, either. Accordingly, we shall assume that ζ_t is white noise.

It may be helpful to briefly consider the rationale for the specification of policy behavior in (5). Regarding the r_{t-1} term, there exists some controversy regarding the reason behind central banks' proclivity for interest rate smoothing--and, indeed, for their use of interest rate instruments. But there is virtually no disagreement with the proposition that the Fed--and other major central banks--have in fact employed such practices during most (if not all) of the last 40 years.[10] (For some useful discussion, see Goodfriend (1991) and Poole (1991).) In addition (5) reflects the assumption that the central bank tends to tighten policy when the spread R_t-r_t is large. One possible rationalization is that the spread is an indicator of monetary policy expansiveness, as suggested by Laurent (1988), so that an unusually high value indicates the need for corrective action. A different idea is that the spread provides an indicator of the state of the economy from a cyclical perspective. Various investigators, including Estrella and Hardouvelis (1991) and Hu (1993), have documented that spread measures have predictive value for future real GNP growth rates. Also, Mishkin (1990) has shown that a spread variable has some predictive content for future inflation rates. Thus an attempt by the central bank to conduct a forward-looking countercyclical policy would call for a response of the type indicated in (5), i.e., a tightening when $R_t - r_t$ is high.[11] Admittedly, in actual practice the Fed has used other predictor variables instead of the spread. But to the extent that these and the spread are useful predictors, the policy response would be much the same as implied by (5).

Relations (1) and (5) constitute only a portion, of course, of a macroeconomic system. But if we assume that the disturbances ξ_t and ζ_t are independent of those in the remaining relations, the system will be recursive and the subsystem (1)(5) will determine r_t and R_t without reference to the other variables or shocks. Whether the remainder of the model does or does not feature relations of the IS-LM type is irrelevant, for example, as is the extent to which prices of goods are flexible. Let us consider, then, a rational expectations solution to the system (1)(5).[12]

Presuming that attention is to be focused on the fundamental or bubble-free solution yielded by the minimal-state-variable (MSV) criterion discussed by McCallum (1983), we combine (1) and (5) to yield

$$(6) \qquad (1 + \lambda)\ r_t = \sigma r_{t-1} + \lambda\ [1/2\ (r_t + E_t r_{t+1}) + \xi_t] + \zeta_t$$

and seek values of the undetermined coefficients ϕ_0, ϕ_1, ϕ_2, and ϕ_3 that will provide a r_t solution of the form

$$(7) \qquad r_t = \phi_0 + \phi_1 r_{t-1} + \phi_2 \xi_t + \phi_3 \zeta_t.$$

Clearly, the latter implies that $E_t r_{t+1} = \phi_0 + \phi_1\ (\phi_0 + \phi_1 r_{t-1} + \phi_2 \xi_t + \phi_3 \zeta_t) + \phi_2 \rho \xi_t$ so we substitute these into (6) to obtain

$$(8) \qquad (1 + \lambda)[\phi_0 + \phi_1 r_{t-1} + \phi_2 \xi_t + \phi_3 \zeta_t] = \sigma r_{t-1} + \lambda\ [1/2(\phi_0 + \phi_1 r_{t-1} + \phi_2 \xi_t + \phi_3 \zeta_t) + 1/2\ (\phi_0 + \phi_1(\phi_0 + \phi_1 r_{t-1} + \phi_2 \xi_t + \phi_3 \zeta_t) + \phi_2 \rho \xi_t) + \xi_t] + \zeta_t.$$

Thus for (7) to be a solution -- to hold for all ξ_t, ζ_t realizations -- it must be true that:

$$(9) \qquad \begin{aligned} (1 + \lambda)\ \phi_0 &= \lambda \phi_0 + 1/2\ \lambda \phi_1 \phi_0 \\ (1 + \lambda)\ \phi_1 &= \sigma + 1/2\ \lambda \phi_1 + 1/2\ \lambda \phi_1^2 \\ (1 + \lambda)\ \phi_2 &= 1/2\ \lambda \phi_2 + 1/2\ \lambda \phi_1 \phi_2 + 1/2\ \lambda \rho \phi_2 + \lambda \\ (1 + \lambda)\ \phi_3 &= 1/2\ \lambda \phi_3 + 1/2\ \lambda \phi_1 \phi_3 + 1 \end{aligned}$$

The second of these is satisfied by two values of ϕ_1, namely,

$$(10) \qquad \phi_1 = \frac{(1 + \lambda/2) \pm [(1 + \lambda/2)^2 - 2\lambda\sigma]^{1/2}}{\lambda},$$

but the MSV criterion implies that the one with the minus sign is relevant.[13] Then the remaining coefficients are straightforwardly given by the other three equalities in (9).

In analyzing the implications of this solution it will be useful to emphasize the important special case involving $\sigma = 1$, which is the value suggested by interest rate smoothing behavior. When $\sigma = 1$, the MSV solution for ϕ_1 becomes $[(1 + \lambda/2) - (1 - \lambda/2)]/\lambda = \lambda/\lambda = 1$ and the other three equalities in (9) are simplified considerably. They yield $\phi_0 = 0$, $\phi_2 = \lambda/(1 - \rho\lambda/2)$, and $\phi_3 = 1$ so the solution for r_t is

(11) $$r_t = r_{t-1} + \frac{\lambda}{(1 - \rho\lambda/2)} \xi_t + \zeta_t.$$

Furthermore, $E_t r_{t+1} - r_t = \phi_2 \rho \xi_t$ so we find that the spread obeys

(12) $$R_t - r_t = 1/2\ (E_t r_{t+1} - r_t) + \xi_t = (1 - \rho\lambda/2)^{-1} \xi_t.$$

Finally, equations (11) and (4) imply

(13) $$r_t - r_{t-1} = \frac{\lambda\rho}{1 - \lambda\rho/2} \xi_{t-1} + \frac{\lambda}{1 - \lambda\rho/2} u_t + \zeta_t$$

so we can combine (12) and (13) to obtain

(14) $$1/2(r_t - r_{t-1}) = \frac{\lambda\rho}{2} (R_{t-1} - r_{t-1}) + \frac{\lambda/2}{1 - \rho\lambda/2} u_t + 1/2\zeta_t.$$

But here u_t and ζ_t are uncorrelated with $R_{t-1} - r_{t-1}$, so (14) represents a population version of the regression described in (3). Thus the slope coefficient in (3) is a consistent estimator of $\rho\lambda/2$, so the analyst should anticipate a slope well below 1.0. Indeed, if ξ_t were white noise, with $\rho = 0$, a slope coefficient of zero would be implied -- even though relation (1) is the main behavioral relation of the system. That result demonstrates, I would suggest, not only that the usual regression test is inappropriate but also that it is misleading to think of the expectations theory in terms of the "predictive content" of the spread for future changes of the short rate. Such predictive content is not a necessary implication of that theory.

In addition, a zero slope coefficient would be implied if $\lambda = 0$, i.e., if the central bank did not respond to the current value of the spread but simply set r_t equal to r_{t-1} (plus, perhaps, ζ_t). This special case, of the special case with $\sigma = 1$, represents the hypothesis of Mankiw and Miron (1986)--that the Federal Reserve has practiced interest rate smoothing and thereby induced short rates to approximate a random walk process in their behavior. Our result strongly supports the general idea of the Mankiw and Miron hypothesis, but shows that it holds even if r_t behavior is not that of a random walk.

A few readers have remarked that (14) appears to be inconsistent with the fact that a regression of form (3) should yield a slope coefficient of 1.0 in the special case in which the term premium ξ_t is a constant. But with $\sigma = 1.0$ in (5), a constant ζ_t implies that $R_t - r_t$ is also constant--see equation (12). Thus there is a degenerate regressor, in this case, so the regression cannot be conducted. And in the case with $\sigma < 1.0$, (14) does not apply so again there is actually no inconsistency.

Let us now briefly consider the situation with $\sigma < 1$. In such cases we would need to include a non-zero constant term in (5) to permit a stationary equilibrium with $E\zeta_t = 0$. The solution in this case yields a relationship analogous to (14) that is less tidy than the latter, and includes additional predetermined variables. But it remains true that the probability limit of the slope coefficient in a regression of $r_t - r_{t-1}$ on $R_{t-1} - r_{t-1}$ is not in general equal to 1.0 and is most likely to be smaller than 1.0; a demonstration is provided in Appendix A. Accordingly, the same general message applies as in the more tractable case with $\sigma = 1$. That message is that the realization of (say) a positive value of ξ_t will drive up R_t relative to r_t via (1). But then $R_t - r_t$ will be negatively correlated with the composite disturbances $-\xi_{t-1} + 0.5\varepsilon_{t+1}$ in (3), implying that least-squares estimation of (4) will yield a slope coefficient that has a probability limit not equal to 1.0.

III. N-Period Case

Now we turn to the more interesting case in which the long rate, R_t, is for a bond with a maturity of more than two periods. In this case an approximation to the expectations-hypothesis relationship between R_t and r_t can be written as

(15) $R_t - N\, E_t(R_{t+1} - R_t) = r_t + \xi_t,$

where N+1 is a measure of the duration of the long rate.[14] In (5) the left-hand side is an approximation to the one-period holding return on the long-rate bond, the inexactness arising because the term R_{t+1} should pertain to a maturity one period less than that for R_t. Thus for many-period maturities, the approximation should be adequate.[15]

In this case the apparent empirical failure to be explained arises from writing (15) as

(16) $N\,(R_{t+1} - R_t) = (R_t - r_t) - \xi_t + N\varepsilon_{t+1},$

where $\varepsilon_{t+1} = R_{t+1} - E_tR_{t+1}$ is an expectational error that with RE is uncorrelated with R_t and r_t. Thus if ξ_t were constant, the slope coefficient in a regression of $N(R_{t+1} - R_t)$ on $R_t - r_t$ should have a probability limit of 1.0 according to the expectations theory. But such regressions again actually yield slopes well below 1.0 with U.S. data. Indeed, the values reported by Evans and Lewis (1994) and Campbell and Shiller (1991) are predominantly negative, as is documented in Table 2, and increase in absolute value with N.

As in the last section we assume that the policy reaction equation (5) obtains with $\lambda<1/N$ and that $\xi_t = \rho\xi_{t-1} + u_t$.[16] Then one can combine (5) and (15) to obtain

(17) $(1 + N)\, R_t = NE_tR_{t+1} + (1 + \lambda)^{-1}\, [\sigma r_{t-1} + \lambda R_t + \zeta_t] + \xi_t.$

The MSV solution will be of the form

(18) $R_t = \pi_1 r_{t-1} + \pi_2\xi_t + \pi_3\zeta_t$

Table 2

Empirical Results, N Period Case

Study	Sample Period	Short Rate	N+1	Slope Coefficient
Evans & Lewis (1994)	1964 - 1988	1 mo	2	-0.17
" " "	" "	"	4	-0.70
" " "	" "	"	6	-1.27
" " "	" "	"	8	-1.52
" " "	" "	"	10	-1.89
Campbell & Shiller (1991)	1952 - 1987	1 mo	2	0.00
" " "	" "	"	4	-0.44
" " "	" "	"	6	-1.03
" " "	" "	"	12	-1.38
" " "	" "	"	24	-1.81
" " "	" "	"	48	-2.66
" " "	" "	"	60	-3.10
" " "	" "	"	120	-5.02
Hardouvelis (1994)	1954 - 1992	3 mo	120	-2.90

implying $E_tR_{t+1} = \pi_1(1 + \lambda)^{-1}\ [\sigma r_{t-1} + \lambda(\pi_1 r_{t-1} + \pi_2\xi_t + \pi_3\zeta_t) + \zeta_t] + \pi_2\rho\xi_t$, which can be substituted with (18) into (17) to give

$$(19) \qquad (1 + N)\ [\pi_1 r_{t-1} + \pi_2\xi_t + \pi_3\zeta_t] = N\pi_1(1 + \lambda)^{-1}\ [\sigma r_{t-1} + \lambda(\pi_1 r_{t-1} + \pi_2\xi_t + \pi_3\zeta_t) + \zeta_t] + N\pi_2\rho\xi_t + (1+\lambda)^{-1}[\sigma r_{t-1} + \lambda(\pi_1 r_{t-1} + \pi_2\xi_t + \pi_3\zeta_t) + \zeta_t] + \xi_t.$$

For (18) to be a solution, then, we must have

$$(20) \qquad \begin{aligned} (1 + N)\ \pi_1 &= N\pi_1(1 + \lambda)^{-1}\ (\sigma + \lambda\pi_1) + (1 + \lambda)^{-1}\ (\sigma + \lambda\pi_1) \\ (1 + N)\ \pi_2 &= N\pi_1(1 + \lambda)^{-1}\ \lambda\pi_2 + N\pi_2\rho + (1 + \lambda)^{-1}\ \lambda\pi_2 + 1 \\ (1 + N)\ \pi_3 &= N\pi_1(1 + \lambda)^{-1}\ (\lambda\pi_3 + 1) + (1 + \lambda)^{-1}\ (\lambda\pi_3 + 1) \end{aligned}$$

The first of these amounts to $(1 + \lambda)\ (1 + N)\pi_1 = (N\pi_1 + 1)\ (\sigma + \lambda\pi_1)$ so we have

$$(21) \qquad \pi_1 = \frac{[(1 + \lambda)(1 + N)-\lambda-N\sigma] \pm \{[(1 + \lambda)(1 + N)-\lambda-N\sigma]^2 - 4N\lambda\sigma\}^{1/2}}{2N\lambda}$$

The term in square brackets will be positive, so the MSV solution for π_1 is the expression in (21) with the minus sign.[17] Given this value, the second and third of equations (20) determine π_2 and π_3.

To facilitate analysis, let us again focus attention on the case with $\sigma = 1$. Then we have $[(1 + \lambda)(1 + N) - (\lambda + N)]^2 = (1 + \lambda)^2(1 + N)^2 - 2(1 + \lambda)(1 + N)(\lambda + N) + (\lambda + N)^2 = 1 + 2N\lambda + N^2\lambda^2$ and the term inside curly brackets in (21) becomes $1 - 2N\lambda + N^2\lambda^2 = (1 - N\lambda)^2$. Consequently, we have $\pi_1 = [(1 + N\lambda) - (1 - N\lambda)]/2N\lambda = 1$, where the last calculation utilizes the assumed condition $1 - N\lambda > 0$. Then with $\pi_1 = 1$, the final equation in (20) implies $\pi_3 = 1$ and $\pi_2 = (1 + \lambda)/[1 + N - N\rho(1 - \lambda)]$. Because $1 > N\lambda$, π_2 is strictly positive. Given these values, we readily see that

$$(22) \qquad R_t = r_{t-1} + \frac{1 + \lambda}{1 + N - N\rho(1 + \lambda)}\ \xi_t + \zeta_t$$

$$(23) \qquad r_t = r_{t-1} + \frac{\lambda}{1 + N - N\rho(1 + \lambda)}\ \xi_t + \zeta_t.$$

Accordingly, the spread variable obeys

$$(24) \qquad R_t - r_t = \frac{1}{1 + N - N\rho(1+\lambda)} \xi_t$$

and using (22) and (4) we also have

$$(25) \qquad R_t - R_{t-1} = \frac{\lambda + 1}{1 + N - N\rho(1+\lambda)} \xi_t - \frac{1}{1 + N - N\rho(1+\lambda)} \xi_{t-1} + \zeta_t$$

$$= \frac{(\lambda\rho + \rho - 1)\xi_{t-1} + (1+\lambda)u_t}{1 + N(1 - \rho(1+\lambda))} + \zeta_t$$

$$= (\lambda\rho + \rho - 1) \; (R_{t-1} - r_{t-1}) + \frac{(1+\lambda)}{1 + N(1 - \rho(1+\lambda))} u_t + \zeta_t.$$

Consequently, we see that a regression of $N(R_t - R_{t-1})$ on $R_{t-1} - r_{t-1}$ will have a slope coefficient whose probability limit is $N(\lambda\rho+\rho-1)$ or $-N(1-\rho(1+\lambda))$. Clearly, the latter will be negative except for very large values of ρ and/or λ, and will be larger in absolute value (for a given ρ) with longer maturities (larger N).[18] In qualitative terms, both of these characteristics match the results of Evans and Lewis (1994) and Campbell and Shiller (1991) reported in Table 2.

IV. Additional Evidence

The paper by Campbell and Shiller (1991) concludes with an attempt to provide a summary characterization of term structure behavior that would be consistent with their battery of empirical findings, which include many more than those reported here. In their words, "The explanations we will consider are not finance-theoretic models of time-varying risk premia, but simply econometric descriptions of ways in which the expectations theory might fail" (1991, p. 510). In terms of the notation of the present paper, the two summary characterizations considered are (for the two-period case)

$$(26) \qquad R_t - r_t = 0.5\, E_t(r_{t+1} - r_t) + c + v_t,$$

where v_t is added noise that is orthogonal to $E_t r_{t+1} - r_t$, and

$$(27) \qquad R_t - r_t = k0.5E_t(r_{t+1} - r_t) + c$$

where k>1. The latter "could be described as an overreaction model of the yield spread," according to Campbell and Shiller (1991, p. 513). They explore the implications of these two summary characterizations of ways in which the expectations theory might fail and conclude that (27) is consistent with the data but that (26) is not.

Let us consider how these characterizations compare with the explanation of the present paper. Looking back at Section II, we see that equation (12) is of a similar form to that of (26), but with the crucial difference that ξ_t in (12) is <u>not</u> orthogonal to $E_t r_{t+1} - r_t$. Thus the inadequacy of (26) does not serve to discredit the model of Section II. Furthermore, using the expression $E_t r_{t+1} - r_t = \phi_2 \rho \xi_t$ to eliminate ξ_t from (12) results in

(28) $$R_t - r_t = (1/\rho\lambda)\ E_t(r_{t+1} - r_t)$$

for the model of Section II. But with $0<\lambda<2$ and $|\rho|<1$, (28) implies that $k>1$ in (27) if ρ is positive. So Campbell and Shiller's summary characterization is consistent with the present paper's rationalization.[19]

It was mentioned above that the slope coefficient reported in Table 1 for the years 1890 - 1914 was closer (than for more recent periods) to the value of 1.0 that has been focused on in previous investigations. As Mankiw and Miron (1986) emphasize, those years precede the founding of the Federal Reserve System and therefore pertain to a period during which interest rate smoothing behavior would be absent. In a similar vein, Kugler (1988, 1990) finds that slope coefficients are closer to 1.0 for Germany and Switzerland than for the United States during recent years. This result he attributes to a smaller degree of interest smoothing behavior by the Bundesbank and the Swiss National Bank, in comparison with the Fed, a hypothesized behavioral difference that is consistent with the beliefs of many students of central banking behavior. Since the model in Sections II and III presumes a substantial degree of interest rate smoothing, this paper's explanation is consistent with both of these findings.[20]

V. Concluding Remarks

The discussion of the foregoing paragraph suggests that one possible way of conducting additional tests of this paper's hypothesis would be to consider different monetary policy regimes corresponding to different time periods for the United States and to different nations. Reaction functions corresponding to (5) would be estimated and the implications of their parameter values for the crucial slope coefficients then compared with values of the coefficients obtained for these different regimes. Now, it may prove possible to make some progress toward execution of such a study. There is, however, a substantial difficulty that needs to be mentioned. Specifically, it is the case that actual central banks do not respond only to term spreads in deciding upon changes in r_t. Thus equation (5) represents a simplification relative to actual behavior of the Fed, which almost certainly responds to recent inflation and output or employment movements as well as the spread. So, if one were to attempt to econometrically estimate actual reaction functions, then measures of inflation and output gaps would need to be included. But in that case values of these variables would need to be explained endogenously, so the system of equations in the model would have to be expanded. Furthermore, the dynamic behavior of inflation and output would need to be modeled "correctly," which is an exceedingly difficult task given the absence of professional agreement about short-run macroeconomic dynamics. In short, this type of study would require specification and estimation of a complete dynamic macroeconometric model.

In light of the foregoing discussion it will be seen that, because of the simplified nature of our policy equation (5), this paper's proposed explanation might be regarded as more of a parable than a fully-worked-out quantitative model. I would argue, however, that this is not a source of embarrassment, for most knowledge in economics is actually of the parable type.[21] The relevant issue is whether a proposed parable is fruitful in understanding important economic phenomena. In this particular case the proposed parable suggests that slope estimates in regressions of the form (3) or (16) differ from 1.0 despite the validity of a version of the expectations theory of the term structure. This version permits the holding-period yields on securities of various maturities to differ by a random discrepancy that is exogenous but perhaps serially correlated. The basic idea of the parable is that the estimated slope coefficient is a composite parameter reflecting policy behavior as well as the behavior of market participants, with the type of policy postulated involving interest rate smoothing and response to the long-short spread, the latter reflecting important aspects of the state of the economy. The fact that essentially the same parable can rationalize a major anomaly in foreign exchange markets must be regarded as a significant mark in its favor.

References

Campbell, John Y., and Robert J. Shiller, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," Review of Economic Studies 58 (1991), 495-514.

Canzoneri, Matthew B., Dale W. Henderson, and Kenneth S. Rogoff, "The Information Content of the Interest Rate and Optimal Monetary Policy," Quarterly Journal of Economics 98 (November 1983), 545-566.

Cook, Timothy, and Thomas Hahn, "Interest Rate Expectations and the Slope of the Money Market Yield Curve," Federal Reserve Bank of Richmond Economic Review 76 (No. 5, 1990), 3-26.

__________ and __________, "The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s," Journal of Montary Economics 24 (November 1989), 331-351.

Dotsey, Michael, and Robert G. King, "Monetary Policy Instruments and Policy Rules in a Rational Expectations Environment," Journal of Monetary Economics 12 (September 1983), 357-382.

Estrella, Arturo, and Gikas A. Hardouvelis, "The Term Structure as a Predictor of Real Economic Activity," Journal of Finance 46 (June 1991), 555-576.

Evans, Martin D.D., and Karen K. Lewis, "Do Stationary Risk Premia Explain It All? Evidence from the Term Structure," Journal of Monetary Economics 33 (April 1994), 285-318.

Fama, Eugene F., "The Information in the Term Structure," Journal of Financial Economics 13 (Dec. 1984), 509-528.

Froot, Kenneth A., "New Hope for the Expectations Hypothesis of the Term Structure of Interest Rates," Journal of Finance 44 (June 1989), 283-305.

Fuhrer, Jeffrey C., and George Moore, "Monetary Policy and the Behavior of Long-Term Real Interest Rates," Finance and Economics Discussion Series 93-16, Board of Governors of the Federal Reserve System, 1993.

Goodfriend, Marvin, "Interest Rate Policy and the Inflation Scare Problem," Federal Reserve Bank of Richmond Economic Quarterly 79 (No. 1, 1993), 1-24.

________________, "Interest Rates and the Conduct of Monetary Policy," Carnegie-Rochester Conference Series on Public Policy 34 (Spring 1991), 7-30.

Hardouvelis, Gikas A., "The Term Structure Spread and Future Changes in Long and Short Rates in the G7 Countries: Is There a Puzzle?" Journal of Monetary Economics 33 (April 1994), 255-283.

Hu, Zuliu, "The Yield Curve and Real Activity," IMF Staff Papers 40 (December 1993), 781-806.

Kugler, Peter, "An Empirical Note on the Term Structure and Interest Rate Stabilization Policies," Quarterly Journal of Economics 103 (November 1988), 789-792.

____________, "The Term Structure of Euro Interest Rates and Rational Expectations," Journal of International Money and Finance 9 (1990), 234-244.

Laurent, Robert D., "An Interest Rate-Based Indicator of Monetary Policy," Federal Reserve Bank of Chicago Economic Perspectives 12 (January/February 1988), 3-14.

Mankiw, N. Gregory, and Jeffrey A. Miron, "The Changing Behavior of the Term Structure of Interest Rates," Quarterly Journal of Economics 101 (May 1986), 211-228.

McCallum, Bennett T., "Price Level Determinacy With an Interest Rate Policy Rule and Rational Expectations," Journal of Monetary Economics 8 (November 1981), 319-329.

__________, "On Non-Uniqueness in Rational Expectations Models: An Attempt at Perspective," Journal of Monetary Economics 11 (1983), 139-168.

__________, "Some Issues Concerning Interest Rate Pegging, Price Level Determinacy, and the Real Bills Doctrine," Journal of Monetary Economics 17 (January 1986), 135-160.

__________, "A Reconsideration of the Uncovered Interest Parity Relationship," Journal of Monetary Economics 33 (February 1994), 105-132.

McCulloch, J. Huston, "The Term Structure of Interest Rates: Appendix B," Handbook of Monetary Economics, ed. by B.M. Friedman and F.H. Hahn. Amsterdam: North-Holland Pub. Co., 1990.

Mishkin, Frederic S., "The Information in the Longer Maturity Term Structure about Future Inflation," Quarterly Journal of Economics 105 (December 1990), 815-828.

Poole, William, "Interest Rates and the Conduct of Monetary Policy: A Comment," Carnegie-Rochester Conference Series on Public Policy 34 (Spring 1991), 31-40.

Roberds, William, David Runkle, and Charles H. Whiteman, "Another Hole in the Ozone Layer: Changes in FOMC Operating Procedure and the Term Structure," Operating Procedures and the Conduct of Monetary Policy: Conference Proceedings, edited by Marvin Goodfriend and David H. Small. Finance and Economics Discussion Series, Federal Reserve Board, March 1993.

Rudebusch, Glenn D., "Federal Reserve Interest Rate Targeting and the Term Structure," Working Paper, May 1994.

Shiller, Robert J., "The Term Structure of Interest Rates," Handbook of Monetary Economics, ed. by B.M. Friedman and F.H. Hahn. Amsterdam: North-Holland Pub. Co., 1990.

________________, "The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure," Journal of Political Economy 87 (December 1979), 1190-1219.

Shiller, Robert J., John Y. Campbell, and Kermit L. Schoenholtz, "Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates," Brookings Papers on Economic Activity (No. 1, 1983), 173-217.

Footnotes

[1]The hypothesized form of policy behavior involves smoothing of (relative) interest rates, with these rates used as instruments, together with policy attempts to "lean against" exchange rate changes. The analysis also assumes random disturbances--from varying risk premia--to the UIP relationship. No departure from rational expectations is involved.

[2]General aspects of the failure are discussed by Cook and Hahn (1990), Campbell and Shiller (1991), and Evans and Lewis (1994), among others.

[3]Since drafting this paper I have become aware of a study with a rather similar objective by Rudebusch (1994), which is also intended to provide a generalization of the Mankiw-Miron hypothesis. The type of policy behavior assumed there is quite different, however, as instrument settings are responsive to current conditions in my setup but are determined exogenously in his. Most significantly, Rudebusch's analysis does not offer an explanation for the empirical phenomena rationalized below at the end of Section III and in Section IV.

[4]The relationship is exact, if the interest rates are based on continuous compounding, or an approximation otherwise: see Shiller (1990).

[5]An analogous result holds for the case of three-month and one-month rates; see Kugler (1988, 1990).

[6]The Roberds, Runkle, and Whiteman (1993) results are for treasury bills. This study also reports results using Federal Funds and Repo securities and finds one slope coefficient close to 1.0 for the former using the sample period 1979.10 - 1982.10.

[7]This possibility has been explored, using survey data on expectations, by Froot (1989).

[8] For values of σ less than 1.0, a constant term should also be included in (5) if $E\zeta_t = 0$. We have not shown it here, however, because the case with $\sigma = 1$ will be featured below and because little interest attaches to the constant term in any case.

[9] In what follows, $\lambda<2$ will be presumed because such a condition seems plausible and also because a theoretical issue, concerning the root of (10) that gives the bubble-free solution, arises when $\lambda>2$. (On this issue see Appendix B.) But the solutions obtained below, and most of the analysis, would continue to prevail with $\lambda\geq2$.

[10] Some analysts are dubious that the Fed's control over the one-day Federal Funds rate translates into effective control over one-month or three-month treasury bill rates that are the operational counterpart of r_t in (5). But the evidence of Cook and Hahn (1989) suggests that three-month rates do, in fact, respond within the day to policy-induced changes in the Federal Funds rate. Furthermore, if the Fed doubted its ability to control treasury bill rates it could (given its holdings) operate directly in the treasury bill markets. Consequently, doubts concerning the controllability of r_t seem to be unfounded.

[11] In an influential recent publication, Goodfriend (1993) suggests that the Fed regards (or should regard?) the long rate as an indicator of "inflation scares," behavior that might be interpreted as descriptive of a rule of the form $r_t = \delta r_{t-1} + \theta(R_t - \bar{R}) + \zeta_t$. The latter can be written in the form (5) by defining $\sigma = \delta/(1-\theta)$ and $\lambda = \theta/(1-\theta)$, but then dynamic stability (non-explosiveness) requires $\delta<1-\theta$ (assuming that $0<\theta<1$). It is not clear that Goodfriend would agree with the above formulation, however: another possibility is $r_t = r_{t-1} + \theta(R_t - R_{t-1}) + \zeta_t$. In any event, the policy behavior pattern in his article has a substantial degree of similarity with formulation (5): both call for an increase in the short rate in response to a ceteris paribus rise in the long rate.

[12] Students of the price level determinacy literature--e.g., McCallum (1981) (1986), Dotsey and King (1983), Canzoneri, Henderson, and Rogoff (1985)--will wonder about the absence of nominal variables in the system (1)(5). But the price level can be brought in by adding (e.g.) an IS-type relation in which a real rate such as $r_t - (E_t p_{t+1} - p_t)$ appears, p_t being the log of the price level. Then determinacy of p_t will require the presence of an additional term in (5), one that includes a nominal variable such as p_t or $E_t p_{t+1}$ or p_{t-1}. Algebraic analysis becomes much more difficult because the counterpart of (10) below will be a cubic in many such cases. But a cubic must have at least one real root, so in principle determinacy can be investigated. My examination of a case with p_t included in (5) indicates that determinacy would be guaranteed unless $\sigma = 1.0$ exactly. Thus for σ close to 1.0, the results would be approximately the same as those emphasized below.

[13] This is the root that yields $\phi_1 = 0$ when $\sigma = 0$, a special case in which it is clear that r_{t-1} would be an extraneous state variable [as discussed in McCallum (1983)].

[14]For pure discount bonds, N+1 is the maturity.

[15]Equation (15) can alternatively be written as $R_t = (1-\delta)\Sigma\delta^k E_t r_{t+k}$ + term premium, with the summation from 0 to ∞. Thus the approximation amounts to an infinite-maturity version of the linearization developed by Shiller (1979), with $N = \delta/(1-\delta)$. This approximation has also been used by Shiller, Campbell, and Schoenholtz (1983), Campbell and Shiller (1991), Fuhrer and Moore (1993), and Hardouvelis (1994).

[16]The condition $\lambda < 1/N$ is the counterpart of $\lambda < 2$ in the two period case (in which $N = 1$) and is again presumed but not strictly required. The larger is N, the smaller will be λ in equation (5) because only one λ can prevail, but many long rates can be considered.

[17]Again this is because with $\sigma = 0$, r_{t-1} should not appear in the solution for R_t.

[18]The policy parameter λ would be expected to be smaller for larger N. This effect reinforces the tendency for the slope coefficient to increase in absolute value with N.

[19]The foregoing discussion implies, incidentally, that there is actually nothing bizarre or irrational about a finding expressible as $k>1$ in (27).

[20]For additional discussion of the Mankiw-Miron hypothesis, see Cook and Hahn (1990).

[21]Consider the usual depiction of a production function as $y_t = f(n_t, k_t)$, where the symbols should not require definition. Can this depiction be considered anything more than a parable?

Appendix A

Here the concern is with the model of Section II when $\sigma<1.0$. From (9), we find that

(A-1) $$r_t = \phi_0 + \phi_1 r_{t-1} + \frac{\lambda}{\delta - \rho\lambda/2}\xi_t + \frac{1}{\delta}\zeta_t$$

where $\delta = 1 - (\phi_1 - 1)\lambda/2$. Then from (A-1) it follows that

(A-2) $$E_t r_{t+1} - r_t = \phi_0 + (\phi_1 - 1)r_t + \lambda\rho/(\delta - \rho\lambda/2)\xi_t$$

and thus using (12) that

(A-3) $$R_t - r_t = (1/2)[\phi_0 + (\phi_1 - 1)r_t + (\rho\lambda/(\delta - \rho\lambda/2))\xi_t] + \xi_t.$$

Now, equation (2) indicates that the plim of the slope coefficient on $R_t - r_t$ in the regression (3) will equal 1.0 minus plim $T^{-1}\xi_t(R_t - r_t)$/plim $T^{-1}(R_t - r_t)^2$. Its value will be smaller than 1.0, then, if $E\xi_t(R_t - r_t)$ is positive.

From (A-3) it is clear that there are two components to $E\xi_t(R_t - r_t)$. One of these is

(A-4) $$\left(\frac{\rho\lambda/2}{\delta - \rho\lambda/2} + 1\right) \sigma_\xi^2,$$

which is necessarily positive since the term in parentheses equals

(A-5) $$\frac{1 + (1 - \phi_1)\lambda/2}{1 + (1 - \phi_1)\lambda/2 - \rho\lambda/2}.$$

Here $(1 - \phi_1)\lambda/2$ is positive, since $\phi_1<1$ when $\sigma<1$ (see below), and $|\rho\lambda/2|<1$. Thus expression (A-5) is unambiguously positive. The second component is

(A-6) $$(1/2)(\phi_1 - 1)Er_t\xi_t,$$

in which the term $\phi_1 - 1$ is negative but will be small for σ (and ϕ_1) close to 1.0. To sign $Er_t\xi_t$, we use (A-1) and (4) as follows, assuming $E\xi_t\zeta_t = 0$:

(A-7) $$Er_t\xi_t = E[\phi_0 + \phi_1 r_{t-1} + \phi_2\xi_t + \phi_3\zeta_t]\xi_t$$
$$= \phi_1 Er_{t-1}\xi_t + \phi_2\sigma_\xi^2 = \phi_1 Er_{t-1}\rho\xi_{t-1} + \phi_2\sigma_\xi^2.$$

Then since $Er_t\xi_t = Er_{t-1}\xi_{t-1}$, we have

(A-8) $$Er_t\xi_t = \frac{\phi_2\sigma_\xi^2}{1 - \phi_1\rho}.$$

The latter is unambiguously positive since $\phi_2>0$ and $|\phi_1\rho|<1$. Thus the second component is negative but will tend to be small relative to the first.

It remains to demonstrate that $\phi_1<1$ when $\sigma<1$. But we have found that

$$\text{(A-9)} \qquad \phi_1 = \frac{(1 + \lambda/2) - [(1 + \lambda/2)^2 - 2\lambda\sigma]^{1/2}}{\lambda}$$

With $0<\sigma<1$, we have $2\lambda>2\lambda\sigma>0$ so the term in square brackets is positive and larger than $(1 - \lambda/2)^2$. Thus the value of ϕ_1 is smaller than when this term equals $(1 - \lambda/2)^2$, i.e., when $\sigma = 1$. But ϕ_1 remains non-negative because the term in brackets is smaller than $(1 + \lambda/2)^2$.

Appendix B

Here the purpose to explain the difficulty concerning the roots of (10), mentioned in footnote 9, that would obtain with $\lambda > 2$. To see the issue, consider the case in which $\sigma = 1$, so that the ϕ_1 solutions are $[(1 + \lambda/2) \pm ((1 - \lambda/2)^2)^{1/2}]/\lambda$. Denote these solutions as $\phi_1^{(1)}$ and $\phi_1^{(2)}$ and define $\phi_1^{(1)}$ as $[(1 + \lambda/2) - ((1 - \lambda/2)^2)^{1/2}]/\lambda$ for $\lambda < 2$. But then if $\lambda > 2$, is the "same" root $\phi_1^{(1)}$ equal to $[(1 + \lambda/2) - (1 - \lambda/2)] = 1$ or to $[(1 + \lambda/2) + (1 - \lambda/2)]/\lambda = 2/\lambda$? Note that the latter is suggested by the convention that $(Z^2)^{1/2}$ is a positive number regardless of the sign of Z. But defining the relevant root as $\phi_1^{(1)} = 1$ for all $\lambda > 0$ seems more appropriate than making $\phi_1^{(1)}$ equal to 1 or $2/\lambda$ depending on whether $\lambda < 2$ or $\lambda > 2$. Then equations (11) - (14) are valid even with $\lambda > 2$. (Similar considerations apply to equations (21) - (25) if we permit $N\lambda > 1$.)