Monetary Policy and the Term Structure of Interest Rates

Bennett T. McCallum

A major puzzle in financial economics is the apparent drastic inconsistency of U.S. data with the expectations theory of the term structure of interest rates.[1] As documented extensively by Campbell and Shiller (1991), both short changes in long rates and long changes in short rates fail to be related to existing long-short spreads in even approximately the manner implied by the expectations theory together with rational expectations; a convenient summary of the evidence is provided by Campbell (1995, Table 2). This failure is analogous, however, to the apparent drastic failure of uncovered interest parity in foreign exchange, which can be rationalized—it is argued by McCallum (1994)—as a consequence of monetary policy behavior that is ignored in the usual regression tests. In the present article it is shown that a similar result is applicable to the term-structure puzzle. In particular, the above-mentioned failure 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

[redacted] For helpful comments on an earlier draft, the author is indebted to John Campbell, Tim Cook, Spencer Dale, Margarida Duarte, Eugene Fama, Kenneth Froot, Marvin Goodfriend, David Gruen, Charles Goodhart, Hubert Janicki, Greg Mankiw, Allan Meltzer, Danny Quah, Tony Smith, John Weinberg, Kenneth West, Julian Wiseman, and Alex Wolman. The views expressed are the author's and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] This article is a slightly revised version of NBER Working Paper 4938, issued in November 1994, which has been cited and utilized by a number of authors but not previously published. A few expositional changes have been made and Section 5 has been added to fill crucial gaps in the argument and to include a few references to subsequent work.

[2] General aspects of the failure are discussed by Cook and Hahn (1990), Campbell and Shiller (1991), Fama (1984), Mankiw and Summers (1984), and Evans and Lewis (1994), among others.

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 article's organization is as follows. In Section 1, the term-structure puzzle is reviewed and the article's rationalization is developed for the simplest two-period case. Then in Section 2, the analysis is extended to long rates of greater maturity. Additional evidence is developed in Section 3 after which the article's original conclusion appears as Section 4. Then a short review of more recent developments is included in Section 5, where an important difficulty neglected in the original version is described together with a resolution due to Romhányi (2002). Important subsequent work by Kugler (1997), Hsu and Kugler (1997), Dai and Singleton (2002), Gallmeyer, Hollifield, and Zin (2005), and others is briefly discussed.

1. TWO-PERIOD CASE

We begin by considering the basic 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 pure 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]

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

Here $E_t r_{t+1} = E(r_{t+1} \mid \Omega_t)$ with $\Omega_t = \{r_t, r_{t-1}, ..., R_t, R_{t-1}, ...\}$ so we are assuming rational expectations. The random variable ξ_t is a "term premium" that is often assumed constant.[5] Defining the expectational error $\epsilon_{t+1} = r_{t+1} - E_t r_{t+1}$, equation (1) implies

$$0.5(r_{t+1} - r_t) = (R_t - r_t) - \xi_t + 0.5\epsilon_{t+1}. \quad (2)$$

[3] After first drafting the article I became 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 Sections 2 and 3.

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

[5] Terminologically, many writers define the expectations hypothesis in a manner that requires that ξ_t is a constant. Campbell (1995), for example, does so and also defines the "pure expectations theory" as implying that the constant is zero. The definition used in this article permits a time-varying ξ_t but requires that (in the present case) R_t must move point for point with $0.5(r_t + E_t r_{t+1})$ for any given value of ξ_t.

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

$$0.5(r_t - r_{t-1}) = \alpha + \beta(R_{t-1} - r_{t-1}) + \text{disturbance} \tag{3}$$

should have a probability limit of 1.0. An estimated value significantly different from 1.0 is inconsistent with either 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.[6] The former, which pertains to observations taken before the founding of the Federal Reserve, will be discussed in Section 3. The latter is accompanied by a rather large asymptotic standard error that, according to Campbell and Shiller (1991, 510), seriously understates "the true uncertainty about the regression coefficients" due to finite-sample bias.[7]

One possible explanation for these findings is, of course, that the expectations theory is simply untrue—but the quantitative extent of the discrepancy seems implausibly large. Another possibility is invalidity of the rational expectations (RE) hypothesis,[8] but it seems unlikely that the same general type of systematic expectational error would prevail over different sample periods. Also, it would again appear that the magnitude of the discrepancy is too large to be explained by a departure from expectational rationality.[9] In any event, my 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 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)]:

$$\xi_t = \rho\xi_{t-1} + u_t. \tag{4}$$

Here u_t is white noise and $|\rho| < 1.0$. To this writer it seems implausible that there would not be *some* period-to-period variability in the discrepancy ξ_t

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

[7] 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.

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

[9] This point has also been made by Dotsey and Otrok (1995).

Table 1 Empirical Results, Two-Period Case

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

in (1), a random component that reflects changes in tastes regarding the need for financial flexibility or any of 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 true 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, i.e., 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 United States (and many other nations) involves manipulation of a short-term interest rate "instrument" or "operating target." Specifically, we assume that [10]

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

[10] 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.

where $\sigma \geq 0$ is presumed to be close to 1.0 and $\lambda \geq 0$ to be smaller than 2.[11] 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 larger 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 50 years.[12] (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.[13] Admittedly,

---

[11] In what follows, $\lambda < 2$ will be presumed because it seems plausible and is useful—sufficient, or, necessary for all possible values of ρ—in avoiding infinite discontinuities in ϕ_2, a coefficient in the solution equation specified in (7) below. But the solutions obtained below, and most of the analysis, would continue to prevail with $\lambda \geq 2$.

[12] 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.

[13] In an influential 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 $\lambda = \theta/(1-\theta)$ so our analysis applies. (In this case dynamic stability (non-explosiveness) requires $\delta < 1 - \theta$, however, assuming that $0 < \theta < 1$). It is not clear that Goodfriend would agree with the above formulation of his argument: another possibility is $r_t = r_{t-1} + \theta(R_t - R_{t-1}) + \zeta_t$, which would greatly increase the complexity of the algebra of our analysis. In any event, the policy behavior pattern in his article has a substantial

in actual practice the Fed has used other predictor variables in addition to or 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).[14]

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

$$(1+\lambda)r_t = \sigma r_{t-1} + \lambda[0.5(r_t + E_t r_{t+1}) + \xi_t] + \zeta_t \tag{6}$$

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

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

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

$$\begin{aligned} &(1+\lambda)[\phi_0 + \phi_1 r_{t-1} + \phi_2 \xi_t + \phi_3 \zeta_t] = \\ &\sigma r_{t-1} + \lambda[0.5(\phi_0 + \phi_1 r_{t-1} + \phi_2 \xi_t + \phi_3 \zeta_t) + \\ &0.5(\phi_0 + \phi_1(\phi_0 + \phi_1 r_{t-1} + \phi_2 \xi_t + \phi_3 \zeta_t) + \phi_2 \rho \xi_t) + \xi_t] + \zeta_t. \end{aligned} \tag{8}$$

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

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.

[14] Students of the price level determinacy literature—e.g., McCallum (1981)(1986), Dotsey and King (1983), Canzoneri, Henderson, and Rogoff (1983)—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 the policy rule (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.

$$\begin{aligned}
(1+\lambda)\phi_0 &= \lambda\phi_0 + 0.5\lambda\phi_1\phi_0 \\
(1+\lambda)\phi_1 &= \sigma + 0.5\lambda\phi_1 + 0.5\lambda\phi_1^2 \\
(1+\lambda)\phi_2 &= 0.5\lambda\phi_2 + 0.5\lambda\phi_1\phi_2 + 0.5\lambda\rho\phi_2 + \lambda \\
(1+\lambda)\phi_3 &= 0.5\lambda\phi_3 + 0.5\lambda\phi_1\phi_3 + 1.
\end{aligned} \tag{9}$$

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

$$\phi_1 = \frac{(1+0.5\lambda) \pm [(1+0.5\lambda)^2 - 2\lambda\sigma]^{1/2}}{\lambda}, \tag{10}$$

but the MSV criterion implies that the one with the minus sign is relevant.[15] 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 $[(1+0.5\lambda)-(1-0.5\lambda)]/\lambda = \lambda/\lambda = 1$ and the other three equalities in (9) are simplified considerably. They yield $\phi_0 = 0$, $\phi_2 = \lambda/(1-0.5\rho\lambda)$, and $\phi_3 = 1$ so the solution for r_t is

$$r_t = r_{t-1} + \frac{\lambda}{(1-0.5\rho\lambda)}\xi_t + \zeta_t. \tag{11}$$

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

$$R_t - r_t = 0.5(E_t r_{t+1} - r_t) + \xi_t = (1-0.5\rho\lambda)^{-1}\xi_t. \tag{12}$$

Finally, equations (11) and (4) imply

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

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

$$0.5(r_t - r_{t-1})\frac{\lambda\rho}{2}(R_{t-1} - r_{t-1}) + \frac{\lambda/2}{1-\rho\lambda/2}u_t + 0.5\zeta_t. \tag{14}$$

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

[15] 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]).

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.[16] 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 extreme 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 much more generally (i.e., even if r_t behavior is not that of a random walk).[17]

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 ξ_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 $\sigma < 1.0$. 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 the Appendix. Accordingly, the same general message applies as in the more tractable case with $\sigma = 1.0$. 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 disturbance $-\xi_{t-1} + 0.5\varepsilon_{t+1}$ in

---

[16] The claim here is not that it is inappropriate to estimate a relation of the form (3), but only that it is inappropriate to view a test of $\beta = 1$ as a test of the expectations theory.

[17] One reader has pointed out correctly that the formal analysis based on (14) presumes that policy response is to the current-period spread, not a lagged value. The argument of the present paragraph suggests that the downward-bias effect would be present, nevertheless, if response was to the lagged spread. In any case, the timing assumed in (5) is consistent with that in much of the recent literature such as Rudebusch and Wu (2004) or Bekaert, Cho, and Moreno (2005) when periods are interpreted as months or six-week intervals.

(3), implying that least-squares estimation of (4) will yield a slope coefficient that has a probability limit not equal to 1.0.

2. 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

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

where N+1 is a measure of the duration of the long rate.[18] In (5) the left-hand side is an approximation to the one-period holding return on the long-rate bond since $N(E_t R_{t+1} - R_t)$ is the (approximate) expected capital loss on the long bond. (The inexactness arises because the term R_{t+1} should pertain to a maturity one period less than that for R_t.) Thus for bonds with a distant maturity date, the approximation should be adequate.[19]

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

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

where $\varepsilon_{t+1} = R_{t+1} - E_t R_{t+1}$ 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.

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

[19] Equation (15) is based on the expression $R_t = (1 - \delta)\Sigma\delta^k E_t r_{t+k}$ + term premium, with the summation from 0 to ∞, i.e., an infinite-maturity version of the linearization developed by Shiller (1979), with $N = \delta/(1 - \delta)$. An exposition is provided by Mankiw and Summers (1984, pp. 226-7). This approximation has also been used by Shiller, Campbell, and Schoenholtz (1983), Campbell and Shiller (1991), Fuhrer and Moore (1993), and Hardouvelis (1994). The reason this approximation is adopted here is so that only two maturities—two endogenous variables—will be involved in the model.

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
Evans & Lewis (1994)	1964–1988	1 mo.	4	-0.70
Evans & Lewis (1994)	1964–1988	1 mo.	6	-1.27
Evans & Lewis (1994)	1964–1988	1 mo.	8	-1.52
Evans & Lewis (1994)	1964–1988	1 mo.	10	-1.89
Campbell & Shiller (1991)	1952–1987	1 mo.	2	0.00
Campbell & Shiller (1991)	1952–1987	1 mo.	4	-0.44
Campbell & Shiller (1991)	1952–1987	1 mo.	6	-1.03
Campbell & Shiller (1991)	1952–1987	1 mo.	12	-1.38
Campbell & Shiller (1991)	1952–1987	1 mo.	24	-1.81
Campbell & Shiller (1991)	1952–1987	1 mo.	48	-2.66
Campbell & Shiller (1991)	1952–1987	1 mo.	60	-3.10
Campbell & Shiller (1991)	1952–1987	1 mo.	120	-5.02
Hardouvelis (1994)	1954–1992	3 mo.	120	-2.90

As in the last section, we assume that the policy reaction equation (5) obtains with $\lambda < 1/N$ and that $\xi_t = \rho\xi_{t-1} + u_t$.[20] Then one can combine (5) and (15) to obtain

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

The MSV solution will be of the form

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

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

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

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

[20] The condition $\lambda < 1/N$ is the condition to prevent infinite discontinuities in π_2. It is analogous to, although different than, the condition $\lambda < 2$ for the two-period case (presumably because of the approximation in (15)) and is again assumed but not strictly required. That the larger is N, the smaller should be λ, is intuitive because the response in (5) is now to only one of the various long rates that could be considered.

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

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

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

The term in square brackets will be positive, so the MSV solution for π_1 is the expression in (21) with the minus sign.[21] 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 $\pi_1 = [(1+N\lambda)-(1-N\lambda)]/2N\lambda = 1$. 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 $1 > N\lambda, \pi_2$ is strictly positive. Given these values, we readily see that

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

and

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

Accordingly, the spread variable obeys

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

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

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

$$
\begin{aligned}
R_t - R_{t-1} &= \frac{\lambda + 1}{1 + N - N\rho(1+\lambda)}\xi_t - \frac{1}{1 + N - N\rho(1+\lambda)}\xi_{t-1} + \zeta_t \\
&= \frac{(\lambda\rho + \rho - 1)\xi_{t-1} + (1+\lambda)u_t}{1 + N(1 - \rho(1+\lambda))} + \xi_t \qquad (25) \\
&= (\lambda\rho + \rho - 1)(R_{t-1} - r_{t-1}) + \frac{(1+\lambda)}{1 + N(1 - \rho(1+\lambda))}u_t + \zeta_t.
\end{aligned}
$$

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).[22] In qualitative terms, both of these characteristics match the results of Evans and Lewis (1994) and Campbell and Shiller (1991) reported above in Table 2.

3. ADDITIONAL EVIDENCE

The article 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, 510). In terms of the notation of the present article, the two summary characterizations considered are (for the two-period case)

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

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

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

where $k > 1$. The latter "could be described as an *overreaction* model of the yield spread," according to Campbell and Shiller (1991, 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 article. Looking back at Section 1, we see that equation (12)

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

is of a similar form to that of (26), but with the crucial difference that ξ_t in (12) is *not* orthogonal to $E_t r_{t+1} - r_t$. Thus the inadequacy of (26) does not serve to discredit the model of Section 2. Furthermore, using the expression $E_t r_{t+1} - r_t = \phi_2 \rho \xi_t$ to eliminate ξ_t from (12) results in

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

for the model of Section 2. 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 article's rationalization.[23]

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) 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 1 and 2 presumes a substantial degree of interest rate smoothing, this article's explanation is consistent with both of these findings.[24]

4. REMARKS

The discussion of the foregoing paragraph suggests that one possible way of conducting additional tests of this article'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

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

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

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.[25]

In light of the foregoing discussion it will be seen that, because of the simplified nature of our policy equation (5), this article'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.[26] 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.

5. ADDITIONAL DISCUSSION

Since the article consisting of the foregoing sections was written, there have been several directly relevant developments. First, Kugler (1997) and Hsu and Kugler (1997) have conducted empirical studies based on the article's framework. In both of these studies, the results are described as supportive of the "policy reaction" hypothesis. In the process of conducting these empirical investigations, Kugler (1997) developed significant extensions of the article's theoretical analysis, one example being an application for the case

[25] Recently, Cochrane and Piazzesi (2002) have developed a "high frequency" empirical strategy that yields results for the United States that are basically consistent with policy behavior of the type hypothesized above.

[26] 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?

in which there are available observations for shorter time periods than those corresponding to the short rate (itself assumed to match the central bank's decision period). This extension is quite useful for econometric analysis of the model linking term-structure and monetary policy behavior.

A more fundamental development concerns a basic problem with the foregoing analysis of the N-period cases in Section 2. Since (15) pertains to different long maturities $N+1$, it should be written more completely as

$$R_t^{(N+1)} - NE_t(R_{t+1}^{(N)} - R_t^{(N+1)}) = r_t + \xi_t^{(N+1)} \tag{15'}$$

for $N = 1, 2, 3, ...$, where we do not retain the approximation $R_{t+1}^{(N)} = R_{t+1}^{(N+1)}$ used in (15). A crucial question, then, is how are the term premia $\xi_t^{(N+1)}$ related to each other? Also, which of the long rates is it that appears in the monetary policy rule? Evidently, the solution equations (22)–(25) cannot be correct for all N since each of them implicitly assumes that the particular long rate being analyzed is the one that appears in the central bank's policy rule.

Both of these flaws in the Section 2 analysis have been addressed by Romhányi (2002), who assumes that $\xi_t^{(N+1)} = N\Psi_t$, with Ψ_t being the same for all N and obeying the first-order AR process $\Psi_t = \rho\Psi_{t-1} + u_t$. Then we have

$$\frac{1+N}{N} R_t^{(N+1)} = \frac{r_t + NE_t R_{t+1}^{(N)}}{N} + \Psi_t,$$

which implies that for every maturity the average holding-period yield discrepancy, between the bond of duration N+1 (on the one hand) and the one-period bond plus N periods with the N-period bond (on the other hand), is the same. This equality is evidently necessary to rule out arbitrage possibilities.

With respect to the central bank's choice of a long rate for definition of the spread that is used in its policy rule, Romhányi (2002) shows that the crucial solution equation (24) becomes

$$R_t^{(N+1)} = r_t + \gamma_q[1 - \frac{1}{N+1}\frac{1-\rho^{N+1}}{1-\rho}]\Psi_t, \tag{24'}$$

where q is the maturity chosen for the policy rule and where γ_q depends upon q as well as λ and ρ but is the same (given q) for all N—see Romhányi (2002). For plausible values of λ, ρ, and q the coefficient γ_q will be positive and decreasing in q. Romhányi's modification therefore eliminates the logical inconsistencies in the argument of Section 2 above.

Over the past decade, 1995–2005, analysis of term-structure relationships has been dominated by no-arbitrage affine factor models, in which (zero-coupon) bond prices are given by a pricing equation that specifies the pricing kernel process as an affine (linear with intercept terms permitted) function of unobservable factors (state variables). Then the prices of bonds of all maturities, which satisfy no-arbitrage conditions, are also affine functions of the state variables. The substantive content of such models lies in the specification of the number of factors and the process generating the state variables. Dai and Singleton (2002) have shown that empirical features of the U.S. term structure data can be well explained by a three-factor affine model in which the "price of risk" is linearly related to the state variables. In this context, Dai and Singleton (2002, 436) report that "it turns out that McCallum's (1994) resolution of the expectations puzzle based on the behavior of a monetary authority is substantively equivalent to our [single-factor] affine parameterization of the market price of risk."

Very recently, Gallmeyer, Hollifield, and Zin (2005) have more extensively explored the relation of this article's model to "endogenous" models of the term premium, i.e., models in which the term premia are constrained to obey no-arbitrage constraints.[27] In addition, they study the way in which this article's policy rule is related to the policy rule of Taylor (1993), which has been extremely influential in both positive and normative analyses of monetary policy over the past two decades. Quantitative analysis indicates that the two parameters of this article's policy rule are plausible for a stochastic volatility specification of state variable behavior but not with a stochastic price-of-risk specification. The latter, however, is shown by Dai and Singleton (2002) to provide a superior match to actual U.S. yield-curve properties. In combination with Romhányi's results, this suggests that this article's policy rule should not be taken literally, a conclusion that is consistent with the discussion in Section 4 above.

APPENDIX

Here the concern is with the model of Section 1 when $\sigma < 1.0$. From (9), we find that

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

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

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

[27] Other notable papers that integrate monetary policy and term-structure analyses include Rudebusch and Wu (2004) and Bekaert, Cho, and Moreno (2005).

and thus using (12) that

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

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

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

which is necessarily positive since the term in parentheses equals

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

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

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

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 $E\xi_t\zeta_t = 0$:

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

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

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

The latter is unambiguously positive since $\phi_2 < 0$ and $\mid \phi_1\rho \mid < 1$. 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

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

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 $(1 + \lambda/2)^2$.

REFERENCES

Items marked (*) have been added in 2005.

Bekaert, Geert, Seonghoon Cho, and Antonio Moreno. 2005. "New-Keynesian Macroeconomics and the Term Structure." NBER Working Paper 11340, May. (*)

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

__________. 1995. "Some Lessons from the Yield Curve." *Journal of Economic Perspectives* 9 (Summer): 124–52. (*)

Canzoneri, Matthew B., Dale W. Henderson, and Kenneth S. Rogoff. 1983. "The Information Content of the Interest Rate and Optimal Monetary Policy." *Quarterly Journal of Economics* 98 (November): 545–66.

Cochrane, John H., and Monika Piazzesi. 2002. "The Fed and Interest Rates—A High-Frequency Identification." *American Economic Review* 92 (May): 90–95. (*)

Cook, Timothy, and Thomas Hahn. 1989. "The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s." *Journal of Monetary Economics* 24 (November): 331–51.

__________. 1990. "Interest Rate Expectations and the Slope of the Money Market Yield Curve." Federal Reserve Bank of Richmond *Economic Review* 76(5): 3–26.

Dai, Qiang, and Kenneth J. Singleton. 2002. "Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure." *Journal of Financial Economics* 63: 415–41. (*)

Dotsey, Michael, and Robert G. King. 1983. "Monetary Policy Instruments and Policy Rules in a Rational Expectations Environment." *Journal of Monetary Economics* 12 (September): 357–82.

__________, and Christopher Otrok. 1995. "The Rational Expectations Hypothesis of the Term Structure, Monetary Policy, and Time-Varying Term Premia." Federal Reserve Bank of Richmond *Economic Quarterly* 81 (Winter) 65–81. (*)

Estrella, Arturo, and Gikas. A. Hardouvelis. 1991. "The Term Structure as a Predictor of Real Economic Activity." *Journal of Finance* 46 (June): 555–76.

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

Fama, Eugene F. 1984. "The Information in the Term Structure." *Journal of Financial Economics* 13 (December): 509–28.

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

Fuhrer, Jeffrey C., and George Moore. 1993. "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.

Gallmeyer, Michael F., Burton Hollifield, and Stanley E. Zin. 2005. "Taylor Rules, McCallum Rules, and the Term Structure of Interest Rates." *Journal of Monetary Economics* 52 (July). (*)

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

______________. 1993. "Interest Rate Policy and the Inflation Scare Problem." Federal Reserve Bank of Richmond *Economic Quarterly* 79(1): 1–24.

Hardouvelis, Gikas A. 1994. "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): 255–83.

Hsu, Chiente, and Peter Kugler. 1997. "The Revival of the Expectations Hypothesis of the U.S. Term Structure of Interest Rates." *Economics Letters* 55 (August): 115–20. (*)

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

Kugler, Peter. 1988. "An Empirical Note on the Term Structure and Interest Rate Stabilization Policies." *Quarterly Journal of Economics* 103 (November): 789–92.

______________. 1990. "The Term Structure of Euro Interest Rates and Rational Expectations." *Journal of International Money and Finance* 9: 234–44.

______________. 1997. "Central Bank Policy Reaction and the Expectations Hypothesis of the Term Structure." *International Journal of Finance and Economics* 2: 217–24. (*)

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

Mankiw, N. Gregory, and Lawrence H. Summers. 1984. "Do Long-Term Interest Rates Overreact to Short-Term Interest Rates?" *Brookings Articles on Economic Activity 1*: 223–42. (*)

______________, and Jeffrey A. Miron. 1986. "The Changing Behavior of the Term Structure of Interest Rates." *Quarterly Journal of Economics* 101 (May): 211–28.

McCallum, Bennett T. 1981. "Price Level Determinacy With an Interest Rate Policy Rule and Rational Expectations." *Journal of Monetary Economics* 8 (November): 319–29.

______________. 1983. "On Non-Uniqueness in Rational Expectations Models: An Attempt at Perspective." *Journal of Monetary Economics* 11: 139–68.

______________. 1986. "Some Issues Concerning Interest Rate Pegging, Price Level Determinacy, and the Real Bills Doctrine." *Journal of Monetary Economics* 17 (January): 135–60.

______________. 1994. "A Reconsideration of the Uncovered Interest Parity Relationship." *Journal of Monetary Economics* 33 (February): 105–32.

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

Mishkin, Frederic S. 1990. "The Information in the Longer Maturity Term Structure about Future Inflation." *Quarterly Journal of Economics* 105 (December): 815–28.

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

Roberds, William, David Runkle, and Charles H. Whiteman. 1993. "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*. Ed. M. Goodfriend and D. H. Small. Finance and Economics Discussion Series, Federal Reserve Board.

Romhányi, Balázs. 2002. "Extending the McCallum Monetary Policy

Reaction Function for More Than Two Periods.” *Economics Letters* 77: 227–32. (*)

Rudebusch, Glenn D. 1994. “Federal Reserve Interest Rate Targeting and the Term Structure.” Working Paper (May). [Also, 1995. *Journal of Monetary Economics* 35 (April): 245–74.]

____________, and Tao Wu. 2004. “A Macro-Finance Model of the Term Structure, Monetary Policy, and the Economy.” Federal Reserve Bank of San Francisco Working Paper 2003-17. (*)

Shiller, Robert J. 1979. “The Volatility of Long-Term Interest Rates and Expectations Models of the Term Structure.” *Journal of Political Economy* 87 (December): 1190–1219.

____________, John Y. Campbell, and Kermit L. Schoenholtz. 1983. “Forward Rates and Future Policy: Interpreting the Term Structure of Interest Rates.” *Brooking Papers on Economic Activity* 1: 173–217.

____________. 1990. “The Term Structure of Interest Rates.” *Handbook of Monetary Economics*. Ed. B. M. Friedman and F. H. Hahn. Amsterdam: North-Holland Pub. Co.

Taylor, John B. 1993. “Discretion Versus Policy Rules in Practice.” *Carnegie-Rochester Conference Series on Public Policy* 39 (December): 195–214.

How Well Do Diffusion Indexes Capture Business Cycles? A Spectral Analysis

Raymond E. Owens and Pierre-Daniel G. Sarte

Regional Federal Reserve banks expend considerable effort preparing for FOMC meetings, culminating in a statement presented to the committee. Statements typically begin with an assessment of regional economic conditions, followed by an update on national economic conditions and other developments pertinent to monetary policy.

This article examines whether the regional economic information produced by the Federal Reserve Bank of Richmond (FRBR), in the form of diffusion indexes, can be tied to the business cycle. Such a link is of direct interest because of its applicability to policy decisions. Very short cycles (such as a month in length) are potentially just noise and of little policy interest. Very long cycles (such as a long-term trend) are typically thought to be driven by technological considerations over which policy has little bearing. In contrast, one generally thinks of monetary policy decisions as affecting primarily medium-length cycles or business cycles. The objective of the research herein, therefore, is to identify which of the FRBR's indexes tend to reflect primarily business cycle considerations. Indeed, indexes for which such considerations are small or nonexistent have little hope of providing any information about the state of aggregate production measures over the business cycle, and their calculation would be of limited value.

At the regional level, economic data are less comprehensive and less timely than at the national level. For example, no timely data are published on state-level manufacturing output or orders. In addition, published data on Gross State Product (GSP) are available with lags of 18 months or more. Also, these

We wish to thank Andreas Hornstein, Yash Mehra, John Walter, and Andrea Waddle for helpful comments and suggestions. The views expressed are the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

published data are available to FOMC members as soon as they are available to the Reserve Banks so that their analysis by the latter adds little to the broader monetary policy process. These shortcomings have led a number of organizations—including several regional Federal Reserve banks—to produce their own regional economic data. These efforts mostly have taken the form of high-frequency surveys. Surveys provide speed and versatility, overcoming the obstacles inherent in the traditional data. But surveys are often relatively expensive per respondent, leading organizations to maintain relatively small sample sizes. Further, to limit the burden on respondents, survey instruments often ask very simple questions, limiting the information set and level of analysis.

The Richmond Fed conducts monthly surveys of both manufacturing and services sector activity. The number of survey respondents is usually around 100 and respondents report mostly whether a set of measures increased, decreased, or was unchanged. However, there are several measures—primarily changes in prices—reported as an annual percentage change. Results from these surveys, along with Beige Book information, comprise the foundation of regional economic input into monetary policy discussions.

That said, there are several reasons why one may be skeptical of diffusion indexes' ability to capture useful variations in the business cycle. Specifically, the usefulness of diffusion indexes hinges critically on the following aspects of survey data:

- Diffusion indexes are produced from data collected at relatively high frequency—with new indexes being typically released every month—and therefore potentially quite noisy.

- The types of questions being asked allow for very little leeway in respondents' answers. For example, the regional diffusion indexes produced by the FRBR are calculated from survey answers that only distinguish between three states from one month to the next. Thus, we ask only whether shipments, say, are up, down, or unchanged relative to last month. In particular, let I, D, and N denote the number of respondents reporting increases, decreases, and no change respectively, in the series of interest. The diffusion index is then simply calculated as

$$I = \left(\frac{I - D}{I + N + D} \right) \times 100. \qquad (1)$$

  Observe that I is bounded between -100 and 100, and takes on a value of zero when an equal number of respondents report increases and decreases.

- The surveys must contain a large enough sample in order that a diffusion index capture potentially meaningful variations at business cycle

frequencies. As a stark example, note that if only two firms were surveyed, the index I above would only ever take on five values, $\{-100, -50, 0, 50,$
$100\}$. If three firms were sampled, I in (1) would only ever take on the values $\{-100, -66, -33, 0, 33, 66, 100\}$. Evidently, I will take on more and more values the more firms are sampled. This may not be a problem for identifying whether the resulting index is driven mainly by business cycle considerations *per se*, but will affect the degree to which such indexes commove with more continuous aggregate measures of production over the cycle.

- Composition effects will also affect this last observation. To see this, suppose that periods of recessions and expansions are characterized by all firms decreasing and increasing their shipments respectively as changes in demand occur. Then, even with a large sample, the diffusion index in (1) could never take on any other value than -100 and 100 and would, therefore, offer no information on the relative strength of economic conditions. This will not be the case, however, when the number of firms reporting decreases or increases in shipments, say, varies in a systematic fashion with the extent of recessions and expansions.

- Finally, respondents possess much discretion in the way they answer survey questions. Thus if a given manufacturer's new orders, say, increased or decreased this month by only a "small" amount relative to last month, she may decide to report no change in her orders. But the key point here is that the definition of "small" is left entirely to the respondent's discretion.

1. SOME KEY CONCEPTS IN FREQUENCY DOMAIN ANALYSIS

Before tackling the issue of whether regional diffusion indexes have anything to do with business cycles, let us briefly review some important concepts that we shall use in our analysis. In particular, the material in this section summarizes central notions of frequency domain analysis that can be found in Hamilton (1994), Chapter 6; Harvey (1993), Chapter 3; as well as King and Watson (1996).

The *spectral representation theorem* states that any covariance-stationary process $\{Y_t\}_{t=-\infty}^{\infty}$ can be expressed as a weighted sum of periodic functions of the form $\cos(\lambda t)$ and $\sin(\lambda t)$:[1]

$$Y_t = \mu + \int_0^{\pi} \alpha(\lambda)\cos(\lambda t)d\lambda + \int_0^{\pi} \delta(\lambda)\sin(\lambda t)d\lambda, \tag{2}$$

where λ denotes a particular frequency and the weights $\alpha(\lambda)$ and $\delta(\lambda)$ are random variables with zero means.

Generally speaking, given that any covariance-stationary process can be interpreted as the weighted sum of periodic functions of different frequencies, a series' *power spectrum* gives the variance contributed by each of these frequencies. Thus, summing those variances over all relevant frequencies yields the total variance of the original process. Moreover, should certain frequencies, say $[\lambda_1, \lambda_2]$, mainly drive a given series, then the variance of cycles associated with these frequencies will account for the majority of the total variance of that series.

A Simple Example

In order to make matters more concrete, consider the following example. Define the following process for a hypothetical economic time series, Y_t,

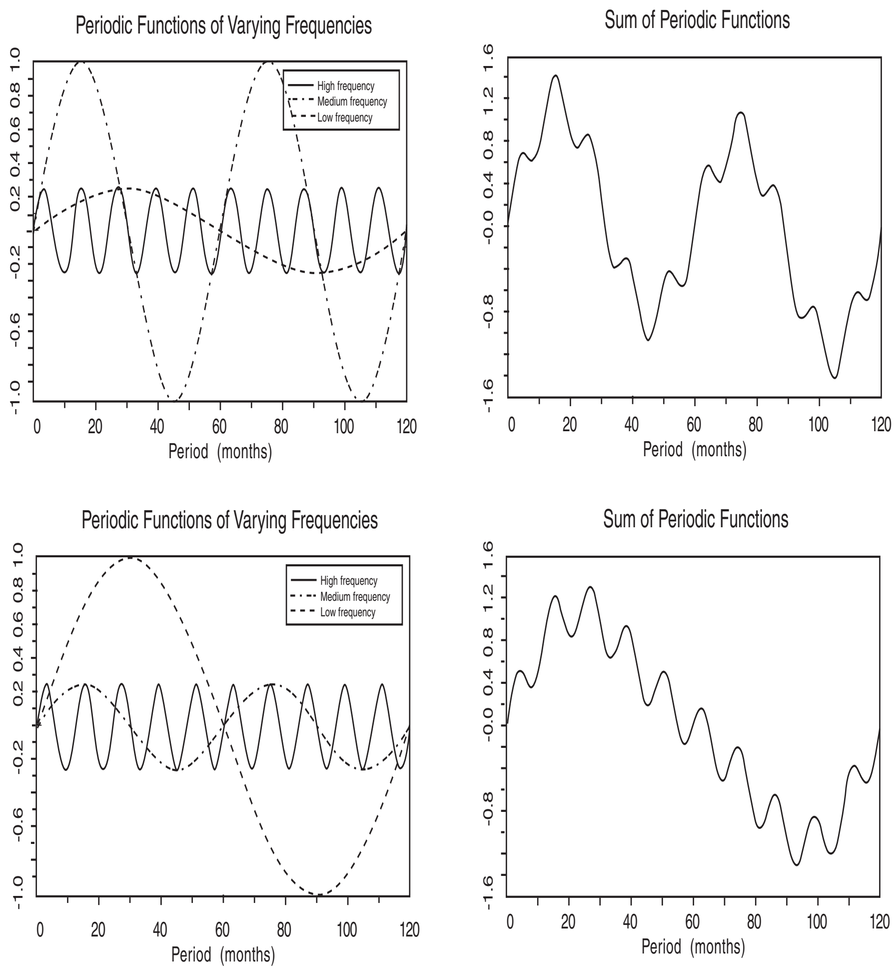
$$Y_t = \alpha_1 \sin(\lambda_1 t) + \alpha_2 \sin(\lambda_2 t) + \alpha_3 \sin(\lambda_3 t), \tag{3}$$

where the α_i's and λ_i's are strictly positive real numbers. A sine function is bounded between -1 and 1, so that the first term on the right-hand side of equation (3) will oscillate between $-\alpha_1$ and α_1, the second term between $-\alpha_2$ and α_2, etc. We refer to α_i as the amplitude of the component of Y_t associated with $\alpha_i \sin(\lambda_i t)$. A function is periodic with period T when the function repeats itself every T periods. The period of a sine function is defined as 2π divided by its frequency. Thus, the first term on the right-hand side of (3) will repeat itself every $2\pi/\lambda_1$ periods, the second term every $2\pi/\lambda_2$ periods, etc. Furthermore, observe that the higher the frequency, the faster a periodic function repeats itself.

For additional concreteness, assume now that one unit of time is a month, and that in the above example, $\{\alpha_1, \lambda_1\} = \{0.25, \frac{\pi}{6}\}$, $\{\alpha_2, \lambda_2\} = \{1, \frac{\pi}{30}\}$, and $\{\alpha_3, \lambda_3\} = \{0.25, \frac{\pi}{60}\}$. Then, the components of Y_t given by $\alpha_1 \sin(\lambda_1 t)$ and $\alpha_3 \sin(\lambda_3 t)$ have the shortest and longest periods, one year (i.e., a seasonal cycle) and 10 years, respectively, as well as the smallest amplitude, 0.25. We refer to these components as the high- and low-frequency components of Y_t, respectively. In contrast, the component of Y_t given by $\alpha_3 \sin(\lambda_3 t)$ repeats itself every $2\pi/(\pi/30) = 60$ months, or five years. Thus, we refer to this component as the medium-frequency or business cycle component of Y_t. Note

[1] A stochastic process, Y_t, is covariance stationary if $E(Y_t) = \mu$ and $E(Y_t T_{t-s}) = \sigma_s \forall t$ and s.

Figure 1 Examples of Aggregation of Periodic Functions



also that $\alpha_2 \sin(\lambda_2 t)$ has the largest amplitude of all three components since $\alpha_2 = 1$. The upper left-hand panel of Figure 1 illustrates these periodic functions separately over a period of 10 years. We can clearly see that the slowest moving periodic function (i.e., the low-frequency component) repeats itself exactly once over that time span. In contrast, the business cycle component repeats itself twice and dominates in terms of its amplitude.

The upper right-hand panel of Figure 1 illustrates the sum of these periodic components. It is clear that Y_t repeats itself twice over the 10-year time span. Put another way, Y_t in this case is primarily driven by its business cycle or medium-frequency component. This is because this component has the largest amplitude and matters most, while the high- and low-frequency components have relatively small amplitude. In particular, the amplitude of Y_t is $\alpha_1 + \alpha_2 + \alpha_3 = 1.5$, with two-thirds of that amplitude being contributed by the medium-frequency component. Since, strictly speaking, the power spectrum relates to variances, the fraction of total variance of Y_t explained by the component $\alpha_2 \sin(\lambda_2 t)$ in this case is $1/(0.25^2 + 0.25^2 + 1)$, or 89 percent.[2]

As an alternative example, suppose that $\alpha_2 = 0.25$ while $\alpha_3 = 1$, with all other parameters unchanged. This case is depicted in the lower left-hand panel of Figure 1, where it is the component that repeats itself just once over 10 years that now evidently dominates in terms of amplitude. The sum of low-, medium-, and high-frequency components, Y_t, is given in the lower right-hand side panel of Figure 1, and notice that it reflects mainly its slowest moving element, $\alpha_3 \sin(\lambda_3 t)$. And indeed, contrary to our earlier example, it is now this low-frequency component that accounts for the bulk of the total variance of Y_t, or two-thirds of its amplitude.

Formally, one defines the population spectrum of Y as

$$f(\lambda) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\lambda j}, \; -\pi \leq \lambda \leq \pi \tag{4}$$
$$= \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j \cos(\lambda j) \right],$$

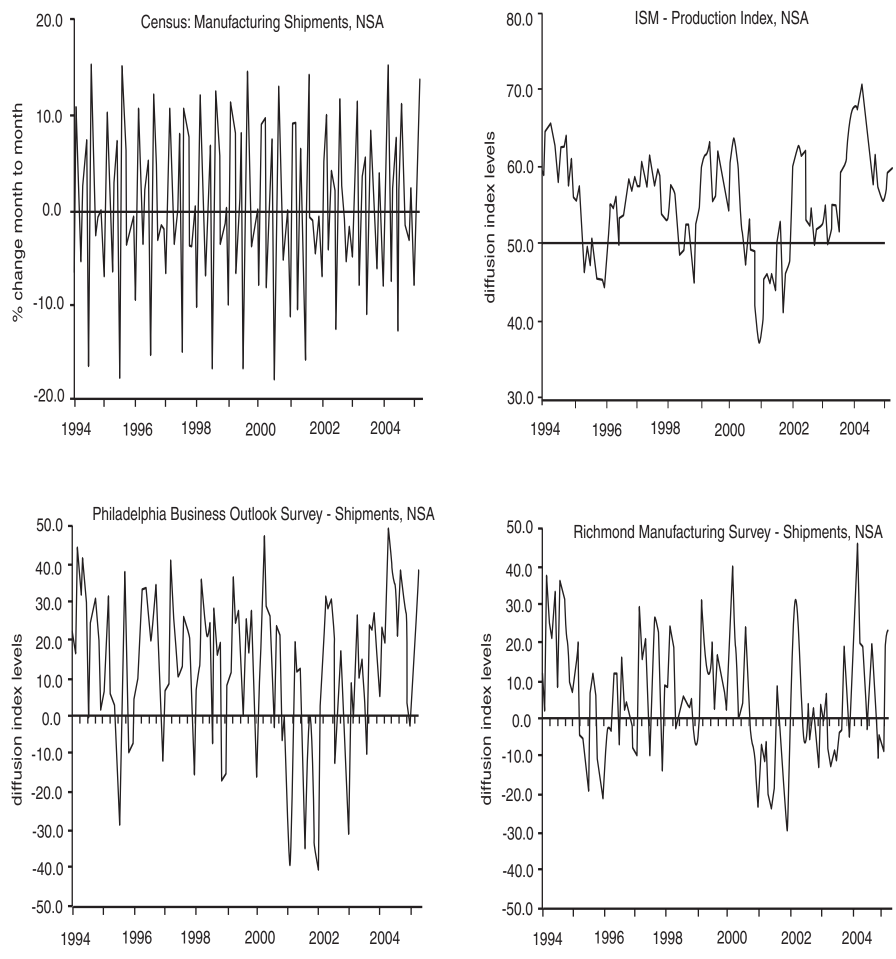
where $i^2 = -1$ and γ_j is the j^{th} auto-covariance of Y, $\text{cov}(Y_t, Y_{t\pm j})$. In a manner similar to our example above, economic time series that are driven principally by business cycle forces will have most of their variance (or amplitude) associated with cycles lasting between one and a half to eight years. We can think of $f(\lambda)$ in equation (4) as the variance of the periodic component with frequency λ. Similarly, in the above example, the components $\alpha_i \sin(\lambda_i t)$ have different amplitude or variance. More specific attributes of the power spectrum are given in Appendix A. Details of estimation and calculations for the results that follow are given in Appendix B.

2. EXAMPLES WITH MANUFACTURING DATA

Figure 2 plots the behavior of manufacturing shipments as actually recorded by the Census at the national level, and as captured by different indexes including the Institute of Supply Management (ISM) index, the Federal Reserve Bank of Philadelphia (FRBP) Business Outlook survey, and the FRBR regional survey.

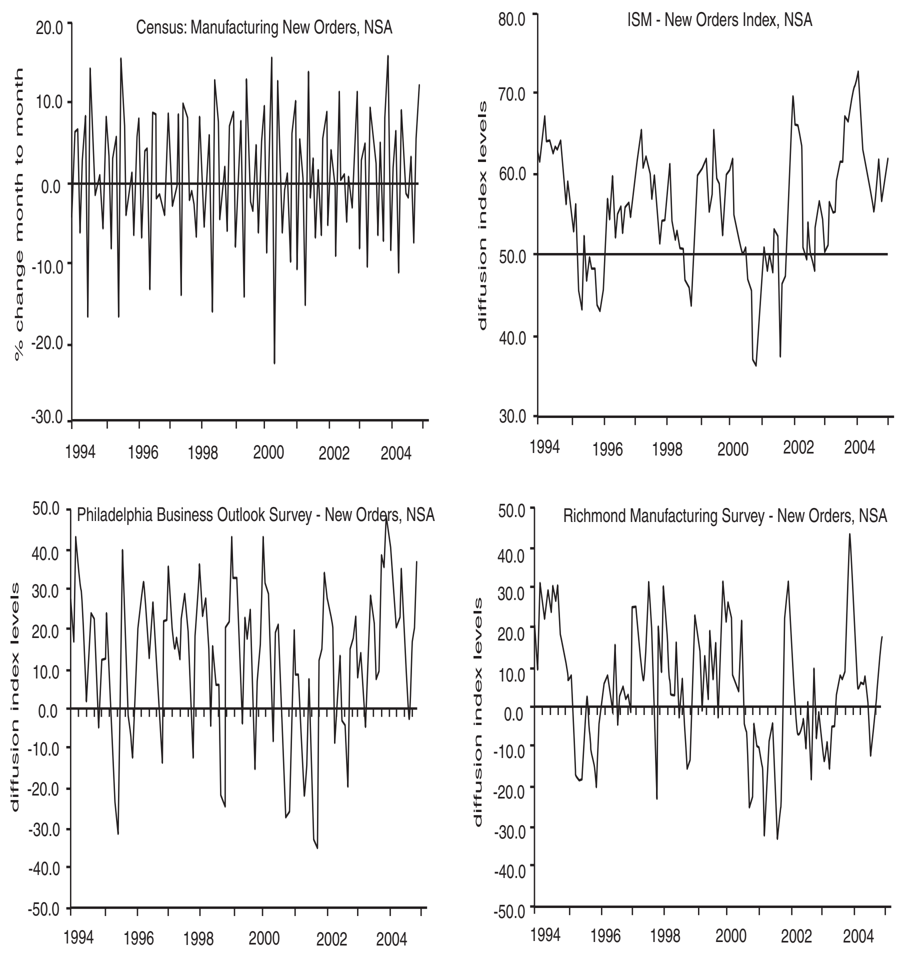
[2] In particular, amplitude and variance are closely related here since $\text{var}(\alpha_i \sin(\lambda_i t)) = \alpha_i^2 \text{var}(\sin(\lambda_i t))$ and $\text{var}(\sin(\lambda_i t)) = \text{var}(\sin(\lambda_j t))$ for $i \neq j$. Therefore, the fraction of total variance explained by the component $\alpha_i \sin(\lambda_i t)$ is $\alpha_i^2 / \sum_i \alpha_i^2$.

Figure 2 Measures of Manufacturing Shipments



Because the FRBR only began to produce its diffusion indexes in November 1993, we chose to homogenize our samples in Figure 2 and show the behavior of the series over the same period. Although the actual monthly manufacturing shipments and the ISM index are meant to reflect similar information, there are clear differences between the two series. The ISM does not make public the formula it uses for translating its respondents' answers into an actual diffusion index, but it is apparent that it produces a much smoother series. At the same time, observe that we can clearly see a common cyclical pattern between the FRBR's manufacturing shipments survey and the corresponding ISM index. The regional diffusion indexes are also smoother than the actual national data,

Figure 3 Measures of Manufacturing New Orders



but this could be indicative of the specific regional industrial makeup of the Third and Fifth Federal Reserve Districts. These observations all apply to the behavior of new orders in Figure 3.

A presumption of our analysis is that manufacturing data fluctuates over time to reflect evolving business cycle conditions. However, this is certainly not obvious from the upper left-hand panel in Figures 2 and 3, where the series seem primarily driven by very fast-moving random components. Economic analysts implicitly recognize this fact when commenting on the behavior of manufacturing data and, indeed, informal discussions of the current data are often framed relative to other episodes. In other words, analysis of the data

Table 1 Aggregate National Data
Percent of variance attributable to cycles with different periods

	periods>8 years	1.5 years< periods<8 years	periods>6 mo.
Shipments	19.00	71.30	97.90
New Orders	17.29	67.89	93.75
Employment	33.80	62.76	99.64

often involves the use filters, whether implicitly or explicitly, in the hope to gain some insight from the series about evolving economic conditions.[3] In principle, one can apply any filter one wishes to the data (that leaves the resulting series covariance stationary) and estimate the corresponding power spectrum to determine to what degree business cycle components are actually being emphasized.

To illustrate this last point, Figure 4 shows estimated power spectra for manufacturing shipments, new orders, and employment data based on both the series' month-to-month and year-to-year changes. The solid vertical lines in the figures cover the frequencies associated with the conventional definition of business cycles, $[\pi/9, \pi/48]$, which correspond to cycles with periods ranging from one and a half to eight years. The dashed vertical line corresponds to cycles with a period of six months, $\lambda = \pi/3$. Observe that cycles have longer and longer periods as we move toward zero on the horizontal axis.

Figure 4 shows that month to month, both national manufacturing shipments and new orders power spectra exhibit multiple peaks at high frequencies. Thus, the monthly observations are driven mainly by short-lived random periodic cycles that are not necessarily informative for the purposes of policymaking. In contrast, the power spectra for the 12-month difference of the manufacturing data series all contain a high notable peak in the business cycle interval, as well as a lower peak at roughly frequency $\lambda = 0.3$ (i.e., cycles of length close to two years). King and Watson (1996) refer to the shape of the power spectra in the right-hand panels of Figure 4 as the typical spectral shape for differences in macroeconomic time series. Cycles that repeat themselves on a yearly basis, and are thus associated with seasonal changes, have frequency $\lambda = \pi/6 = 0.53$, and we can see that the spectra in the right-hand panels of Figure 4 also display a small peak just to the right of that frequency.

Table 1 gives the fraction of total variance attributable to cycles of different lengths for the manufacturing series depicting year-to-year changes.

As in the analysis of King and Watson (1996), the business cycle interval contains the bulk of the variance of the yearly change in these macroeconomic

[3] By filters, we mean a transformation of the original time series such as a moving average or an $n > 1$ order difference.

Figure 4 Power Spectra for Actual Manufacturing Data

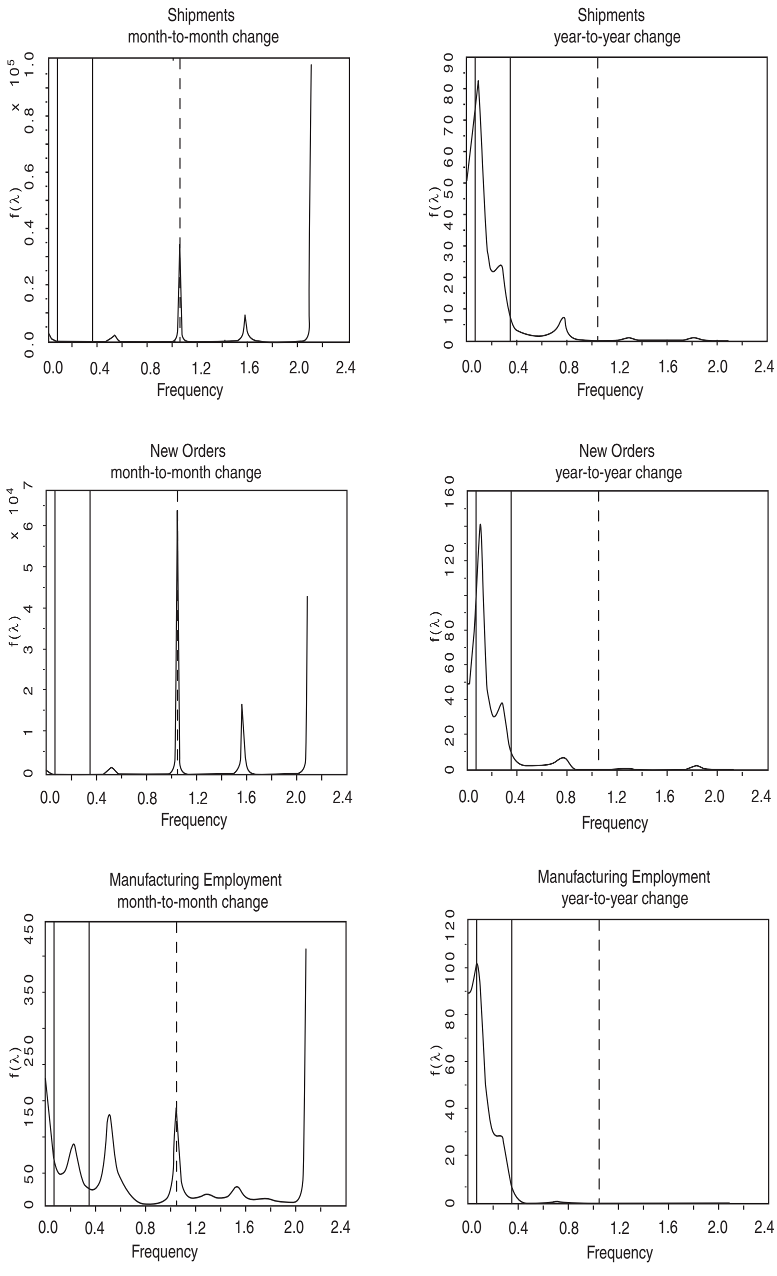


Table 2 ISM Indexes
Percent of variance attributable to cycles with different periods: ISM indexes

	periods>8 years	1.5 years<periods<8 years	periods>6 mo.
Composite Index	18.31	59.64	94.86
Shipments	11.40	57.12	87.78
New Orders	11.69	56.40	89.12
Employment	17.83	61.62	95.80

time series. Some nontrivial contribution to total variance does stem from longer-lived cycles (i.e., those with periods greater than eight years). At the other extreme, virtually no contribution to variance is attributable to cycles with periods less than six months. Observe also in Figure 4 that, outside of the business and seasonal cycles, the power spectra are close to zero.[4]

3. POWER SPECTRA FOR DIFFUSION INDEXES

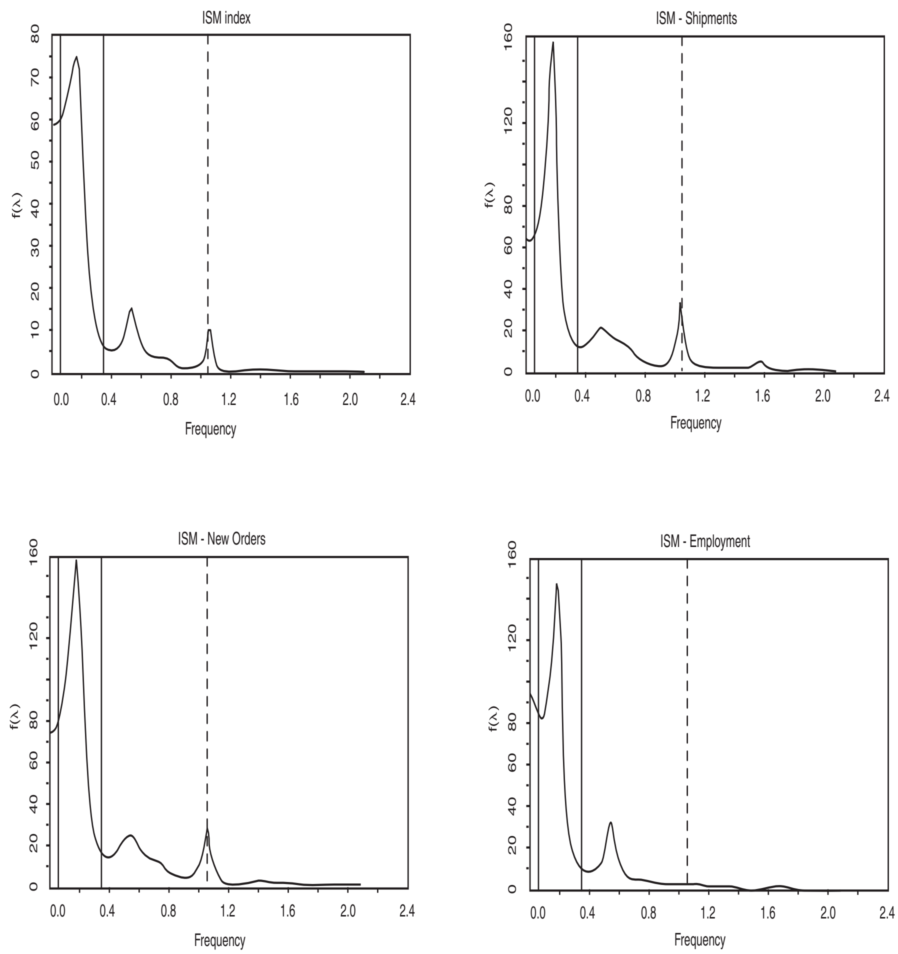
Figure 5 displays estimated power spectra for the ISM diffusion indexes corresponding to the manufacturing series in Figure 4. Interestingly, even though the indexes not filtered in any way, all possess the typical spectral shape associated with differences in macroeconomic time series. In particular, a principle and notable peak in each case occurs well within the business cycle interval. The spectra for the diffusion indexes associated with shipments and new orders suggest an important six-month cycle, and all indexes further emphasize a yearly cycle with a peak occurring almost exactly at frequency $\lambda = \pi/6$. The estimated spectra associated with the ISM indexes suggest virtually no contribution from cycles with periods less than six months.

Thus, although many caveats are associated with survey-generated indexes, it appears that these indexes nonetheless capture systematic aspects of changes in economic time series that virtually mimic those of actual data. This observation is particularly important in that survey data can be much less costly, and always much faster, to produce than measuring changes in actual economic data. In the case of federal regional districts, for instance, state manufacturing data is not even collected; but corresponding diffusion indexes can be produced by the various Federal Reserve Banks in a relatively inexpensive and timely manner.

Finally, the power spectra shown in Figure 5 are indicative of two important aspects of changes in economic conditions. First, it is noteworthy that the untransformed survey data and the year-over-year changes in the national

[4] Results in this case do not depend only on the natural properties of the data, but also on the specific form of the filter. For instance, a 12-month difference filter will by construction eliminate all variations in cycles shorter than one year.

Figure 5 Power Spectra for ISM Diffusion Indexes



aggregate display similar spectral shapes. Second, and related to this last observation, while surveys allow for much discretion in the way respondents answer questions, this discretion does not obscure the informational content of the responses in such a way as to simply produce statistical noise, or even emphasize high-frequency changes.

Table 2 gives a decomposition of variance for the different diffusion indexes in Figure 5 according to cycles of different frequencies.

As with actual manufacturing data in Table 1, the bulk of the overall variance in diffusion indexes is contained within the business cycle frequencies, albeit to a somewhat lesser extent. This reinforces the notion that diffusion indexes capture specific aspects of changes in economic conditions. In this case in particular, and unlike the 12-month difference of actual manufacturing data, the power spectra suggest that some nontrivial portion of the overall variance in the indexes stem from shorter seasonal cycles, those associated with six-month and one-year periods. Shorter cycles, however, appear to play no role in respondents' answers.

Power Spectra for FRBR Regional Diffusion Indexes

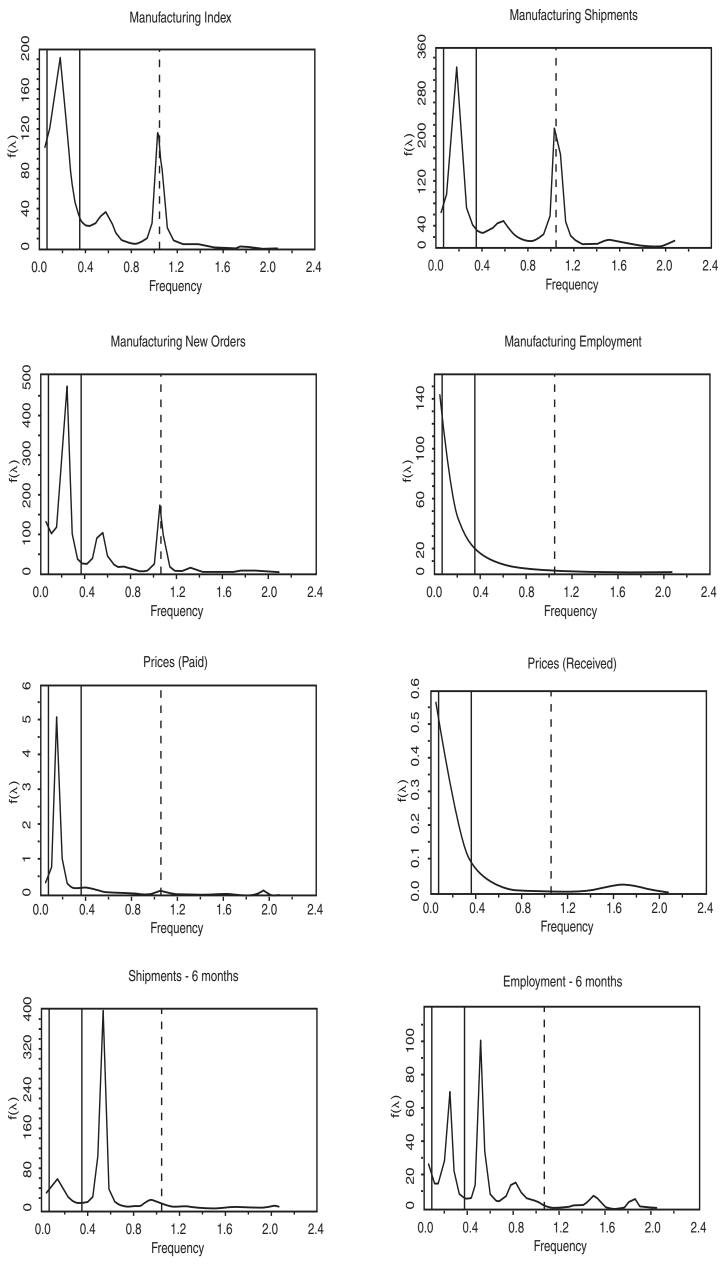
The FRBR's manufacturing survey produces diffusion indexes according to the formula described in the introduction for shipments, new orders, employment, and an overall index. Fifth District businesses are also surveyed regarding prices, as well as expected shipments and employment six months ahead.

Cyclical Properties of Manufacturing Indexes in the Fifth Federal Reserve District

Figure 6 shows estimated power spectra for the various raw (i.e., unfiltered) diffusion indexes produced by the FRBR in manufacturing. Perhaps most surprisingly, it is not the case that the power spectra are indicative of mostly short-lived cyclical noise, even at the relatively narrow regional level. On the contrary, the diffusion indexes display distinctive patterns. More specifically, it appears that the survey respondents do not strictly answer the questions they are asked—(relating simply to changes relative to the previous month)—but instead carry out some implicit deseasonalization. In particular, as with the ISM, the spectrum for the untransformed survey display distinct similarities with the year-over-year changes in the national aggregates. The overall manufacturing index, as well as shipments and new orders, display three distinctive peaks: one in the business cycle interval, a smaller one that captures approximately a 12-month seasonal cycle at $\lambda = 0.53$, as well as distinct evidence of a six-month cycle. Prices paid and received reported by survey respondents also emphasize business cycle frequencies, rather than shorter-lived cycles where the power spectrum is essentially zero. Therefore, it appears that despite the simplicity of the questions asked, which essentially restrict respondents to three states, the questions are asked of enough agents that the corresponding diffusion index captures time variations that move strongly either with business or seasonal cycles.

The figures for expected shipments and employment six months ahead are somewhat less informative. Indeed, the power spectra capture variations that are principally driven by a 12-month seasonal cycle, possibly suggesting that respondents are basing their answers mainly on what they expect during the course of a given year. Thus, key dates that occur on a yearly basis,

Figure 6 Power Spectra for FRBR Manufacturing Diffusion Indexes



such as Christmas or even, say, yearly shut-down periods driven by retooling considerations, seem to play a key role in shaping their expectations.

Table 3 FRBR Manufacturing Diffusion Indexes: (Unadjusted)
Percent of variance attributable to cycles with different periods

	periods>8 years	1.5 years<periods<8 years	periods>6 mo.
Overall Index	8.26	47.76	77.52
Shipments	3.48	38.27	65.68
New Orders	7.34	43.83	72.64
Employment	26.29	41.05	82.88
Prices (paid)	5.43	75.90	94.83
Prices (received)	20.45	52.89	83.06
Shipments–6M	2.36	14.32	52.92
Employment–6M	6.43	20.88	61.13

Table 3 gives the fraction of variance attributable to cycles of different periods for the various manufacturing regional indexes. On the whole, these indexes capture more movement stemming from short-lived cycles relative to the actual manufacturing data in Table 1. Cycles with periods greater than six months can leave up to 47 percent of the total series' variance unaccounted for (e.g., expected shipments six months ahead). However, except for expected future employment and shipments, the business cycle interval does contain a nontrivial fraction of the total variance for the various series, ranging from 38.27 to 75.90 percent. Prices paid, as simply reported in the monthly survey, appear to move most strongly with business cycle frequencies. As suggested above, expected employment and shipments six months ahead have the least to do with business cycles.

Because the unfiltered manufacturing diffusion indexes are driven to a non-negligible extent by relatively short-lived cycles that are presumably less relevant to policymaking decisions, we also consider a six-month difference of all the regional diffusion indexes. The idea is to eliminate variations in the indexes that are quickly reversed in order to acquire a sharper picture of the business cycle. In particular, it should be clear by now that spectral analysis represents a natural tool in searching for a filter that helps isolate changes associated with these specific frequencies.

Figure 7 displays power spectra associated with the six-month difference of the diffusion indexes produced by the FRBR. Except for expected shipments and employment six months ahead, all power spectra now have the typical spectral shape for differences, and their main peaks lie squarely in the business cycle interval. Evidence of a small seasonal cycle lasting one year is also clearly distinguishable. Furthermore, as indicated in Table 4, the business cycle interval now contains a very large fraction of the total variation

Figure 7 Power Spectra for FRBR Manufacturing Diffusion Indexes 6-Month Difference

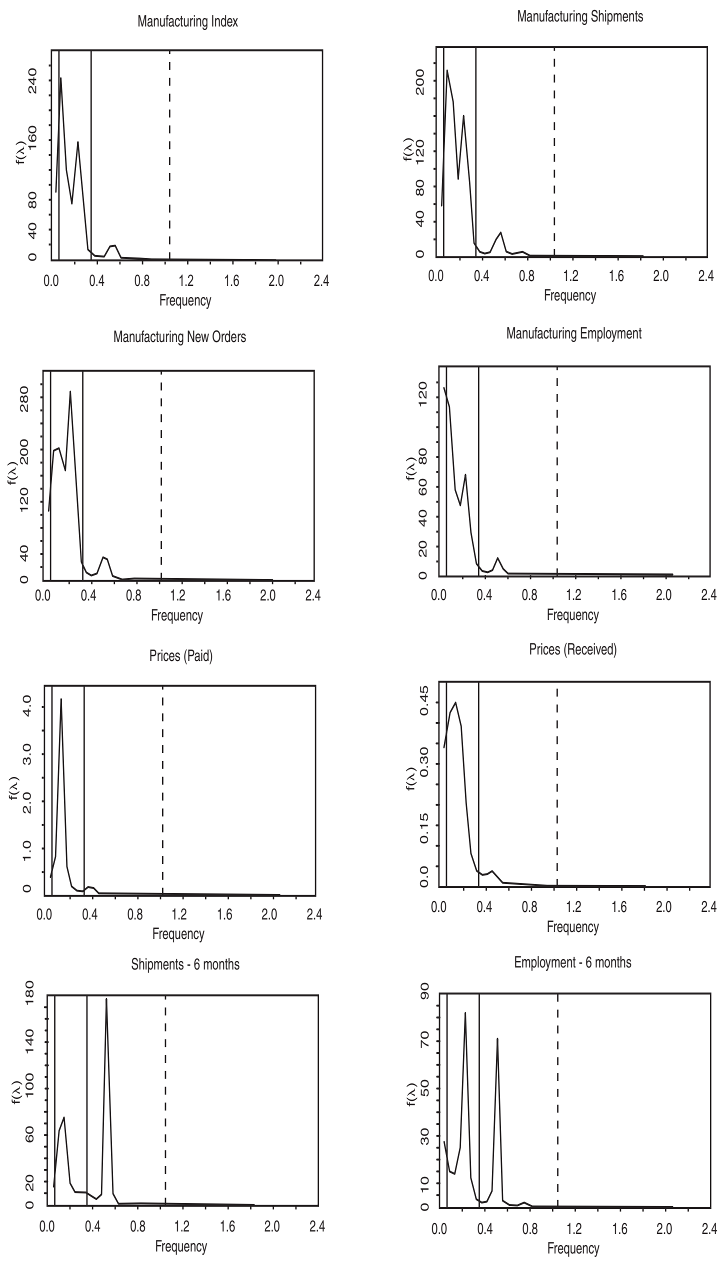


Table 4 FRBR Manufacturing Diffusion Indexes: (6-Month Difference)
Percent of variance attributable to cycles with different periods

	periods>8 years	1.5 years<periods<8 years	periods>6 mo.
Overall Index	12.63	82.87	1.00
Shipments	7.62	84.67	1.00
New Orders	10.62	82.45	1.00
Employment	30.07	63.67	1.00
Prices (paid)	8.15	79.30	96.25
Prices (received)	20.37	72.59	99.43
Shipments–6M	4.60	49.18	79.53
Employment–6M	12.07	36.33	69.52

in the series. Interestingly, the six-month difference filter leaves the spectra associated with prices relatively unchanged.

4. FINAL REMARKS

Information on economic activity gathered from high-frequency surveys offers a timely gauge of conditions in the sector surveyed. The value of this timely information to monetary policymakers depends not only on whether the information accurately reflects conditions within the sector, but also on whether the information infers something about conditions that monetary policy can address, such as movements in the business cycle. That is, if survey results typically deviate from trend very often or very seldom, the information gained from the results may suggest changes in economic conditions at frequencies largely immune to monetary policy capabilities and may be of little value to policymakers, even if the results are an accurate reading of sector conditions. In contrast, if the deviations occur with a frequency similar to that of the business cycle, monetary policymakers can use the information to better shape policy.

In this article, we estimate power spectra for the results from two high-frequency surveys and show that deviations from trend generally occur at business-cycle-length frequencies in manufacturing indexes. The proportion of variation captured in business-cycle-length frequencies is strongest for a six-month moving average of the Richmond results.

APPENDIX A

Some important features of the power spectrum are as follows:

- $\gamma_0 = \int_{-\pi}^{\pi} f(\lambda)d\lambda$. In other words, the area under the population spectrum between $-\pi$ and π integrates to the overall variance of Y.
- Since $f(\lambda)$ is symmetric around 0, $\gamma_0 = 2\int_0^{\pi} f(\lambda)d\lambda$. More generally, $2\int_0^{\lambda_1} f(\lambda)d\lambda$ represents the portion of the variance in Y that can be attributed to periodic random components with frequencies less than or equal to λ_1.
- Recall that if the frequency of a cycle is λ, the period of the corresponding cycle is $2\pi/\lambda$. Thus, a conventional frequency domain definition of business cycles, deriving from the duration of business cycles isolated by NBER researchers using the methods of Burns and Mitchell (1946), is that these are cycles with periods ranging between 18 and 96 months. Therefore, in the frequency domain, business cycles are characterized by frequencies $\lambda \in [\pi/48, \pi/9] \approx [0.065, 0.35]$.
- The power spectrum is not well defined for frequencies larger than π radians. The frequency $\lambda = \pi$ is known as the *Nyquist* frequency and corresponds to a period of $2\pi/\pi = 2$ time units. To see the relevance of this concept, note that with quarterly data, no meaningful information can be obtained regarding cycles shorter than two quarters since, by definition, the shortest observable changes in the data are measured from one quarter to the next. Hence, changes within the quarter are not observable. In contrast, with monthly data, one can refine the calculation of the power spectrum up to a two-month cycle.
- When Y is a white noise process, $Y_t \sim^{iid} N(0, \sigma^2)$, $f(\lambda)$ is simply constant and equal to $\sigma^2/2\pi$ on the interval $[-\pi, \pi]$. If survey-generated data were mainly noise, therefore, one might expect a relatively flat power spectrum with no specific frequencies being emphasized.

APPENDIX B

Estimation of the power spectrum:

Given data $\{Y_t\}_{t=1}^{T}$, the power spectrum can be estimated using one of two approaches: a non-parametric or a parametric approach. Evidently, the simplest (non-parametric) way to estimate the power spectrum is by replacing (4) by its sample analog,

$$\widehat{f}(\lambda) = \frac{1}{2\pi}\left[\widehat{\gamma}_0 + 2\sum_{j=1}^{T-1}\widehat{\gamma}_j \cos(\lambda j)\right], \tag{5}$$

where the "hat" notation denotes the sample analog of the population autocovariances. Since our hypothetical sample contains only T observations, autocovariances for j close to T will be estimated very imprecisely and, although unbiased asymptotically, $\widehat{f}(\lambda)$ will generally have large variance. One way to resolve this problem is simply to reduce the weight of the autocovariances in (5) as j approaches T. The Bartlett kernel, for example, assigns the following weights:

$$\omega_j = \begin{cases} 1 - \frac{j}{k+1} & \text{for } j = 1, 2, ..., k \\ 0 & \text{for } j > k \end{cases},$$

where k denotes the size of the Bartlett bandwidth or window. When k is small, $\widehat{f}(\lambda)$ has relatively small variance since the autocovariances that are estimated imprecisely (i.e., those for which j is close to T) are assigned small or zero weight. However, given that the true power spectrum is based on all the autocovariances of Y, $\widehat{f}(\lambda)$ also becomes asymptotically biased. The reverse is true when k is large; the periodogram becomes asymptotically unbiased but acquires large variance. How does one then choose k in practice? Hamilton (1994) suggests that one "practical guide is to plot an estimate of the spectrum using several different bandwidths and rely on subjective judgment to choose the bandwidth that produces the most plausible estimate."

Another popular way to go about estimating the spectrum of a series is to adopt a parametric approach. Specifically, one can show that for any AR(P) process, $Y_t = \mu + \phi_1 Y_{t-1} + ... + \phi_p Y_{t-p} + \varepsilon_t$ such that $\text{var}(\varepsilon_t) = \sigma^2$, the power spectrum (4) reduces to

$$f(\lambda) = \frac{\sigma^2}{2\pi}.\left\{\left|1 - \sum_{j=1}^{p}\phi_j e^{-i\lambda j}\right|\right\}^{-1}, \text{ where } i^2 = -1. \tag{6}$$

Therefore, since any linear process has an AR representation, one can estimate an AR(P) by OLS and substitute the coefficient estimates, $\widehat{\phi}_1, ..., \widehat{\phi}_p$, for the parameters $\phi_1, ..., \phi_p$ in (6). Put another way, one can fit an AR(P) model to the data, and the estimator of the power spectrum is then taken as the *theoretical* spectrum of the *fitted* process. Note that the spectrum estimated in this way will converge to the true spectrum (as the sample size becomes large) under standard assumptions that guarantee that the coefficient estimates, $\widehat{\phi}_1, ..., \widehat{\phi}_p$, converge to the true parameters, $\phi_1, ..., \phi_p$. Of course, the difficulty lies in deciding on the order of the AR process. When P is small, the estimated spectrum may be badly biased but a large P increases its variance. The trade-off, therefore, is similar to that encountered in using the non-parametric approach described above. Harvey (1993) suggests that one solution that works well in practice is to actively determine the order of the model on a goodness-of-fit criterion, such as maximizing the adjusted R^2 statistic or minimizing Akaike's information criterion.

For the purpose of this article, power spectra will be estimated using the parametric method we have just described. Since we shall be analyzing time series with monthly data, we fit an AR(P) to each series with P being at most 24. The actual value of P is then chosen by maximizing the adjusted R^2 in each series' estimation.

REFERENCES

Hamilton, James, D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

Harvey, Andrew, C. 1993. *Time Series Models*, 2nd Edition. Cambridge: MIT Press.

King, Robert, G., and Mark W. Watson. 1996. "Money, Prices, Interest Rates and the Business Cycle." *Review of Economics and Statistics* 78 (1): 35–53.

Technological Design and Moral Hazard

Edward Simpson Prescott

The classic moral hazard model studies the problem of how a principal should provide incentives to an agent who operates a project for him. In this model, the principal only observes the realized output and not the agent's effort. Consequently, the agent must be induced to work hard with compensation that depends on performance. Because many contracts explicitly tie rewards and punishments to performance, the model is a workhorse of modern economics, with applications to insurance contracts, employee and executive pay, sharecropping contracts, corporate finance, and bank regulation, to name just a few.

In the moral hazard model the exact dependence of compensation on performance depends on the relationship between the agent's input and the project's output. Most analysis takes this relationship, or technology, as given; that is, something that cannot be modified by the principal.

There are many situations, however, where the principal has some control over this technology. For example, a principal can design a production process so some outcomes are more likely than others when certain inputs are applied. A production line can be designed so that if sufficient care is not supplied, it will break down. In agriculture, the fertilizer, the type of seed, and other inputs all effect the stochastic properties of production. Debt contracts frequently include loan covenants that put restrictions on the activities of a borrower, such as working capital requirements.[1] Financial regulation works similarly. Banks have limits on their activities. For example, a bank cannot lend more than a set fraction of its assets to a single borrower. Similarly, money market mutual funds are limited to investing in short-term, safe, commercial paper, and as a

The views expressed in this article do not necessarily represent those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] See Black, Miller, and Posner (1978) for more information on loan covenants as well as connections with bank regulation.

consequence they have a very different risk profile than banks, even though the money market liabilities are close substitutes to some bank liabilities.

In each of the above examples, the principal has some choice over the functional relationship between the agent's actions and his output. In the agricultural case, the connection is through use of inputs. For debt contracts, loan covenants are used to keep a borrower away from potentially dangerous conditions. In the money market and bank regulation examples, the investment restrictions reduce the variance of returns.

The purpose of this article is to work out some of the implications of this line of thought.[2] Only some are explored because there are many different dimensions along which the technology could be changed. Consequently, the analysis is necessarily limited and mainly exploratory. Still, it emphasizes the principles at work and demonstrates why this margin of choice is potentially important.

Most of the issues are illustrated with two examples in which the principal can adjust the technology. The first example gives the principal wide latitude in determining the technology and starkly illustrates how powerful this margin may be. It also demonstrates that this margin strongly affects the optimal contractual form. The second example limits the principal to choosing between only two technologies, but it demonstrates that the principal may be willing to choose a less productive technology, as measured by expected output, because it reduces the incentive problem. In this example, the limited choice among technologies is motivated by the interaction between a principal's decision and inferences made from financial market prices. In particular, the principal's decision to liquidate the firm alters the informativeness of market prices.

1. THE MODEL

In the basic moral hazard model, the agent chooses an action a that combines with a random shock to produce output q. In this article the principal has some control over how a interacts with the randomized shock to produce output. The choice made by the principal is called the *technological* choice and is indexed by i. The relationship between the inputs and the output is described by the conditional probability distribution function $p_i(q|a)$. For simplicity, q is assumed to take on only a finite number of values. Of course, $\sum_q p_i(q|a) = 1$. Finally, based on the output, the principal pays the agent his consumption c.

[2] The moral hazard literature touches on this issue, but the implications and importance of this idea may not be fully appreciated, since the results are scattered across different applications. One important application of this idea is in the task assignment model of Holmstrom and Milgrom (1991) and Itoh (1991). They study how to assign workers to tasks, which in turn affects the technology faced by an agent. One paper, however, that explicitly considers the principal's choice of technology is Lehnert, Ligon, and Townsend (1999).

Preferences

The agent cares about his consumption and his effort. His utility function is $U(c) - V(a)$, with U strictly concave and V increasing in a. The principal is risk neutral so he only cares about the project's surplus, that is, $q - c$.

The principal offers the agent a contract that consists of three items: the principal's choice of technology, what action the agent takes, and how the agent is paid as a function of the output. Formally,

Definition 1 *A contract is a technological index i, a recommended action a, and an output-dependent compensation schedule $c(q)$.*

The agent has an outside opportunity that gives him $\bar{U}$ units of utility. For this problem, this means that the contract has to give him at least that amount of expected utility before he will agree to work for the principal. Therefore, a feasible contract must satisfy

$$\sum_q p_i(q|a)U(c(q)) - V(a) \geq \bar{U}. \tag{1}$$

The point of the moral hazard problem is to generate nontrivial dependence of consumption on output. But from what has been described to this point, there is little reason to expect such a nontrivial dependence. For example, if the agent is risk-averse, that is, U is strictly concave, then an optimal contract fully insures the agent against variations in output, paying the agent a constant wage with the principal absorbing all the risk in output.

Dependence of consumption on output is generated by assuming that the agent's action is *private information*; that is, the principal does not observe it. Consequently, the principal must set up the compensation schedule $c(q)$ to *induce* the agent to take the recommended action. Inducement means here that given $c(q)$ it is in the agent's best interest to take the recommended action. More formally, if the principal wants the agent to take a, then the contract must satisfy the following constraint:

$$\sum_q p_i(q|a)U(c(q)) - V(a) \geq \sum_q p_i(q|\hat{a})U(c(q)) - V(\hat{a}), \quad \forall \hat{a}. \tag{2}$$

A contract is called *feasible* if it satisfies constraints (1) and (2).[3] The principal chooses a feasible contract that maximizes his utility. We can find such a contract by solving the following constrained maximization program:

[3] In mathematics, a *program* refers to the problem of choosing an object that maximizes (or minimizes) an objective function subject to that object satisfying a set of constraints.

Moral Hazard Program

$$\max_{i,a,c(q)} \sum_{q} p_i(q|a)(q - c(q))$$

s.t. (1) and (2).

Analysis

To keep the analysis simple, assume that there are only two actions, a_l and a_h, with $a_l < a_h$. The latter action gives the principal more expected output but gives the agent more disutility. Also assume that in the optimal contract the agent is supposed to take a_h. In this case, there is only one incentive constraint. It is

$$\sum_{q} p_i(q|a_h)U(c(q)) - V(a_h) \geq \sum_{q} p_i(q|a_l)U(c(q)) - V(a_l). \quad (3)$$

Taking the first-order conditions to the program, with (2) replaced by (3), gives

$$-p_i(q|a_h) + \lambda U'(c(q))p_i(q|a_h) + \mu(p_i(q|a_h) - p_i(q|a_l))U'(c(q)) = 0,$$

where λ is the Lagrangian multiplier on constraint (1) and μ is the multiplier on (3). Simplifying gives

$$\forall q, \quad \frac{1}{U'(c(q))} = \lambda + \mu(1 - \frac{p_i(q|a_l))}{p_i(q|a_h)}). \quad (4)$$

Equation (4) describes the relationship between $c(q)$ and the parameters of the problem. Each Lagrangian multiplier is nonnegative and will be positive if its corresponding constraint binds. These variables affect consumption but not as much as the *likelihood ratio* $\frac{p_i(q|a_l)}{p_i(q|a_h)}$ does.

The likelihood ratio determines how c changes with q. If it decreases with q then consumption increases with q. Inspection of (3) reveals why. When this ratio is low, a high level of consumption rewards the agent relatively more for taking a_h than for taking a_l. Conversely when this ratio is high, a low level of consumption punishes the agent relatively more for taking a_l than for taking a_h. The likelihood ratio determines when the principal should use the carrot and when he should use the stick.

Conditions under which the likelihood ratio is decreasing in q include the normal distribution and some others. (For more information see Hart and Holmström [1987] and Jewitt [1988].) Still, most distributions do not satisfy this monotone likelihood property. This lack of robustness has always been a concern for this class of models because most contracts are monotonic as well as being simpler than those that solve the moral hazard program. For example, many sharecropping contracts are linear in output, while salesman are often paid a fixed wage plus a percentage of sales, sometimes with a bonus for hitting performance targets.[4]

By choosing i, the principal is essentially choosing these ratios and, in doing so, directly affects the severity of the incentive constraints. The above analysis suggests that the principal will want to make this ratio high for some outputs in order to use the stick and low for others in order to provide the carrot. If, as was argued earlier, the principal has some control over the properties of the technology, then this will strongly affect technological design as well as compensation schedules. The following examples are designed to explore this idea.

2. A SIMPLE EXAMPLE

This example examines the extreme case where the principal can *only* control the probability distribution of output for the low action. The only restriction on these probabilities is that the chosen distribution still needs to produce the same expected output. In this problem, it is assumed that the solution is such that the principal wants to implement the high action.

Each choice of the technology index i corresponds to a choice of the entire probability distribution over q given a_l. For this reason, it is convenient to drop explicit reference to i and just let the principal choose the entire function $p(q|a_l)$.

The programming problem for this example is

$$\max_{p(q|a_l)\geq 0, c(q)} \sum_q p(q|a_h)(q - c(q))$$

subject to the participation constraint

$$\sum_q p(q|a_h)U(c(q)) - V(a_h) \geq \bar{U}, \tag{5}$$

the incentive constraint

$$\sum_q p(q|a_h)U(c(q)) - V(a_h) \geq \sum_q p(q|a_l)U(c(q)) - V(a_l), \tag{6}$$

and the constraints on technology

$$\sum_q p(q|a_l) = 1, \text{ and} \tag{7}$$

$$\sum_q p(q|a_l)q = \bar{q}, \tag{8}$$

[4] See Townsend and Mueller (1998), however, for a description of sharecropping contracts that are formally linear but in practice are much more complicated.

where $\bar{q}$ is the expected output amount that the distribution must produce.

The first set of first-order conditions is identical to (4). It is

$$\forall q, \quad \frac{1}{U'(c(q))} = \lambda + \mu(1 - \frac{p(q|a_l))}{p(q|a_h)}). \tag{9}$$

The second set of incentive constraints comes from taking the derivative with respect to $p(q|a_l)$. Letting η be the Lagrangian multiplier on (7) and ν the multiplier on (8), these conditions are

$$\forall q, \quad -U(c(q))\mu + \eta + \nu q \leq 0, \quad (= 0 \text{ if } p(q|a_l) > 0). \tag{10}$$

There is not necessarily an interior solution to this problem, so it is possible that (10) holds at an inequality.[5]

Equation (10) implies that $c(q)$ is monotonic over q such that $p(q|a_l) > 0$. Whether it is increasing or decreasing depends on the sign of ν. It is important to note that this does not mean that $c(q)$ is monotonically increasing everywhere. The result only applies for outputs for which $p(q|a_l)$ is chosen to be strictly positive. For outputs in which $p(q|a_l) = 0$, equation (9) implies that

$$\frac{1}{U'(c(q))} = \lambda + \mu. \tag{11}$$

Thus, consumption is the same for all of these values of output. Comparing (11) with (9) and noting that $\mu > 0$ reveals that consumption is higher for values of q that satisfy $p(q|a_l) = 0$ than for values that satisfy $p(q|a_l) > 0$.

These results are summarized by the following proposition:

Proposition 1 *The optimal contract is characterized by the following features: i) Consumption is a constant for all values of q such that $p(q|a_l) = 0$; ii) Consumption is monotonically increasing or decreasing in q over values of q such that $p(q|a_l) > 0$; iii) Consumption levels for q such that $p(q|a_l) = 0$ are higher than consumption levels for q such that $p(q|a_l) > 0$.*

At first glance, the monotonicity result is appealing because many contracts are monotonic. But since monotonicity only applies to outputs for which $p(q|a_l) > 0$, the degree of monotonicity will depend on the range of values of output with this property. As the following analysis demonstrates, there only needs to be two such outputs.

Proposition 2 *Let $Q = \{q|p(q|a_l) > 0\}$. There exists a solution in which there are no more than two outputs with $q \in Q$.*

Proof: The variables $p(q|a_l)$ *only* affect the right-hand side of the incentive constraint, (6), and the constraints on the technological design, (7) and

[5] Notice that it has been implicitly assumed that $c(q)$ will be an interior solution.

(8). The lower the value of the right-hand side of (6) the less binding the incentive constraint will be and the better the consumption schedule that can be implemented. Consequently, for a given consumption schedule a solution will solve the following program for the principal,

$$\min_{\forall q \in Q, p(q|a_l) \geq 0} \sum_{q \in Q} p(q|a_l)U(c(q))$$

subject to

$$\sum_{q \in Q} p(q|a_l) = 1, \tag{12}$$

$$\sum_{q \in Q} p(q|a_l)q = \bar{q}. \tag{13}$$

This program is a linear program. If a solution exists to a linear program, which is true by assumption here, then a basic feasible solution exists. A basic feasible solution is one in which the number of non-zero valued variables is less than or equal to the number of constraints. In this problem there are only two constraints so there is a solution where all but two variables are necessarily zero. So if Q had more than two elements, their probabilities would be zero and they would not be in Q. **Q.E.D.**

The $p(q|a_l)$ function is set to make the utility from taking the low action as low as possible. Consequently, the program puts as much weight as possible on the lowest values of $c(q)$. Proposition 2 shows that there needs to be only two such points, which helps us characterize the compensation schedule. Since only two outputs are in Q, $p(q|a_l) = 0$ for all other values of q. So by (9) the value of consumption for these outputs is a constant. The compensation schedule in this problem is a wage except for the one or two outputs for which $p(q|a_l) > 0$. Which one or two outputs will be chosen cannot be determined without solving for the multipliers and all the other variables.

The goal of this exercise is to demonstrate the striking effect that technological choice by the principal can have on the properties of an optimal contract. Still, the optimal contract with its punishments on two levels of intermediate outputs does not look like contracts used in practice. Indeed, the analysis suggests that the incentive problem will be relatively minor since the low levels of consumption are only paid infrequently for the two outputs with $p(q|a_l) > 0$. Furthermore, the contract looks a lot more like a wage contract than the pay for performance contracts the theory was designed to describe. Fortunately, as the next section demonstrates, adding a small modification to the problem generates an optimal contract that is much more appealing on "realism" grounds. In particular, it will be monotonically increasing and relatively simple.

A Monotonicity Extension

Assume now that the agent can costlessly destroy output and that the principal does not know if he destroyed any output. All the principal observes is what is left. The agent does not consume any of the destroyed output. This assumption adds another source of private information to the problem; one that is easy to analyze. The ability to destroy output requires that the compensation schedule $c(q)$ be weakly monotonically increasing in output.[6] Otherwise, the agent could destroy some output and claim the higher consumption. If there are n possible output realizations and q_j refers to the jth output, then this assumption requires adding the following constraints to the program.

$$\text{for } j = 1, ..., n-1, \quad c(q_j) \leq c(q_{j+1}). \tag{14}$$

An analysis of the amended program is not too different from that of the earlier program. The first-order conditions analogous to (9) only have some additional multipliers for the monotonicity constraints. The remaining first-order conditions are the same as (10).

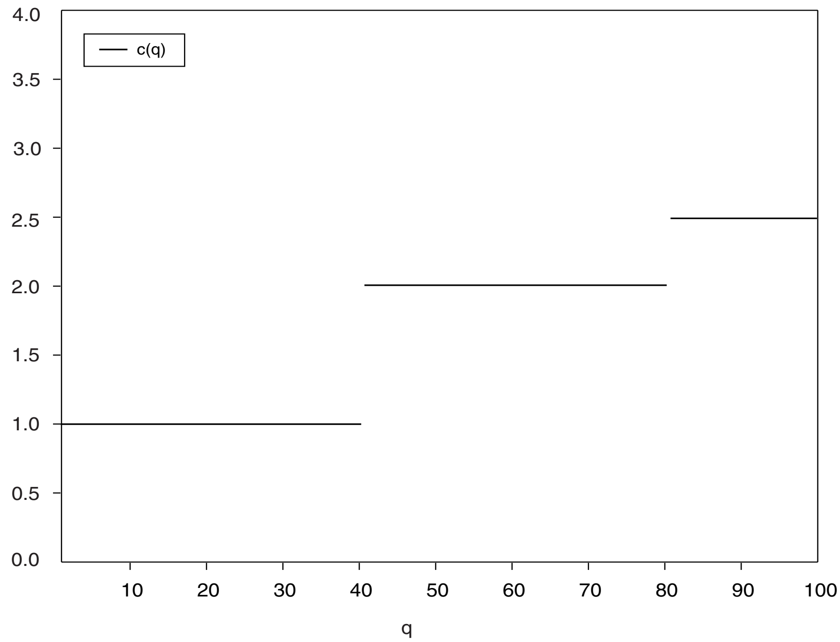
The addition of the monotonicity constraints prevents the solution to the earlier program from being optimal. If the agent produced the one or two outputs that correspond to $p(q|a_l) > 0$—the outputs with the lowest consumption—he could simply destroy some of the output and receive a higher level of consumption that goes with his new lower output.

The optimal contract to the program with the monotonicity constraints retains some of the same features as the solution to the earlier program. There are still only one or two outputs for which $p(q|a_l) > 0$. If there are two such outputs, they split the contract into three distinct regions: one region less than the lower of the two outputs, a middle region between the two outputs, and a third region above the higher output. In the lower range consumption is a constant and $p(q|a_l) = 0$, in the second region consumption is also a constant but higher than the consumption of the first point as well as higher than the consumption in the first region, and in the third region consumption is yet again a constant but at an even higher level. The contract is a step function with three steps. It resembles a contract with a wage and two levels of bonuses (or, a contract with a wage and one bonus level and one lower wage level for poor performance, etc.). Figure 1 illustrates. The point is that the contract keeps the desired monotonicity property and is relatively simple.

Proposition 3 *An optimal contract to the program with the monotonicity constraint is characterized by a compensation schedule that is: i) monotonic; ii) characterized by the three regions described above; and iii) consumption for*

[6] Adding this source of private information to the standard model is enough to generate monotonic consumption but the optimal contract can still be complicated with lots of contingencies. As will be shown, the combination of technological choice and monotonicity also simplifies the contract in important ways.

Figure 1 Optimal Compensation Schedule



Notes: Example of what an optimal compensation-sharing rule might be. It is broken into three regions with a linear portion in each region. Consumption is weakly monotonically increasing.

each of two outputs that separate the two regions is equal to the consumption in one of the adjacent regions.

Proof: *i)* Follows directly from the monotonicity constraints (14). To prove *ii),* use the same argument as before to argue that there are only two output levels for which $p(q|a_l) > 0$. These points are the boundaries. Below, between, and above, there is full insurance within each region because if there was not, consumption could be smoothed, which would deliver the risk averse agent the same utility at lower cost to the principal and not affect incentives. Finally, for *iii),* if consumption of either of these two points was not equal to consumption in one of the adjacent regions, then consumption in the regions could be made closer together without altering the agent's utility. Analogous to the argument in *ii),* this is a less expensive way for the principal to provide utility to a risk-averse agent. **Q.E.D.**

There are two lessons to this example. First, the modification generates contracts that are appealing on observational grounds. Second, the principal

Table 1 Probability Distributions of Output and Expected Output for Each Action

	0	1	2	$E(q\|a)$
a_l	1/3	1/3	1/3	1
a_h	ε	1/3	2/3-ε	5/3 - 2ε

Notes: The parameter ε is nonnegative.

will try and design the technology so that if the agent slacks off (takes a_l), certain outputs will be very likely. In particular, he wants off-equilibrium probability distributions to be as revealing as possible when the agent slacks off.

3. LIQUIDATION EXAMPLE

The next example demonstrates that because of incentives, sometimes the principal prefers a less-productive technology, as measured by expected output. The reason for this counterintuitive result is that sometimes a less-productive technology alters the likelihood ratios in such a way that the incentive constraint is relaxed enough to outweigh the loss in output.

There are three outputs, $q_1 = 0$, $q_2 = 1$, and $q_3 = 2$. As before, there are only two actions, a_l and a_h. The production function is described in Table 1. The exercise is to assume that $\varepsilon \geq 0$ and then to vary it to illustrate how that affects the solution. Expected output for the high action is higher than that of the low output for any $\varepsilon < 1/3$.

The literal description of this problem is different from the standard model. Mathematically, however, it will be identical to the moral hazard program. The description is useful because it better motivates the example.

Now, assume that q represents an intermediate valuation of the project's long-term prospects. The principal does *not* observe q. There is a market, however, that trades securities based on the long-term value of the project. The market observes q and prices its securities accordingly. Alternatively, market participants have varying sources of information that are combined and communicated, however imperfectly, through the market price. Importantly, the principal observes the market price and makes an inference about the true q from it.

So far, the problem is no different than that of the standard model; the principal does not observe q directly but he can infer it from the market price of the security. Now, however, the principal has the option of liquidating the

Table 2 Probability Distributions of Output and Expected Output for Each Action

	0	1	2	$E(q\|a)$
a_l	0	2/3	1/3	4/3
a_h	0	$1/3+\varepsilon$	$2/3-\varepsilon$	$5/3-\varepsilon$

Notes: Principal liquidates whenever the traded security indicates that $q = 0$ or $q = 1$.

project right after q is created (and observed and traded upon by the market). If he liquidates, the value of the project becomes one.

Markets, as always, are forward looking. In this context, this means that the market takes into account the effect of the principal's liquidation strategy on the value of the project. For example, if the strategy is to liquidate the project when the market price indicates that $q = 0$ or $q = 1$, then the market will trade the security at a price that indicates that $q = 1$. Indeed, under this liquidation strategy the security would never trade at a price of zero![7]

The problem for the principal here is to decide on the best liquidation strategy. If he does not liquidate, the technology is the one described in Table 1. If the principal liquidates when the market price indicates $q = 0$ or $q = 1$, then the principal has essentially chosen the probability distributions to be those described in Table 2. No other feasible liquidation strategy is preferable, so the other ones are not explicitly considered.

In the liquidation case, $q = 1$ is not literally the amount produced since the agent may have produced $q = 0$, but liquidating turns it into an output level of one. Because the principal chooses whether to liquidate, the principal is essentially choosing between the probability distribution in Table 1 and the one in Table 2. Thus, the program has been mapped into the mathematical structure of the moral hazard program. Furthermore, for $\varepsilon > 0$ expected output in Table 1 is less than that in Table 2 for both actions. It is in this sense that the first technology is technologically inferior to the second technology. Yet, as we will shortly see, the first technology is sometimes superior when incentive considerations are taken into account.

The two problems can be compared by merely contrasting the incentive constraints. As before, assume that the principal wants to implement a_h. The no-liquidation incentive constraint, i.e., the one from choosing the technology

[7] Related, a liquidation strategy of liquidating only when $q = 0$ would create an equilibrium existence problem. Under this strategy, $q = 0$ would never be observed because the price would be one. But if the price is one, the supervisor would never liquidate! Liquidating when the market price is zero or one avoids this circularity.

in Table 1, is

$$(1/3 - \varepsilon)(U(c(q_3) - U(c(q_1))) \geq V(a_h) - V(a_l). \tag{15}$$

The liquidation incentive constraint, i.e., the one from choosing the technology in Table 2, is

$$(1/3 - \varepsilon)(U(c(q_3) - U(c(q_2))) \geq V(a_h) - V(a_l). \tag{16}$$

The only difference between the two constraints is the replacement of $U(c(q_1))$ in (15) with $U(c(q_2))$ in (16). This should not be surprising. Output q_1 is not produced in the liquidation case, so it is not a factor in that case.

The striking feature of this example is that for small enough values of ε the principal prefers the no-liquidation technology even though the liquidation technology produces a higher expected output (for either action). The best way to see this is to analyze the limiting case where $\varepsilon = 0$. Consider the contract where the principal chooses the no-liquidation technology, sets $c(q_1)$ equal to its minimum level and sets $c(q_2) = c(q_3)$. Assuming that $c(q_1)$ can be set low enough so that the incentive constraint (15) holds, then this solution provides full insurance. Indeed, the incentive constraint (15) does not bind. Because the agent chooses a_h, the principal has not given any output up by not liquidating and the low consumption for producing $c(q_1)$ is a very powerful way of preventing the agent from choosing a_l.[8] In contrast, if the principal liquidated the project with this consumption schedule, the agent would take a_l because he would never suffer the penalty from producing q_1.

As ε gradually increases from zero, the principal starts foregoing output by not liquidating. Still, for small values of ε the incentive effect of setting $c(q_1)$ to a low value outweighs the loss in output as well as any cost to the agent from producing q_1. (As ε grows, the optimal contract will no longer provide constant consumption over q_1 and q_2, and $c(q_1)$ will increase.) Consequently, the "inferior" no-liquidation technology is preferred to the liquidation technology for incentive reasons. Of course, as ε gets large enough, the output effect will dominate the incentive effect and only then will the principal prefer the liquidation strategy.

4. CONCLUSION

This article worked through two examples to illustrate the importance of technological design on moral hazard. The first example gave the principal wide latitude in designing the probability distribution. It illustrated the mechanics of the approach and demonstrated that large effects on optimal compensation schedules were possible. The second example studied a problem in which the

[8] The likelihood ratio is infinite in this case. What this is indicating here is that consumption is not an interior solution and, in this case, is set to its lower bound.

liquidation strategy affected the informativeness of output. It demonstrated that sometimes the principal was willing to forgo output in return for a more informative distribution of output.

The main conceptual difficulty in these examples is determining how much latitude to give the principal in setting the probability distributions. What choice to offer the principal will depend on the application. The second example, with its problem of inferring true output from a market price, had a natural way of limiting the principal's control over the technology. Other applications will suggest different dimensions to this choice. Regardless of the application, what the analysis makes clear is that the technological design dimension to the moral hazard problem is an important one. It affects the surplus for the principal and the shape of the compensation schedule.

REFERENCES

Black, Fischer, Merton H. Miller, and Richard A. Posner. 1978. "An Approach to the Regulation of Bank Holding Companies." *Journal of Business* 51 (3): 379–412.

Hart, Oliver and Bengt Holmström. 1987. "The Theory of Contracts." In *Advances in Economic Theory, Fifth World Congress*. Cambridge, England: Cambridge University Press.

Holmström, Bengt and Paul Milgrom. 1991. "Multitask Principal-Agent Analysis." *Journal of Law, Economics & Organization* 7: 24–52.

Itoh, Hideshi. 1991. "Incentives to Help in Multi-Agent Situations." *Econometrica* 59 (3): 611–37.

Jewitt, Ian. 1988. "Justifying the First-Order Approach to Principal-Agent Problems." *Econometrica* 56 (5): 1177–90.

Lehnert, Andreas, Ethan Ligon, and Robert M. Townsend. 1999. "Liquidity Constraints and Incentive Constraints." *Macroeconomic Dynamics* (1): 1–47.

Townsend, Robert M. and Rolf A. E. Mueller. 1998. "Mechanism Design and Village Economies: From Credit to Tenancy to Cropping Groups." *Review of Economic Design* 1 (1): 119–72.

Trend Inflation, Firm-Specific Capital, and Sticky Prices

Andreas Hornstein and Alexander L. Wolman

Research on monetary policy, both at academic and monetary policy institutions, has increasingly been performed within an analytical framework that assumes limited nominal price adjustment, "sticky prices" for short. At the heart of much of this analysis is a so-called New Keynesian (NK) "expectational" Phillips curve that relates current inflation, π_t, to expected future inflation and the deviation of marginal cost from trend $\hat{s}_t$:

$$\pi_t = \beta E_t \pi_{t+1} + \xi \hat{s}_t, \quad (1)$$

with $\beta, \xi > 0$. Empirical estimates of the coefficient on the marginal cost term, ξ, in this NK Phillips curve tend to be positive but small in absolute value, e.g., Sbordone (2002) and Galí and Gertler (1999).[1] This represents a problem for the sticky-price framework since the coefficient ξ is directly related to the frequency with which nominal prices are assumed to be adjusted: the coefficient is smaller the less frequently prices are adjusted. Within standard sticky-price models, estimated values of ξ imply that prices are adjusted less than once per year. This macro estimate of price stickiness is implausibly high from the perspective of the micro estimates surveyed in Wolman (forthcoming).

It has been conjectured widely that nominal rigidities, such as sticky prices, have more persistent real effects if they interact with real rigidities. For example, the basic NK Phillips curve (1) has been derived for an environment with nominal frictions, but essentially no real rigidities: firms rent factors

For helpful comments we would like to thank Andrew Foerster, Bob Hetzel, Ned Prescott, and Pierre Sarte. The views expressed in this article are those of the authors and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] Expression (1) is derived in Woodford (2003, ch. 3) for an economy with Calvo-type sticky prices. Woodford's (2003) textbook presents a unified framework for thinking about monetary policy based on sticky-price models. For a critical review of this line of research, see Green (2005).

of production—capital and labor—in frictionless markets. Now, suppose that there is a real rigidity in addition to the sticky prices. In particular, assume that capital is specific to individual firms, and it is costly for these firms to adjust their capital stock. Introducing firm-specific capital adjustment costs into sticky-price models substantially complicates the analysis, yet Woodford (2005) manages to derive an almost closed-form solution to this problem. In particular, Woodford (2005) again derives an NK Phillips curve of the form (1), but now the marginal cost coefficient, ξ, depends not only on the extent of price stickiness, but also on the magnitude of capital adjustment costs: the coefficient is smaller the less frequently prices are adjusted and the more costly it is to adjust capital. Thus low estimated values of ξ do not necessarily imply a high degree of price stickiness.

Woodford's (2005) clean analytical solution of the modified NK Phillips curve does come with a cost. His approach is based on the linear approximation of an economy with Calvo-type nominal price adjustment around an equilibrium with zero average inflation. The assumption of zero average inflation makes the theoretical analysis of the firm aggregation problem possible, yet it is not empirically plausible. Even though in recent years inflation has been remarkably stable in many industrialized countries, average inflation has been positive. Furthermore, most estimates of the NK Phillips curve use data from periods of moderate inflation. Thus, it is important to know whether the behavior of these models is sensitive to the steady-state inflation rate.[2]

In this article we evaluate the relative impact of positive average inflation versus zero inflation in an economy with nominal rigidities and firm-specific capital adjustment costs. Unlike Woodford (2005), we model nominal rigidities as Taylor-type staggered price adjustment, and not as Calvo-type probabilistic price adjustment. This approach is necessary since at this time there are no aggregation results for our economic environment with Calvo-type pricing and nonzero average inflation. We show that for small values of positive average inflation, the Taylor principle, which states that a central bank should increase the nominal interest rate more than one-for-one in response to a deviation of inflation from its target, is no longer sufficient to guarantee that monetary policy does not become a source of unnecessary fluctuations in our economy.

The fundamental difficulty with incorporating firm-specific capital into a model with sticky prices is that firm-specific capital can amplify the heterogeneity associated with price stickiness. With Calvo price setting, firms face a constant exogenous probability of being able to readjust their price. If there

[2] Furthermore, even though overall inflation has been low and stable, trends have remained in disaggregated measures of prices—for example, services prices have a positive trend and durable goods prices have a negative trend. This means that in a multi-sector model with zero inflation, the steady state would involve trends in individual nominal prices and thus a nondegenerate distribution of prices across sticky-price firms (Wolman 2004).

are no state variables specific to the firm (other than price), then all firms that adjust in a given period choose the same price. In that case, even though the true distribution of prices is infinite, it is possible to summarize the relevant distribution with just a small number of state variables.[3] If instead capital is firm specific, firms that adjust in the same period generally do not have the same capital stock. Their marginal cost is not the same, and in general they will not choose the same price. Thus, combining Calvo pricing and firm-specific capital appears to lead to an intractable model.

The model is intractable in its exact form, but Sveen and Weinke (2004) and Woodford (2005) have shown how to derive a tractable linear approximation to the model, under the assumption that the average inflation rate is zero. The key to these derivations is the fact that in the zero-inflation steady state there is no heterogeneity: all firms charge the same price.

Given the tractability problem, there is little hope of being able to learn how the Calvo model with firm-specific capital behaves away from a zero-inflation steady state. Fortunately, there is another class of sticky-price models that remains tractable when combined with firm-specific capital. The staggered pricing framework associated with Taylor (1980) assumes that there are J different types of firms; each period a fraction $1/J$ of firms adjusts their prices, and their prices remain fixed for J periods. Firm-specific capital presents no problems in the Taylor model, because it remains the case that all firms that adjust in a given period enter with the same capital stock and thus will choose the same price.

We solve the linear approximation to the Taylor model numerically and ask whether the model's dynamics are sensitive to the steady-state inflation rate around which we linearize.[4] We find that a small but positive inflation rate can have a big impact on the set of parameters for monetary policy rules and investment adjustment costs for which a rational expectations (RE) equilibrium is unique.[5] If the equilibrium is not unique, that is, there is equilibrium indeterminacy, then possible equilibrium outcomes can depend on shocks that do not constrain the resource feasible allocations in an economy. In these equilibria self-fulfilling expectations that coordinate on such nonfundamental shocks, known as "sunspots," introduce unnecessary fluctuations into the economy.

---

[3] We say the true distribution is infinite because a positive fraction of firms charges a price set arbitrarily many periods in the past.

[4] Others have worked with the Taylor model with firm-specific capital; see, for example, Coenen and Levin (2004) and de Walque, Smets, and Wouters (2004). They have not studied the role of steady-state inflation.

[5] Since we are studying linear approximations of equilibria, all of our statements have to be understood as applying to local properties of the equilibria for small deviations from the steady state. Wolman and Couper (2003) discuss the potential pitfalls of this type of analysis, especially as it relates to statements about the uniqueness of equilibrium.

In standard sticky-price models, monetary policy rules that set the nominal interest rate in response to deviations of inflation from its target value achieve a unique RE equilibrium, if they follow the Taylor principle. The principle states that the nominal interest rates increase more than one-for-one with an increase of the inflation rate. This policy response does not have to be very big, as long as it is greater than one. We show that in the sticky-price model with firm-specific capital, positive steady-state inflation generally increases the region of the parameter space for which there is indeterminacy of equilibrium. In other words, for the same magnitudes of price-stickiness and capital-adjustment costs, monetary policy has to be much more responsive to deviations of inflation from its target in order to maintain a unique RE equilibrium outcome. These results suggest that it may be misleading to interpret history and make policy recommendations based on findings from the zero steady-state inflation case. Our results complement those in Sveen and Weinke (2005), who show that moving from a rental market to firm-specific capital leads to a larger region of the parameter space for which there is indeterminacy of equilibrium when steady-state inflation is zero.

In Section 1 we describe the economy with firm-specific capital adjustment cost and the two types of sticky prices: Calvo-type and Taylor-type nominal price setting. In Section 2 we outline how Woodford (2005) solves the aggregation problem for Calvo-type pricing and derives the modified NK Phillips curve. In Section 3 we characterize the economy with Taylor-type pricing, and in Section 4 we study the impact of capital adjustment costs and nonzero average inflation on the economy with Taylor-type pricing.

1. STICKY-PRICE MODELS WITH FIRM-SPECIFIC CAPITAL

This section presents the common features of Calvo and Taylor sticky-price models. There is an infinitely lived representative household and a continuum of differentiated firms. The firms act as monopolistic competitors in their differentiated output markets, but they are competitive in their differentiated labor markets. The differentiated output goods of the firms are used to produce an aggregate output good in a competitive market. The aggregate output good can be used for consumption or investment. Firms use investment goods to augment their firm-specific capital stocks, subject to capital adjustment costs. Firms set the nominal price of their differentiated output good, and only infrequently do they have the opportunity to adjust their nominal price.

The Representative Household

The household values consumption, c_t, and experiences disutility from the supply of differentiated labor to a continuum of markets, $h_t(j)$. The expected present value of utility is

$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{c_t^{1-\sigma} - 1}{1-\sigma} - \gamma \int_0^1 \frac{h_t(j)^{\nu+1}}{\nu+1} dj \right\}, \tag{2}$$

with discount factor, β. Period utility is an increasing (decreasing) concave (convex) function of consumption (work time), $\sigma, \nu, \gamma > 0$. The representative household owns shares in the continuum of firms and holds nominal bonds. The household's budget constraint is

$$\begin{aligned} P_t c_t + \int_0^1 Q_t(j)\, a_{t+1}(j)\, dj + B_{t+1} &= \int_0^1 W_t(j)\, n_t(j)\, dj \\ &+ \int_0^1 \left[Q_t(j) + D_t(j)\right] a_t(j)\, dj + (1+i_t)\, B_t, \end{aligned} \tag{3}$$

where P_t is the nominal price of the aggregate output good, $Q_t(j)$ is the nominal price of a share in firm j, $W_t(j)$ is the nominal wage paid by firm j, $D_t(j)$ is the nominal dividend paid by firm j, i_t is the nominal interest rate on nominal bond holdings B_t, and $a_t(j)$ is the household's firm-share holdings.

Optimal choice of work effort implies the following firm-specific labor supply functions

$$w_t(j) = \gamma h_t(j)^{\nu} / \lambda_t, \tag{4}$$

where $w_t(j) = W_t(j) / P_t$ is the real wage paid by firm j, and λ_t is marginal utility of consumption

$$\lambda_t = c_t^{-\sigma}. \tag{5}$$

Optimal asset and bond holdings imply the following Euler equations for bonds and firm shares

$$1 = E_t \left[\beta \frac{\lambda_{t+1}}{\lambda_t} \frac{1+i_t}{P_{t+1}/P_t} \right] \text{ and} \tag{6}$$

$$1 = E_t \left[\beta \frac{\lambda_{t+1}}{\lambda_t} \frac{\left[Q_{t+1}(j) + D_{t+1}(j)\right]/Q_t(j)}{P_{t+1}/P_t} \right]. \tag{7}$$

The representative household chooses consumption such that the household is indifferent between consuming slightly more, with a corresponding reduction in asset holdings, and consuming slightly less, with a corresponding increase in asset holdings. The Euler equations embody this indifference. In an equilibrium, the representative household owns all firms, $a_t(j) = 1$.

Aggregate Output

The aggregate output, y_t, is produced from the continuum of differentiated inputs,$y_t\,(j)$, using a constant-elasticity-of-substitution production function

$$y_t = \left[\int_0^1 y_t(j)^{(\theta-1)/\theta} dj \right]^{\theta/(\theta-1)}, \tag{8}$$

where $\theta \geq 1$ denotes the elasticity of substitution between goods. This is the Dixit-Stiglitz (1977) formulation used by Blanchard and Kiyotaki (1987). Production is competitive and given nominal prices, $P_t\,(j)$, for the differentiated inputs, cost minimization implies the following nominal price index/marginal cost for the aggregate output

$$P_t = \left[\int_0^1 P_t\,(j)^{1-\theta}\, dj \right]^{\frac{1}{1-\theta}}. \tag{9}$$

Given aggregate output, the demand for a differentiated good is a function of its relative price, $p_t\,(j) \equiv P_t\,(j)\,/P_t$,

$$y_t\,(j) = p_t\,(j)^{-\theta}\, y_t. \tag{10}$$

Aggregate output can be used for consumption or for the accumulation of firm-specific capital by the producers of differentiated goods, $x_t\,(j)$. Market clearing for goods implies that aggregate output equals the sum of consumption and aggregate investment

$$y_t = c_t + \int_0^1 x_t\,(j)\,dj. \tag{11}$$

Firms

The differentiated goods are produced by a continuum of monopolistically competitive firms, and these are the same firms to which the household supplies labor. The differentiated goods are produced using the inputs capital and labor, both of which are specific to each firm. The differentiated firms can adjust the nominal prices they set for their product only infrequently.

Production

Production is constant-returns-to-scale; in particular, we assume that the production function is Cobb-Douglas:

$$y_t\,(j) = k_t\,(j)^{\alpha}\,[A_t h_t\,(j)]^{1-\alpha}\,; \tag{12}$$

$y_t\,(j)$ is firm j's output in period t, and $k_t\,(j)$ and $h_t\,(j)$ are, respectively, the capital input and labor input used by firm j in period t. There is an aggregate productivity disturbance given by A_t. At the beginning of period t, firm j's capital input is predetermined as a result of the investment decision firm j made in period $t-1$. Furthermore, there are convex costs of changing the capital stock, which we will specify further below. Labor is hired in competitive markets, but because households receive distinct disutility from the labor they provide to each firm, the wage can differ across firms.[6]

In order to change its capital stock from k_t in period t to k_{t+1} in period $t+1$, a firm needs x_t units of the aggregate output good

$$x_t(j) = k_t(j)\, G\left[k_{t+1}(j)/k_t(j)\right]. \tag{13}$$

The firm incurs capital adjustment costs determined by the increasing and convex function, $G(k_{t+1}/k_t)$. As in Woodford (2005), $G(1)=\delta$, $G'(1)=1$ and $G''(1)=\epsilon_\psi$, where $\epsilon_\psi > 0$ is a parameter. If the firm exactly replaces depreciated capital, then the marginal investment cost is one, but if the firm increases its capital stock, then the marginal cost of each additional unit of capital is greater than one and increasing with the rate at which the capital stock increases.

Prices

Firms in the model face limited opportunities for price adjustment. In particular, we assume that any firm faces an exogenous probability of adjusting its price in period t and that the probability may depend on when the firm last adjusted its price. The key notation describing limited price adjustment will be a vector Φ (possibly with a countably infinite number of elements); the s^{th} element of Φ, called ϕ_s, is the probability that a firm adjusts its price in period t, conditional on its previous adjustment having occurred in period $t-s$.

There is a time invariant distribution of firms according to when they last adjusted their price, since the price-adjustment probabilities do not vary with time. Let ω_s denote the fraction of firms in period t, charging prices set in periods $t-s$, with the corresponding vector, Ω. Given the price-adjustment probabilities, the time invariant distribution satisfies

$$\begin{aligned} \omega_s &= \left(1-\phi_s\right)\omega_{s-1}, \text{ for } s = 1, 2, ..., \text{and} \\ \omega_0 &= 1 - \sum_{s=1}^{J-1} \omega_s. \end{aligned} \tag{14}$$

The most common pricing specifications in the literature are those first described by Taylor (1980) and Calvo (1983). Taylor's specification is one of uniformly staggered price setting: every firm sets its price for J periods, and at any point in time a fraction $1/J$ of firms charge a price set s periods ago. The J-element vector of adjustment probabilities for the Taylor model is $\Phi = [0, ..., 0, 1]$, and the J-element vector of fractions of firms is

[6] Labor market clearing is implicitly imposed by not differentiating between the labor supplied to the j^{th} type of firm and the labor demanded by the j^{th} type of firm.

$\Omega = [1/J, 1/J, ..., 1/J]$. In contrast, Calvo's specification involves uncertainty about when firms can adjust their price. No matter when a firm last adjusted its price, it faces a probability ϕ of adjusting. Thus, the infinite vector of adjustment probabilities is $\Phi = [\phi, \phi, ...]$, and the infinite vector of fractions of firms is $\omega_s = \phi (1-\phi)^s$, $s = 0, 1, ...$

Firm Value

We assume that a firm pays out each period's profits as dividends to its shareholders:

$$d_t (j) = p_t (j)\, y_t (j) - w_t (j)\, h_t (j) - x_t (j)\,. \tag{15}$$

Conditional on the firm's relative price, $p_t (j)$, sales, $y_t (j)$, are determined by the demand curve (10). The firm's demand for labor is

$$h (j) = H [y (j) , k (j) , A] = \left[\frac{y (j)}{k (j)^{\alpha}}\right]^{1/(1-\alpha)} A^{-1}. \tag{16}$$

The rationale behind solving for labor input in (16) is that in period t the firm's capital stock is predetermined, and thus the labor input it must employ is determined by its technology, given the level of demand, $y_t (j)$. Conditional on the available capital stock, the marginal (labor) cost of output is then

$$s_t (j) = \frac{1}{1-\alpha} \frac{w_t (j)\, y_t (j)}{h_t (j)}. \tag{17}$$

Investment is determined by the capital stock the firm operates at the beginning of the period and the capital stock the firm plans to operate in the next period, equation (13). With some abuse of notation we can rewrite the real dividends of a firm as a function of its idiosyncratic state and control variables: the relative price and the beginning-of-period and end-of-period capital stocks,

$$d_t (j) = d_t \left[p_t (j) , k_t (j) , k_{t+1} (j)\right]. \tag{18}$$

The dependence on the aggregate state of the economy (aggregate demand, productivity, wages) is subsumed in the time subscript t for the function d.

The firms maximize the discounted expected present value of future dividends. The relevant discount factor is the representative household's intertemporal marginal rate of substitution, since the firms are owned by the household,

$$\max E_t \sum_{\tau=0}^{\infty} \beta^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} d_{t+\tau} (j)\,. \tag{19}$$

Let $v_t (p_{-1}, k, j)$ denote the value of a firm with relative price, p_{-1}, in the last period and beginning of period capital stock k. Let j denote when the firm last adjusted its nominal price. If $j = 0$, the firm can adjust its nominal price in the current period, that is, p_{-1} does not affect the firm's value and we write $v_t(k, 0)$. We can write the value of a firm as a function of its own state variables recursively,

$$\begin{aligned} v_t(k_t, 0) &= \max_{p_t^*, k_{t+1}} \left\{ d_t\left(p_t^*, k_t, k_{t+1}\right) + E\left[\beta \frac{\lambda_{t+1}}{\lambda_t} \left\{\phi_1 v_{t+1}(k_{t+1}, 0) \right.\right.\right. \\ &\quad \left.\left.\left. + (1-\phi_1) v_{t+1}\left(p_t^*, k_{t+1}, 1\right)\right]\right\}, \qquad (20) \\ v_t(p_{t-1}, k_t, j) &= \max_{k_{t+1}} \left\{ d_t(p_t, k_t, k_{t+1}) + E\left[\beta \frac{\lambda_{t+1}}{\lambda_t} \left\{\phi_{j+1} v_{t+1}(k_{t+1}, 0) \right.\right.\right. \\ &\quad \left.\left.\left. + \left(1-\phi_{j+1}\right) v_{t+1}(p_t, k_{t+1}, j+1)\right]\right\}, \qquad (21) \\ &\text{and } p_t = p_{t-1} \frac{P_{t-1}}{P_t} \end{aligned}$$

Note that for Calvo pricing, $\phi_j = \phi$, and therefore $v_t(p_{-1}, k, 1) = v_t(p_{-1}, k, j)$ for all $j \geq 1$. On the other hand, for Taylor pricing the firm value functions are only defined for $j \leq J - 1$, since $\phi_J = 1$.

Government Policy

We assume that there is neither taxation nor government spending. Monetary policy chooses a desired steady-state level for the inflation rate, π^*. Given the steady-state real interest rate, $1/\beta$, the steady-state nominal interest rate, i^*, consistent with the inflation rate, π^*, is

$$1 + i^* = \frac{1 + \pi^*}{\beta}. \qquad (22)$$

Monetary policy is assumed to set the period nominal interest rate in response to deviations of the inflation rate and output from their respective steady-state values,

$$i_t = i^* + f_\pi \left[P_t/P_{t-1} - \left(1 + \pi^*\right)\right] + f_y \left[\frac{y_t - y^*}{y^*}\right]. \qquad (23)$$

2. THE CALVO MODEL

We now outline how the equilibrium of the economy with Calvo pricing can be characterized for a log-linear approximation around a steady state with zero inflation. In particular, we show that despite the fact that firms differ according to their relative prices and their capital stocks, calculating simple averages over all these firms yields a consistent aggregation. We do not provide a complete characterization of the equilibrium; for this we refer the reader to Woodford (2005). Although our results below on equilibrium indeterminacy are for the Taylor model, we present the equilibrium characterization for the Calvo model because it helps to explain the appeal of firm-specific capital. It is only in the zero-inflation Calvo model that one can solve for a simple NK Phillips curve involving aggregate marginal cost and see how the coefficient on marginal cost depends on investment adjustment costs as well as price stickiness.

The crucial element of the procedure is that the approximation proceeds around a deterministic steady state where all firms are identical, so that the log-linearized first-order conditions are the same for all firms. This feature makes it possible to derive a first-order aggregation over firms that may temporarily deviate from the deterministic steady state, and may therefore be characterized by firm-specific state variables, $k_t(j)$ and $P_t(j)$.

Since firms differ only because they may or may not have the chance to adjust their prices, there are only two possibilities for firms to be the same in the steady state despite the fact that they do not all adjust their prices at the same time. First, there is zero steady-state inflation. In this case there is no need for firms to adjust their prices and they will all be the same anyway. Second, there is indexation: if firms cannot adjust their price optimally to their current state, their price is nevertheless adjusted according to the average inflation rate. Thus the firm's relative price also does not change. In the following we study the first case, zero steady-state inflation.

To summarize, we study the log-linear approximation of an economy with a deterministic steady state where all firms are identical. That is, we have $p_t^{ss}(j) = 1$ and $k_t^{ss}(j) = k^*$.

Optimal Capital Accumulation

Taking the firm's price decision as given for the time being, optimal choices of $k_{t+1}(j)$ and $x_t(j)$ maximize the expectation of (19) subject to the firm's product demand function (10), capital adjustment costs (13), and demand for labor (16).

The first-order conditions for k_{t+1} imply the following Euler equation:

$$G'\left(\frac{k_{t+1}(j)}{k_t(j)}\right) \tag{24}$$

$$= E_t\left[\beta\frac{\lambda_{t+1}}{\lambda_t}\left\{G\left(\frac{k_{t+2}(j)}{k_{t+1}(j)}\right)\cdot\left(\frac{G'\left[k_{t+2}(j)/k_{t+1}(j)\right]}{G\left[k_{t+2}(j)/k_{t+1}(j)\right]}\cdot\frac{k_{t+2}(j)}{k_{t+1}(j)}-1\right)\right.\right.$$
$$\left.\left.+u_{t+1}(j)\right\}\right],$$

where $u_{t+1}(j)$ denotes the value of having an additional unit of capital in period $t + 1$. This value, u, is the marginal labor cost reduction from the additional capital:

$$u_{t+1}(j) = -w_{t+1}(j)\frac{\partial H\left[y_{t+1}(j), k_{t+1}(j), A_{t+1}\right]}{\partial k_{t+1}(j)} = \frac{\alpha}{1-\alpha} w_{t+1}(j) h_{t+1}(j) / k_{t+1}(j). \tag{25}$$

The Euler equation is somewhat complicated, but it embodies the fact that a marginal increase in next period's capital stock has three effects. It subtracts from resources available for current consumption; it adds to resources available for future consumption; and it reduces future labor costs.

We now derive the log-linear approximation of the firm's Euler equation for capital (24). Let $\hat{x}$ denote the percentage deviation of a variable from its steady-state value x^*, $\hat{x} = dx/x^*$. Because $k_{t+1}^{ss}(j)/k_t^{ss}(j) = 1$, the log-linear approximation of the Euler equation is

$$\frac{G''(1)}{G'(1)}\left[\hat{k}_{t+1}(j) - \hat{k}_t(j)\right] = E_t\left[\beta\frac{G''(1)}{G'(1)}\left[\hat{k}_{t+2}(j) - \hat{k}_{t+1}(j)\right] + \left[1-\beta(1-\delta)\right]\hat{u}_{t+1}(j) + \hat{\lambda}_{t+1} - \hat{\lambda}_t\right]. \tag{26}$$

Note that $G''(1)/G'(1) = \epsilon_\psi$. The log-linear approximation of the marginal value of capital (24) is

$$\hat{u}_{t+1}(j) = \hat{w}_{t+1}(j) + \hat{h}_{t+1}(j) - \hat{k}_{t+1}(j). \tag{27}$$

After substituting for firm-specific labor supply using (5), this equation can be written as

$$\hat{u}_{t+1}(j) = \left[\nu\hat{h}_{t+1}(j) - \hat{\lambda}_{t+1}\right] + \hat{h}_{t+1}(j) - \hat{k}_{t+1}(j). \tag{28}$$

Next, substituting for the equilibrium employment from (16) and then substituting for firm j's output using the demand function (10), we get the marginal value of a unit of firm-specific capital in terms of the firm-specific variables (relative price and capital stock) and the aggregate variables (aggregate demand, marginal utility, and technology):

$$\hat{u}_{t+1}(j) = -\theta\frac{\nu+1}{1-\alpha}\hat{p}_{t+1}(j) - \left[\frac{(\nu+1)\alpha}{1-\alpha} + 1\right]\hat{k}_{t+1}(j) + \frac{\nu+1}{1-\alpha}\hat{y}_{t+1} - \hat{\lambda}_t - (\nu+1)\hat{A}_{t+1}. \tag{29}$$

Notice that the Euler-equation approximations (26) and (29) are the same for all firms, independent of their idiosyncratic state. We can now average/aggregate over these approximate first-order conditions of all firms. For the following, let

$$\hat{k}_t \equiv \int_0^1 \hat{k}_t(j)\, di \tag{30}$$

be the deviation of the aggregate capital stock from its steady-state value, and similarly for all other variables. Aggregating over the first-order conditions (26) and (29), we have

$$\begin{aligned}
\epsilon_{\psi}\left(\hat{k}_{t+1}-\hat{k}_{t}\right) &= E_{t}\left[\hat{\lambda}_{t+1}-\hat{\lambda}_{t}+\beta\epsilon_{\psi}\left(\hat{k}_{t+2}-\hat{k}_{t+1}\right)\right. \\
&\quad \left.+\left\{1-\beta\left(1-\delta\right)\right\}\hat{u}_{t+1}\right]; \qquad (31) \\
\hat{u}_{t+1} &= \frac{\nu+1}{1-\alpha}\hat{y}_{t+1}-\hat{\lambda}_{t+1}-\left[\frac{\left(\nu+1\right)\alpha}{1-\alpha}+1\right]\hat{k}_{t+1} \\
&\quad -\left(\nu+1\right)\hat{A}_{t+1}. \qquad (32)
\end{aligned}$$

For the aggregate marginal value of capital we have used the fact that (9) implies

$$\int_{0}^{1}\hat{p}_{t}\left(j\right)dj=0. \qquad (33)$$

Now define a firm's capital stock deviation from the aggregate deviation from the steady state as

$$\tilde{k}_{t}\left(j\right)=\hat{k}_{t}\left(j\right)-\hat{k}_{t} \qquad (34)$$

and subtract the aggregate conditions (31) and (32) from the firm-specific conditions (26) and (29) to yield

$$\begin{aligned}
\epsilon_{\psi}\left\{\tilde{k}_{t+1}\left(j\right)-\tilde{k}_{t}\left(j\right)\right\} &= E_{t}\left[\beta\epsilon_{\psi}\left\{\tilde{k}_{t+2}\left(j\right)-\tilde{k}_{t+1}\left(j\right)\right\}\right. \\
&\quad \left.+\left\{1-\beta\left(1-\delta\right)\right\}\tilde{u}_{t+1}\left(j\right)\right], \qquad (35) \\
\tilde{u}_{t+1}\left(j\right) &= -\frac{\nu+1}{1-\alpha}\theta\hat{p}_{t+1}\left(j\right) \\
&\quad -\left\{\frac{\left(\nu+1\right)\alpha}{1-\alpha}+1\right\}\tilde{k}_{t+1}\left(j\right). \qquad (36)
\end{aligned}$$

Note that (35) and (36) define an autonomous system for the firm-specific relative capital stock and relative price that is independent of aggregate variables. In order to complete this system, we need the expression for the unconditional expectation of the firm's relative price in the next period. There are two possibilities for next period's relative price. First, with probability $1-\phi$, the firm will be unable to adjust its nominal price, and its relative price declines with the aggregate inflation rate π. Second, with probability ϕ, the firm can adjust its nominal price and the optimal relative price choice is $\hat{p}^{*}$:

$$E_{t}\hat{p}_{t+1}\left(j\right)=\left(1-\phi\right)\left[\hat{p}_{t}\left(j\right)-E_{t}\pi_{t+1}\right]+\phi E_{t}\hat{p}_{t+1}^{*}\left(j\right). \qquad (37)$$

The analysis so far suggests that we can solve for the evolution of the firm's relative state variables independently of the evolution of aggregate state variables, but it also implies that optimal capital accumulation and optimal price setting will interact.

The Interaction of Price Setting and Capital Accumulation

We first show how aggregate inflation is related to the average price chosen by all the firms that can adjust prices. Once we conjecture that a particular price-adjusting firm's deviation from this average optimal price depends only on its relative capital stock, we can show how to solve for the evolution of the firm's relative capital stock. Conditional on the law of motion for the firm's optimal relative capital stock, one can then solve the firm's optimal price-setting problem. For an equilibrium, the conjecture on the optimal price-setting rule in the first step has to be consistent with the solution of the price-setting problem in the second step. This second step involves quite a bit of algebra, and we refer the reader to Woodford (2005) for the solution. We do state the Phillips curve equation that follows from these steps. The form of the Phillips curve illustrates the appeal of firm-specific capital.

Aggregate Inflation

In the Calvo setup, aggregate inflation is determined as a weighted average of the current distribution of relative prices and the optimal relative prices set by price-adjusting firms. At the beginning of period $t+1$, a fraction $1-\phi$ of all firms keeps their price and a fraction ϕ adjusts their price conditional on their state. For both groups we can use the unconditional distribution of all firms in the economy. Thus, the deviation of the aggregate price level from the steady state is

$$\hat{P}_{t+1} = (1-\phi)\int_0^1 \hat{P}_t(j)\,dj + \phi\int_0^1 \hat{P}_t^*(j)\,dj = (1-\phi)\,\hat{P}_t + \phi\hat{P}_{t+1}^*. \quad (38)$$

Subtract $\hat{P}_t$ from both sides and the aggregate inflation rate is

$$\pi_{t+1} = \hat{P}_{t+1} - \hat{P}_t = \phi\left(\hat{P}_{t+1}^* - \hat{P}_t\right). \quad (39)$$

Adding and subtracting $\hat{P}_{t+1}$ on the right-hand side and using the definition of the inflation rate, we get the inflation rate proportional to the average optimal relative price

$$(1-\phi)\,\pi_{t+1} = \phi\left(\hat{P}_{t+1}^* - \hat{P}_{t+1}\right) = \phi\hat{p}_{t+1}^*. \quad (40)$$

Using expression (40) for the inflation rate in the definition of next period's unconditional expected relative price (37) we get

$$\begin{aligned} E_t\hat{p}_{t+1}(j) &= (1-\phi)\left(\hat{p}_t(j) - E_t\left[\frac{\phi}{1-\phi}\hat{p}_{t+1}^*\right]\right) + \phi E_t^*\hat{p}_{t+1}(j) \\ &= (1-\phi)\,\hat{p}_t(j) + \phi E_t\left[\hat{p}_{t+1}^*(j) - \hat{p}_{t+1}^*\right]. \quad (41) \end{aligned}$$

Now assume that the deviation of a firm's optimal relative price from the average optimal relative price is a function of the firm's relative state only:

$$\hat{p}_t^*(j) = \hat{p}_t^* - \mu\tilde{k}_t(j). \tag{42}$$

Then equations (35), (36), (41), and (42) define an autonomous system for the firm-specific relative capital stock, $\tilde{k}(j)$, and relative price, $\hat{p}(j)$, that is independent of aggregate variables. We are interested in a recursive solution to this system, that is, a solution such that the firm's choice for next period's relative capital stock, $\tilde{k}_{t+1}(j)$, is a function of its own relative state only, $\left[\tilde{k}_t(j), \hat{p}_t(j)\right]$:

$$\tilde{k}_{t+1}(j) = \Lambda\tilde{k}_t(j) - \tau\hat{p}_t(j). \tag{43}$$

Optimal Price Setting

Woodford (2005) solves the optimal price-setting problem conditional on the optimal capital accumulation rule (43). In particular, the optimal price-setting rule is shown to be of the form assumed in equation (42): the deviation of a particular firm's optimal relative price from the average optimal relative price, $\hat{p}_t^*(i) - \hat{p}_t^*$, is a function of the firm's relative state, $\tilde{k}_t(i)$. Woodford (2005) shows how one can obtain the coefficients Λ, τ, and μ through the method of undetermined coefficients.

The solution of the optimal pricing problem yields an expression for the average optimal price as a function of the average marginal labor cost of production, $\hat{s}_t$, and expected future optimal prices and inflation:

$$\hat{p}_t^* = \frac{1-(1-\phi)\beta}{\Gamma}\hat{s}_t + (1-\phi)\beta E_t\left[\pi_{t+1} + \hat{p}_{t+1}^*\right], \tag{44}$$

where Γ is a coefficient to be determined by the solution procedure. In particular, Γ will depend on the the price-adjustment probability ϕ and the degree of capital adjustment costs, ϵ_ψ. Average marginal cost is by definition

$$\begin{aligned}\hat{s}_t &\equiv \int_0^1 \left[\hat{w}_t(j) + \hat{h}_t(j) - \hat{y}_t(j)\right] dj \\ &= \left(\frac{\nu+1}{1-\alpha} - 1\right)\hat{y}_t - \hat{\lambda}_t - (\nu+1)\left[\frac{\alpha}{1-\alpha}\hat{k}_t + \hat{A}_t\right].\end{aligned} \tag{45}$$

We can now use again the expression for aggregate inflation in the Calvo model in (40) and derive the "standard" New Keynesian Phillips curve

$$\pi_t = \frac{[1-(1-\phi)\beta]\phi}{(1-\phi)\Gamma}\hat{s}_t + \beta E_t\left[\pi_{t+1}\right]. \tag{46}$$

For a simple Calvo model with no firm-specific capital, $\Gamma = 1$. Thus the modified Calvo model with firm-specific capital adjustment costs generates almost the same NK Phillips curve as the basic Calvo model, except for Γ.

In particular, higher capital adjustment costs increase Γ and thereby reduce the coefficient on the marginal cost term. Woodford (2005) and Eichenbaum and Fisher (2004) thus argue that a low estimated coefficient on marginal cost does not necessarily imply that the price-adjustment probability is very low; it can also mean that the capital adjustment costs are very high.

3. THE TAYLOR MODEL

In the Taylor model, price adjustment occurs every J periods for an individual firm, and in any given period by a fraction $1/J$ of firms. Because there is no uncertainty regarding when a firm will adjust its price, the state space does not explode as it does in the Calvo model. Therefore, the Taylor model with firm-specific capital can be approximated easily around a steady state with nonzero inflation. Here we present the exact equations of the model. We then linearize them and compute the model's local dynamics.

Pricing

An individual firm that can adjust its price in period t chooses a sequence of nominal prices, $\left\{P^*_{t+Js}(j)\right\}$, every J periods, and a sequence of capital stocks $\left\{k^*_{t+1}(j)\right\}$ every period, that maximizes the objective function

$$\max E_t \sum_{s=0}^{\infty} \beta^{Js} \sum_{\tau=0}^{J-1} \beta^{\tau} \frac{\lambda_{t+Js+\tau}}{\lambda_t} \times \tag{47}$$

$$\left\{ \left[\frac{P^*_{t+Js}(j)}{P_{t+Js+\tau}} \right]^{1-\theta} y_{t+Js+\tau} - w_{t+Js+\tau}(j)\, h_{t+Js+\tau}(j) - x_{t+Js+\tau}(j) \right\},$$

subject to the demand for the firm's goods (10) and the firm's demand for labor (16). Note that in contrast to the Calvo model, the expectation operator in (47) is the unconditional expectation operator—there is no uncertainty in the price adjustment process. The first-order conditions for optimal price setting are

$$E_t \sum_{\tau=0}^{J-1} \beta^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} (1-\theta) \frac{1}{P_{t+\tau}} \left(\frac{P^*_t(j)}{P_{t+\tau}} \right)^{-\theta} y_{t+\tau}$$

$$+\theta E_t \sum_{\tau=0}^{J-1} \beta^{\tau} s_{t+\tau}(j) \frac{\lambda_{t+\tau}}{\lambda_t} \frac{1}{P_{t+\tau}} \left(\frac{P^*_t(j)}{P_{t+\tau}} \right)^{-\theta-1} y_{t+\tau} = 0, \tag{48}$$

where $s_t(j)$ is the firm's marginal (labor) cost of production, (17). The first-order conditions for optimal capital accumulation are the same as in the Calvo model, equations (25) and (26).

To simplify (48) we will solve for the optimal price $P_t^*(j)$, at the same time dividing both sides of the equation by P_t :

$$\frac{P_t^*(j)}{P_t} = \left(\frac{\theta}{\theta - 1}\right) \frac{E_t \sum_{\tau=0}^{J-1} \beta^\tau s_{t+\tau}(j) \frac{\lambda_{t+\tau}}{\lambda_t} \left(\frac{P_t}{P_{t+\tau}} \frac{1}{P_t}\right)^{-\theta} y_{t+\tau}}{E_t \sum_{\tau=0}^{J-1} \beta^\tau \frac{\lambda_{t+\tau}}{\lambda_t} \frac{P_t}{P_{t+\tau}} \left(\frac{P_t}{P_{t+\tau}} \frac{1}{P_t}\right)^{-\theta} y_{t+\tau}}. \tag{49}$$

Next, note that P_t^θ cancels from the numerator and denominator:

$$\frac{P_t^*(j)}{P_t} = \left(\frac{\theta}{\theta - 1}\right) \frac{E_t \sum_{\tau=0}^{J-1} \beta^\tau s_{t+\tau}(j) \frac{\lambda_{t+\tau}}{\lambda_t} \left(\frac{P_{t+\tau}}{P_t}\right)^{\theta} y_{t+\tau}}{E_t \sum_{\tau=0}^{J-1} \beta^\tau \frac{\lambda_{t+\tau}}{\lambda_t} \left(\frac{P_{t+\tau}}{P_t}\right)^{\theta-1} y_{t+\tau}}. \tag{50}$$

Until now we have carried around the firm's index j, which lies in the interval $[0, 1]$. However with Taylor pricing, it is only necessary to keep track of J different types of firms—any firms that set their price in the same period behave identically. Of course, this is not the case in the Calvo model.[7] Henceforth the index j denotes the finite types J. For example, the marginal cost for a firm that set its price in period $t - j$ will be $s_{j,t}$; the price in period t charged by a firm that last set its price in period $t - j$ will be $P_{j,t}$. Thus, instead of $P_t^*(j)$ we will write $P_{0,t}$.

$$\frac{P_{0,t}}{P_t} = \frac{\theta}{\theta - 1} \cdot \frac{E_t \sum_{j=0}^{J-1} \beta^j s_{j,t+j} \lambda_{t+j} \left(\frac{P_{t+j}}{P_t}\right)^{\theta} y_{t+j}}{E_t \sum_{j=0}^{J-1} \beta^j \lambda_{t+j} \left(\frac{P_{t+j}}{P_t}\right)^{\theta-1} y_{t+j}}. \tag{51}$$

Imposing the fact that there are only J prices charged, the price index can be written as

$$P_t = \left\{ \frac{1}{J} \sum_{j=0}^{J-1} P_{0,t-j}^{1-\theta} \right\}^{\frac{1}{1-\theta}}, \tag{52}$$

and the demand equations are

$$y_{j,t} = p_{j,t}^{-\theta} y_t, \qquad j = 0, 1, ..., J - 1. \tag{53}$$

Also, from the household side we have the labor supply equations

$$\frac{\gamma h_{j,t}^{\upsilon}}{\lambda_t} = w_{j,t}, \qquad j = 0, 1, ..., J - 1. \tag{54}$$

[7] We could also study the Taylor model under the assumption that firms that set their price in the same period have initial conditions that involve heterogeneous capital. Under this assumption, there would be multiple prices chosen in the same period. However, as long as the size of the initial state was manageable, it would be feasible to analyze such a situation.

Investment and Labor Demand

Here, for convenience, we collect the equations that were stated in Section 1 for the general model and the equations for optimal capital accumulation from the Calvo model. We express these equations in a form specific to the Taylor model. The technology is

$$y_{j,t} = k_{j,t}^{\alpha} \left(A_t h_{j,t}\right)^{1-\alpha}. \tag{55}$$

Adjustment costs for the capital stock are

$$x_{j,t} = k_{j,t} G \left(k_{j+1,t+1}/k_{j,t}\right). \tag{56}$$

The first-order condition for next period's capital stock depends on the stage of the price cycle that a firm is in. To simplify notation, let "$j+i$" denote $(j+i) \bmod (J-1)$ for $j = 0, 1, ..., J-1$. For example for $j = J-2$ and $i = 3$, $j+i = 2$. The rewritten first-order condition (24) for next period's capital stock is then

$$\begin{aligned} G'\left(\frac{k_{j+1,t+1}}{k_{j,t}}\right) &= \beta E_t \left[\frac{\lambda_{t+1}}{\lambda_t}\left\{\frac{k_{j+2,t+2}}{k_{j+1,t+1}} G'\left(\frac{k_{j+2,t+2}}{k_{j+1,t+1}}\right)\right.\right. \\ &\qquad \left.\left. -G\left(\frac{k_{j+2,t+2}}{k_{j+1,t+1}}\right) + u_{j+1,t+1}\right\}\right]. \end{aligned} \tag{57}$$

Real profits in period t for firm j are given by

$$d_{j,t} = p_{j,t} y_{j,t} - w_{j,t} h_{j,t} - x_{j,t}. \tag{58}$$

The marginal cost of production is

$$s_{j,t} = \frac{w_{j,t} h_{j,t}}{(1-\alpha)\, y_{j,t}}. \tag{59}$$

4. RESULTS FOR THE TAYLOR MODEL

In this section we present results describing how the behavior of the Taylor model with firm-specific capital varies with the steady-state inflation rate around which it is linearized. We follow Sveen and Weinke's (2005) analysis of the Calvo model with firm-specific adjustment costs and zero steady-state inflation. First, we report on the range of parameters for the monetary policy rule and adjustment costs for which we can find unique RE equilibria. This range is sensitive to the steady-state inflation rate: higher inflation rates reduce the set of parameters for which there is a unique RE equilibrium. Next, we compare impulse response functions to a productivity shock for zero and moderate inflation. They differ, but not dramatically.

The model is parameterized as follows. We interpret a period as a quarter, and set the discount factor, $\beta = 0.99$; the risk aversion parameter, $\sigma = 2$; the inverse labor supply elasticity, $\nu = 1$; the capital depreciation rate, $\delta = 0.03$; and the capital income share, $\alpha = 0.36$. This is a standard parameterization.

We set the investment adjustment cost parameter, $\epsilon_{\psi} = 3$, as in Woodford (2005). Based on evidence from aggregate data, Eichenbaum and Fisher (2005) suggest that this value represents a lower bound for adjustment costs. Around a zero-inflation steady state, there is no need to specify the function $G(.)$ beyond the two parameters, δ and ϵ_{ψ}. Around steady states with nonzero inflation however, it is necessary to specify the entire function. We use

$$G(x) = \left(\delta - \frac{1}{1+\epsilon_{\psi}}\right) + \frac{x^{1+\epsilon_{\psi}}}{1+\epsilon_{\psi}}, \qquad (60)$$

which satisfies the desired properties $G(1) = \delta$, $G'(1) = 1$ and $G''(1) = \epsilon_{\psi}$.

Equilibrium Determinacy

A good monetary policy rule should imply a unique RE equilibrium. If the RE equilibrium is not unique, then at any point in time several different equilibrium time paths for current and future outcomes are possible. In other words, the equilibrium is indeterminate. In this situation the path that is expected to be chosen will occur, but many can be chosen. The choice of equilibrium path then may depend on random shocks that are not fundamental to the economy, that is, they do not constrain the set of resource-feasible allocations in the economy. In these "sunspot" equilibria self-fulfilling expectations that coordinate on the nonfundamental shocks introduce unnecessary fluctuations into the economy.[8] Since the representative agent is risk-averse, she will prefer a smooth consumption path relative to the same smooth consumption path with some added mean zero random fluctuations. This means that, in general, "sunspot" equilibria are sub-optimal, and a good monetary policy should not give rise to equilibrium indeterminacy.

Taylor (1993) proposed a monetary policy rule of the form $f_{\pi} = 1.5$ and $f_y = 0.125$ based on the outcomes of model simulations.[9] This policy rule reflects the Taylor principle that monetary policy should increase nominal interest rates more than one-for-one for any increase of inflation. In basic sticky-price models with reasonable specifications of price rigidity and without capital, this principle will, in general, imply a unique RE equilibrium. Sveen and Weinke (2005) evaluate the role of the policy parameter, f_{π}, and the degree of price stickiness, ϕ, for the existence of unique RE equilibria in the Calvo model with firm-specific capital. They show that as the degree of price stickiness increases, the set of policy parameters for which there is local uniqueness becomes smaller. For the Taylor model we provide an analog to

---

[8] For a textbook treatment of sunspot equilibria, see, for example, Farmer (1993).

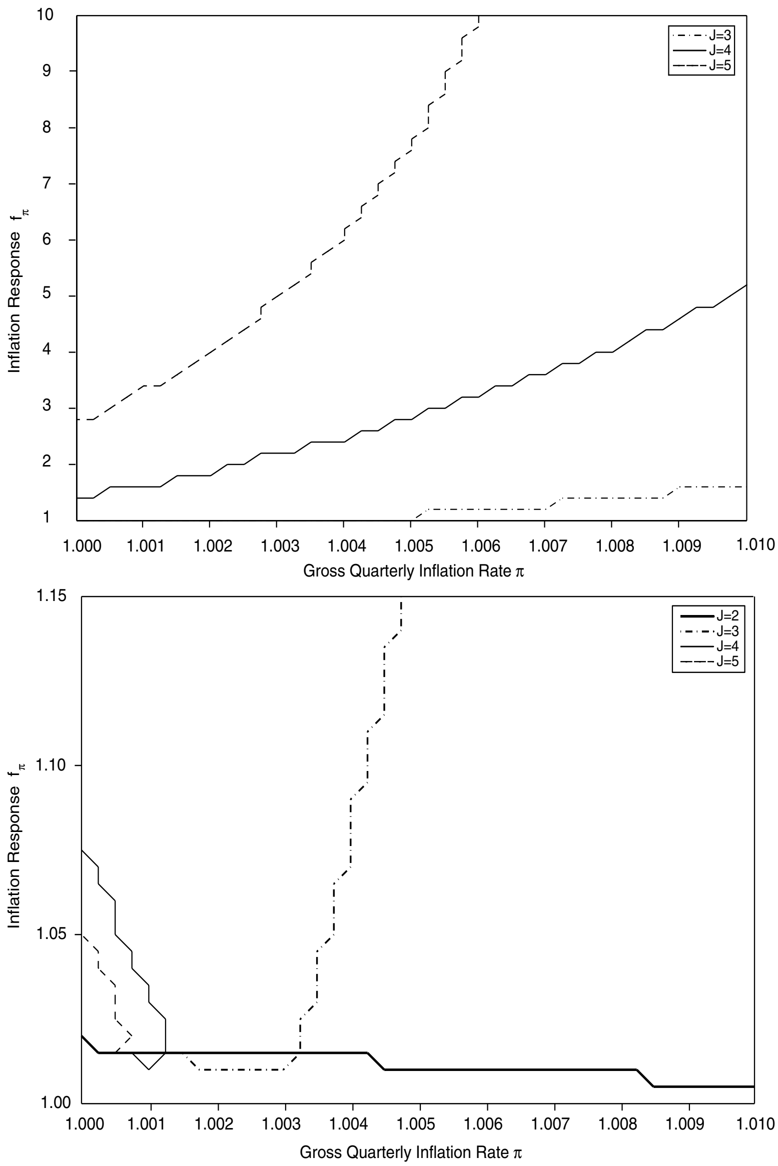
[9] Taylor (1993) writes the policy rule for annual data, thus his $f_y = 0.5$ coefficient on output deviations translates to $0.125 = 0.5/4$ in our quarterly model. Taylor's proposed policy rule has also spawned an empirical literature that tries to estimate whether actual monetary policy conforms to some version of this policy rule, for example, Clarida, Galí, and Gertler (2000).

their results (price stickiness is now represented by J). We also study the impact of the steady-state inflation rate, π, and investment adjustment costs, ϵ_{ψ}, on equilibrium indeterminacy. We find that local uniqueness becomes less likely for higher inflation rates. Depending on the degree of price stickiness, high or low values of the adjustment cost parameter ϵ_{ψ} can lead to indeterminacy.

In Figure 1, we plot several graphs in (π, f_{π})-space that represent the border between indeterminacy and uniqueness for a policy rule that does not respond to output, $f_y = 0$. We present this information in two panels because for very low values of f_{π}, it is not possible to convey the relevant information unless the f_{π}-axis scale is very fine. The inflation rate, π, is the rate of price change from one period to the next, and since a period represents a quarter, a gross inflation rate of 1.01 represents a 4 percent annual inflation rate. Each graph corresponds to a different value of J. In the top panel of Figure 1, which corresponds to relatively high values of f_{π}, the region of equilibrium indeterminacy (uniqueness) for an economy with price stickiness, J, is between the graph and the southeast (northwest) corner of the figure. There is no graph for $J = 2$ in the top panel because uniqueness holds everywhere in the figure when $J = 2$. The bottom panel, corresponding to low values of f_{π}, is less straightforward: for $J = 2$ there is indeterminacy below the graph; for $J = 3, 4$ and 5 there is indeterminacy generally below and to the right of the graphs.

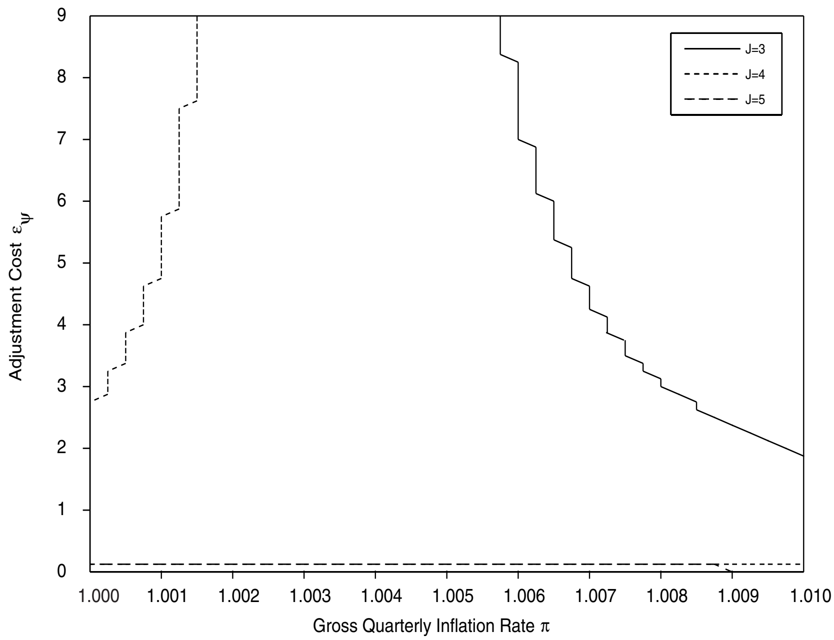
We find that for moderate steady-state inflation, if prices are fixed for more than two periods then policy needs to respond to inflation significantly more than one-to-one in order for the RE equilibrium to be unique. First, for all values of J and π that we consider, equilibrium is indeterminate if f_{π} is less than approximately 1.01 (the precise number varies with J and π), as seen in the lower panel of Figure 1. In contrast, for the Calvo model with zero inflation, Sveen and Weinke (2005) find that there is a neighborhood of $f_{\pi} = 1$ such that equilibrium is unique. Second, for fixed degrees of price stickiness, $J > 2$, the policy response f_{π} required to maintain a unique equilibrium can become quite large as we increase the steady-state inflation rate, as seen in the upper panel of Figure 1. This occurs even though the steady-state inflation rates that we consider are moderate, less than 4 percent per year. For example, if prices are fixed for three periods, around a zero-inflation steady state there is a unique equilibrium if $f_{\pi} \gtrsim 1.02$; in contrast, around a 4 percent inflation steady state there is a unique equilibrium only if $f_{\pi} \gtrsim 1.73$. The sensitivity to steady-state inflation becomes more extreme for higher degrees of price stickiness. If prices are fixed for four periods, around a zero-inflation steady state there is a unique equilibrium if $f_{\pi} \in \{(1.02, 1.074) \cup (1.47, \infty)\}$; in contrast, around a 4 percent inflation steady state there is a unique equilibrium only if $f_{\pi} \gtrsim 5.29$. Finally, for a given steady-state inflation rate, the region of

Figure 1 The Monetary Policy Response to Inflation and Equilibrium Indeterminacy



indeterminacy is increasing in the degree of price rigidity. This is consistent with Sveen and Weinke (2005, Figure 1).

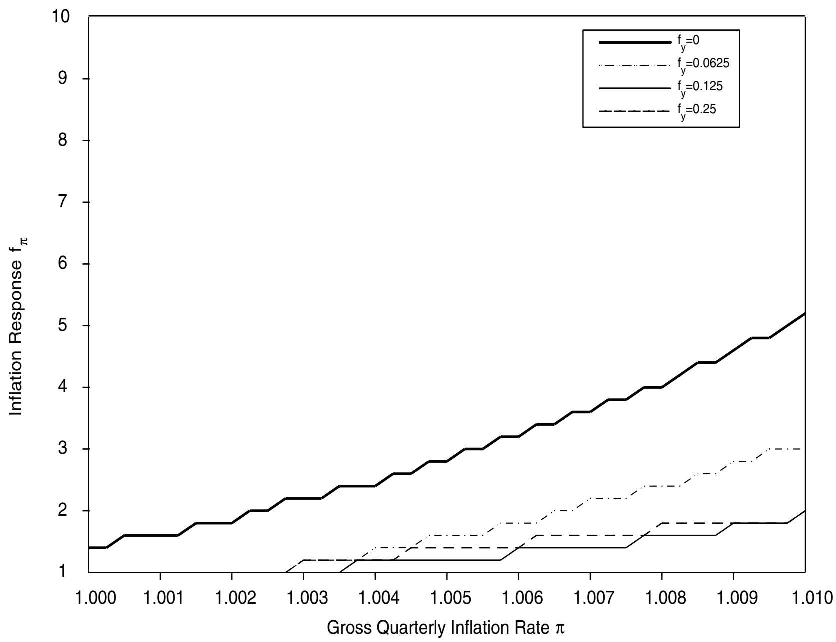
Figure 2 Investment Adjustment Costs and Equilibrium Indeterminacy



For steady-state inflation rates that are even moderately high, the RE equilibrium tends to be indeterminate for a wide range of values of the adjustment cost parameter, ϵ_ψ, but the precise relationship is sensitive to the degree of price stickiness. In Figure 2 we graph the borders between indeterminacy and uniqueness in (π, ϵ_ψ)-space for different values of price stickiness J and a policy rule with $f_\pi = 1.5$ and $f_y = 0$. For parameter combinations between a graph and the left (right) border of the figure, the RE equilibrium is locally unique (indeterminate) for $J = 3$ and $J = 4$ (there is also a region of uniqueness near $\epsilon_\psi = 0$ for $J = 4$). For $J = 5$ there is indeterminacy (uniqueness) above (below) the graph. For $J = 2$ there is uniqueness across the entire figure. For $J = 3$ and $J = 4$ the region of indeterminacy is increasing in the steady-state inflation rate. However, as the inflation rate increases, for $J = 3$ indeterminacy first appears at high values of ϵ_ψ, whereas for $J = 4$ indeterminacy first appears at low values of ϵ_ψ.

Sveen and Weinke (2005) argue that if a monetary policy rule responds not only to the inflation rate but also to output, then it is more likely that the RE equilibrium is unique. Indeed the Taylor rule (1993) specifies the coefficient on output as 0.125. In Figure 3 we graph the borders between indeterminacy and uniqueness in (π, f_π)-space for different values of the coefficient on output

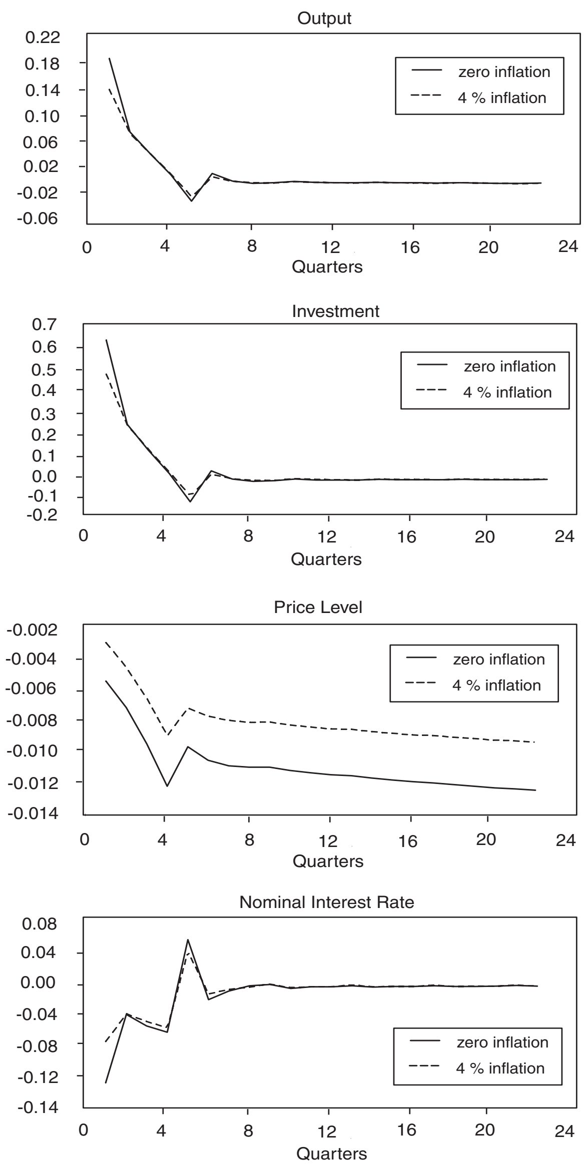
Figure 3 The Monetary Response to Inflation and Output, and Equilibrium Indeterminacy



in the policy rule f_y and fixed price stickiness $J = 4$. For parameter combinations between a graph and the left (right) border of the figure, the RE equilibrium is locally unique (indeterminate). Again, as the steady-state inflation rate increases, it becomes more likely that the RE equilibrium is not locally unique. For fixed steady-state inflation, the RE equilibrium is unique if the policy response to output is sufficiently large. This confirms the findings of Sveen and Weinke (2005). Note, however, that even for moderate steady-state inflation, it takes a large coefficient on output to generate determinacy in a rule that includes the standard Taylor coefficient, $f_y = 0.125$, on output. For example, for annual inflation of 4 percent (corresponding to $\pi = 1.01$ in Figure 3), the coefficient on inflation needs to be greater than 2 in order to maintain a unique RE equilibrium. This is substantially more than the 1.5 value suggested by Taylor.

The overall message of these figures is that when the average inflation rate is even moderately high—say, above 3.5 percent annually—the coefficient on inflation must be large relative to conventional values such as Taylor's 1.5 in order to generate a unique RE equilibrium.

Figure 4 Impulse Response to Productivity Shock



Model Dynamics

Figure 4 plots the response of several of the model's aggregate variables to a white noise productivity shock. We set $J = 4$ and $f_\pi = 5.5$. The solid lines correspond to a steady state of zero inflation, and the dashed lines correspond to a steady state of 4 percent annual inflation. The responses to a productivity shock differ somewhat across very low and moderate inflation, but the differences are not dramatic, and they essentially disappear after the impact period. Given our findings about indeterminacy in Figures 1 and 2, it may seem surprising that the impulse responses do not differ more across steady-state inflation rates. There is, however, a good explanation for this. Unlike a crossing from uniqueness to nonexistence, a crossing from uniqueness to multiplicity need not be "foreshadowed" by large changes in the model's dynamics. As we change a model's parameters and uniqueness disappears, the solution we were tracking does not vanish—it is simply complemented with other solutions.

5. CONCLUSIONS

Sveen and Weinke (2004) and Woodford (2005) have made important contributions in showing how one can linearly approximate the Calvo sticky-price model when capital is tied to the individual firm. Their work shows that capital adjustment costs at the firm level are complementary to price stickiness in generating a small coefficient on marginal cost in the New Keynesian Phillips curve. Around a steady state with nonzero inflation, it is not (yet) known how to approximate the Calvo model with firm-level investment; in such a steady state there would be heterogeneity in both prices and capital stocks. Much recent empirical work on the NK Phillips Curve has used data which is inconsistent with the zero-inflation approximation, so we would like to have some means of evaluating the generality of results from the zero-inflation case. In the Taylor sticky-price model it is straightforward to incorporate firm-specific capital even with nonzero steady-state inflation. Comparing zero- and moderate (4 percent) rates of steady-state inflation, one finds that *if there is a locally unique equilibrium*, quantitatively the model's dynamics are not very sensitive to the rate of inflation. This is consistent with the work of Ascari (2004), who finds that the dynamics of the basic Taylor model (i.e., without firm-specific capital) are relatively insensitive to average inflation, in comparison to the Calvo model. However, we find that the range of parameter values for which the model has a locally unique equilibrium is extremely sensitive to even small changes in steady-state inflation—for example going from zero to 4 percent annual inflation causes a dramatic increase in the size of the parameter space for which there is local indeterminacy. The ability to deal with nonzero inflation in the Taylor model points toward the value of conducting empirical work on the New Keynesian Phillips curve in the Taylor model framework. See Guerrieri (forthcoming) for an important step in this direction.[10] However, the sensitivity of the local equilibrium uniqueness to the average inflation rate presents obstacles to further empirical progress.

[10] Cogley and Sbordone (2005) is an important example of empirical work on the Phillips curve that allows for the possibility of nonzero steady-state inflation. They use a Calvo model with firm-specific capital but without firm-specific investment.

REFERENCES

Ascari, Guido. 2004. “Staggered Prices and Trend Inflation: Some Nuisances.” *Review of Economic Dynamics* 7: 642–667.

Blanchard, Olivier J. and Nobuhiro Kiyotaki. 1987. “Monopolistic Competition and the Effects of Aggregate Demand.” *American Economic Review* 77 (4): 647–66.

Calvo, Guillermo. 1983. “Staggered Prices in a Utility Maximizing Framework.” *Journal of Monetary Economics* 12: 383–98.

Clarida, Richard, Jordi Galí, and Mark Gertler. 2000. “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory.” *Quarterly Journal of Economics* 115 (1): 147–80.

Cogley, Timothy and Argia M. Sbordone. 2005. “A Search for a Structural Phillips Curve.” Federal Reserve Bank of New York Staff Report 204.

Coenen, Guenter and Andrew Levin. 2004 “Identifying the Influences of Real and Nominal Rigidities in Aggregate Price-Setting Behavior.” ECB Working Paper 418.

de Walque, Gregory, Frank Smets and Raf Wouters. 2004. “Price Setting in General Equilibrium: Alternative Specifications.” Manuscript.

Dixit, Avinash K. and Joseph E. Stiglitz. 1977. “Monopolistic Competition and Optimum Product Diversity.” *American Economic Review* 67 (3): 297–308.

Eichenbaum, Martin and Jonas Fisher. 2004. “Evaluating the Calvo Model of Sticky Prices.” NBER Working Paper 10617.

Farmer, Roger E.A. 1993. *The Macroeconomics of Self-Fulfilling Expectations*. MIT Press: Cambridge MA.

Galí, Jordi and Mark Gertler. 1999. “Inflation Dynamics: A structural econometric analysis.” *Journal of Monetary Economics* 44 (October): 195–222.

Green, Edward J. 2005. “A Review of Interest and Prices: Foundations of a Theory of Monetary Policy.” *Journal of Economic Literature* 43 (March): 121–34.

Guerrieri, Luca. Forthcoming. “The Inflation Persistence of Staggered Contracts.” *Journal of Money Credit and Banking*.

Sbordone, Argia M. 2002. “Prices and Unit Labor Costs: a new test of price stickiness.” *Journal of Monetary Economics* 49 (March): 265–92.

Sveen, Tommy and Lutz Weinke. 2004. "Pitfalls in the Modelling of Forward-Looking Price Setting and Investment Decisions." Central Bank of Norway Working Paper 2004/1.

______________. 2005. "New Perspectives on Capital, Sticky Prices, and the Taylor Principle." *Journal of Economic Theory* 123 (July): 21–39.

Taylor, John B. 1980. "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy* 88: 1–22.

______________. 1993. "Discretion versus Policy Rules in Practice," Carnegie-Rochester Conference Series on Public Policy 39: 195–214.

Wolman, Alexander L. Forthcoming. "The Frequency and Costs of Individual Price Adjustment." *Managerial and Decision Economics*.

______________. 2004. "The Optimal Rate of Inflation with Trending Relative Prices." Federal Reserve Bank of Richmond manuscript.

______________, and Elise A. Couper. 2003. "Potential Consequences of Linear Approximation in Economics." Federal Reserve Bank of Richmond *Economic Quarterly* 89 (1): 51–68.

Woodford, Michael. 2003. *Interest and Prices*. Princeton University Press. Princeton, NJ.

______________. 2005. "Firm-Specific Capital and the New-Keynesian Phillips Curve." *International Journal of Central Banking.* Forthcoming.