Tests of Parameter Stability

with Application to the Money-Income Relation in the United States

James H. Stock*

Kennedy School of Government
Harvard University
Cambridge, MA 02138

November 1993

*The author thanks Graham Elliott, Jesus Gonzalo and Mark Watson, for helpful discussions and/or comments. The analytical material in section 3 of this paper is in part translated from Stock (1993). The empirical results in section 4 are new.

ABSTRACT

Although there is a large literature which examines monetary aggregates as predictors of real output, there is much less formal evidence on the stabiity of this link. This paper reviews three classes of tests for parameter stability: likelihood ratio tests for a single break, tests based on cumulative forecast errors, and tests against the alternative of time-varying parameters. These tests are then applied to data on M1, M2, real GDP and interest rates for the United States from 1959 to 1992. Predictive relationships based on M1 and interest rate spreads have been weak and/or unstable. In contrast, the M2-real output link is found to have been stable over this period, with M2 have significant predictive power for real output.

1. Introduction

Economic theory suggests that many empirical time series relations might be unstable over time. In the realm of monetary economics, for example, the introduction of different financial instruments, changes in operating procedures of the central bank, and increasing openness of the economy to international capital flows all might result in changes in the relation between money and output. Understanding these changes is important not only for economic forecasting but also for formulating economic policies, for example the formulation of monetary policies to stabilize short-run economic fluctuations. The first step in understanding these changes in empirical economic relations is recognizing that they have in fact occurred.

This article examines tests for parameter stability in time series regressions and uses these tests to assess the stability of the link between money and output in postwar data for the United States. Although tests for parameter stability have been part of the theoretical econometrics literature for the past three decades, only recently has a sufficiently general distribution theory been developed so that these and newer tests can be applied in standard time series regressions settings with lagged dependent variables and additional stochastic and deterministic regressors. The first part of this article therefore reviews several tests for parameter stability and outlines the derivation of their asymptotic null distribution. With these tests in hand, we then turn to the substantive problem of assessing the stability of the money-output relation. Our main empirical conclusions are that, for the United States since 1960, the M1-output relation is weak and unstable but the M2-output relation has been stable and moreover M2 has had significant marginal predictive content for output, given interest rates. The predictive content of M2 is strengthened and remains stable when one controls for lagged M2 velocity, more precisely, when one estimates a single-equation error correction model with the error correction term being the cointegrating residual from a long-run money demand relation.

The plan of the paper is as follows. Section 2 reviews the problem of detecting a single change in the mean of a serially uncorrelated process. This typically is not of direct interest in time series applications but the arguments leading to the asymptotic distribution are simplest in this case. Tests for stability in multivariate time series regression and their limiting null distributions are reviewed in section 3. The empirical results on the money-output relation are summarized in section 4. Section 5 concludes.

2. Tests for Changes in the Mean of a Serially Uncorrelated Process

Let the observed series be y_t and let μ_t denote its mean at time t. In this section, we suppose y_t is serially uncorrelated with a mean that possibly changes:

$$(2.1) \qquad y_t = \mu_t + \epsilon_t, \quad t = 1, 2, \ldots, T,$$

where ϵ_t is a martingale difference sequence (that is, $E(\epsilon_t | \epsilon_{t-1}, \epsilon_{t-1}, \ldots) = 0$) with four moments and constant conditional variance $E(\epsilon_t^2 | \epsilon_{t-1}, \epsilon_{t-1}, \ldots) = \sigma^2$. A leading special case of the martingale difference sequence model is that ϵ_t is i.i.d. with mean zero and variance σ^2.

A simple model for a structural break is that there has been a single shift in the mean of the process at date r:

$$(2.2) \qquad \mu_t = \mu_0, \ t \le r, \ \text{ and } \mu_t = \mu_0 + \delta, \ t > r.$$

Suppose that the break date r is known. Then the statistical problem of testing $\delta=0$ is the well-known problem of testing for the equality of the means in two independent samples of i.i.d.

variates, when it is assumed that the error variance is the same in the two distributions. There are several ways to implement this test, a leading example being to run the regression of y_t on a constant and a dummy variable which takes on the value 1 for t>r and 0 otherwise. The t-statistic on this dummy variable is the Wald test of equality of the two means and has an asymptotic standard normal distribution. The Wald test statistic is,

$$(2.3) \qquad F_T(r/T) = T\{SSR_{1,T} - (SSR_{1,r} + SSR_{r+1,T})\}/(SSR_{1,r} + SSR_{r+1,T})$$

where $SSR_{t_1,t_2} = \sum_{t=t_1}^{t_2}(y_t-\bar{y}_{t_1,t_2})^2$, where $\bar{y}_{t_1,t_2} = (t_2-t_1+1)^{-1}\sum_{t=t_1}^{t_2}y_t$ is the sample average of y_t over $[t_1, t_2]$. The statistic is asymptotically equivalent to the Gaussian likelihood ratio (LR) statistic. If r is known and $r/T \to s'$, $0<s'<1$, then $F_T(r/t)$ has an asymptotic χ_1^2 null distribution (see Engle (1983)). Thus the hypothesis can be tested by computing $F_T(r/T)$ and checking whether it exceeds the χ_1^2 critical value at the desired significance level.

This textbook problem of testing the means of two sample for equality becomes much harder when r is unknown. In this situation, a researcher might be tempted first to plot the data and to guess at a likely spot for a break, then to split the sample at this point and perform the test. However, by performing the preliminary analysis, the null distribution of the subsequent break test is suspect. This is an essential point: suppose the researcher was in fact given i.i.d. data with the same mean, but he or she sought split points in which there appeared to be a change in the mean. By choosing those splits which seemed most likely to be significant, the researcher would tend to reject too often. Worse, the true null distribution of the test could not be computed, because the break date selection procedure was subjective and specific to that researcher.

An empirical econometrician might object to this metaphor: in performing a break test, he or she does not consciously examine the data for the most likely break prior to computing the test. Rather, the break date is based on some economic event, for example the 1974-5 oil price shock. But, upon reflection, arguably this amounts to the same thing: the 1974-5 oil shock is of interest because it has been associated with substantial changes in the performance of industrialized economies, in particular their subsequently slower average growth rates of real per capita GDP and productivity. Thus the pretesting – searching for the break which is most likely to be significant – has already been done, if not by the researcher by other economists whose observations led to selecting a particular break date.

This reasoning suggests developing tests for a break in which the break date is formally treated as unknown *a-priori*. Quandt (1960) suggested an intuitively appealing approach to this problem: compute the $F_T(r/T)$ statistic over a range of break dates, $r_0 \le r \le r_1$, then take the maximum. This leads to the so-called Quandt (1960) likelihood ratio statistic[1],

$$\text{QLR} = \max_{r=r_0,\ldots,r_1} F_T(r/T). \tag{2.4}$$

Quandt (1960) observed that the distribution of this statistic will not be χ_1^2 because the QLR statistic is the maximum of dependent χ_1^2 statistics. However, the limiting distribution of the QLR statistic can be obtained using the functional central limit theorem (FCLT), which has recently been used to great effect in the econometric theory of unit roots and cointegration. We therefore briefly digress to review the key elements of the FCLT.

Consider the partial sum of the errors,

$$\xi_T(s) = \sigma^{-1}T^{-\frac{1}{2}}\textstyle\sum_{t=1}^{[Ts]}\epsilon_t \tag{2.5}$$

where $[\cdot]$ is the greatest lesser integer function. For a fixed fraction s', $0 \le s' \le 1$, the standard central limit theorem implies that $\xi_T(s')$ has a N(0,s') distribution. The FCLT extends this to the stochastic process created treating ξ_T as a function of s. Accordingly, interpret (2.5) as referring to the random function of s which takes on the value given by the partial sum through [Ts], so that this random function is a step function. Then the FCLT states that,

$$\text{(2.6)} \qquad \xi_T \Rightarrow W$$

where W is a standard Brownian motion and "=>" denotes convergence in distribution on the space D[0,1] of functions with countably many discrete jumps.[2]

To obtain a limiting distribution for the QLR statistic, define $N_T(r/T) = SSR_{1,T} - (SSR_{1,r} + SSR_{r+1,T})$. Note that under the null, $SSR_{1,r} = \sum_{t=1}^{r}(y_t-\bar{y}_{1,r})^2 = \sum_{t=1}^{r}(\epsilon_t-\bar{\epsilon}_{1,r})^2 = \sum_{t=1}^{r}\epsilon_t^2-r\bar{\epsilon}_{1,r}^2$, where $\bar{\epsilon}_{1,r} = r^{-1}\sum_{t=1}^{r}\epsilon_t$, and that similar expressions apply to $SSR_{r+1,T}$ and $SSR_{1,T}$. Thus one obtains,

$$\text{(2.7)} \qquad \begin{aligned} N_T(r/T) &= r\bar{\epsilon}_{1,r}^2 + (T-r)\bar{\epsilon}_{r+1,T}^2 - T\bar{\epsilon}_{1,T}^2 \\ &= \sigma^2\{(T/r)\xi_T(r/T)^2 + (T/(T-r))(\xi_T(1)-\xi_T(r))^2 - \xi_T(1)^2\}. \end{aligned}$$

Thus N_T is a random function on the unit interval which is itself a continuous functional of ξ_T. By the continuous mapping theorem, the limiting representation of N_T is therefore obtained as that functional of the limiting representation of ξ_T.[3] That is, $N_T \Rightarrow \sigma^2F^*$, where

$$\text{(2.8)} \qquad F^*(s) = W(s)^2/s + \{W(1)-W(s)\}^2/(1-s) - W(1)^2.$$

Because $N_T \Rightarrow \sigma^2F^*$ and $SSR_{1,T}/T \xrightarrow{p} \sigma_\epsilon^2$ under the null, $(SSR_{1,r} + SSR_{r+1,T})/T \xrightarrow{p} \sigma^2$ uniformly in r in the sup-norm, that is, $P[\sup_{r\in[r_1,r_2]}|(SSR_{1,r} + SSR_{r+1,T})/T - \sigma^2| > \eta] \to 0$ for all $\eta>0$. Thus $F_T \Rightarrow F^*$.

The functional F* in (2.8) has a simpler representation, which is obtained by direct algebraic manipulation:

$$(2.9) \qquad F^*(s) = B_1^{\mu}(s)^2/(s(1-s))$$

where $B_1^{\mu}(s) = W(s)-sW(1)$ is a 1-dimensional Brownian bridge. This representation specializes to the usual χ_1^2 distribution when r is known and $r/T \rightarrow s'$, say, where s' is fixed. (Because B_1^{μ} is a 1-dimensional Gaussian process, F*(s') will be proportional to a χ_1^2. By direct evaluation, it can be shown that $EB_1^{\mu}(s')^2 = s'(1-s')$; thus the constant of proportionality is unity, so F*(s') has the usual χ_1^2 distribution for fixed s'.) The result (2.9) is stronger: it provides a representation for the limiting distribution of the random function F*, that is, the entire sequence of break-test statistics.

The final step in obtaining the distribution of the QLR statistic is to apply the continuous mapping theorem to (2.9). The supremum is a continuous functional, so the limiting null representation of the QLR statistic is,

$$(2.10) \qquad QLR = \max_{r=r_0,\ldots,r_1} F_T(r/T) \Rightarrow \sup_{s_0 \leq s \leq s_1} F^*(s)$$

where $r_0/T \rightarrow s_0$ and $r_1/T \rightarrow s_1$. Thus the QLR statistic has the asymptotic null distribution of the supremum of a squared Brownian bridge, normalized by its variance process s(1-s).

While (2.10) provides a formal answer to the question of the limiting distribution of the QLR statistic, it remains to obtain the critical values for use in practical work. Several approaches are available for computing these critical values, but the simplest is to recognize the limit (2.10) as stating the existence of a limiting distribution, which in turn can be computed by Monte Carlo simulation under the null for a suitably large sample size. This is equivalent to generating pseudo-data $\tilde{y}_t$ from a Gaussian random walk with $y_0 = 0$ and unit innovation variance and replacing W by its discretized realization. For example, $W(1)^2$ would be replaced by $(T^{-\frac{1}{2}}\tilde{y}_T)^2$. For T sufficiently large, the FCLT ensures that the limiting distribution of these pseudo-random variates converges to those of the functionals of Brownian motion. The main disadvantage of this approach is that high numerical accuracy requires many Monte Carlo repetitions.

The QLR test was developed with the alternative of a single break in mind, and other alternatives and test statistics can be constructed. Rather than discuss them here, however, these alternative approaches and extensions will be reviewed in the next section in the context of the standard linear time series regression model.

3. Tests for Structural Breaks in Time Series Regression

The leading application of tests for parameter instability is in the linear time series regression model,

(3.1) $$y_t = \beta_t' X_{t-1} + \epsilon_t$$

where X_t is a k-dimensional vector of regressors. Under the null hypothesis $\beta_t = \beta$ for all t. It is assumed that ϵ_t is a martingale difference sequence with respect to the σ-fields generated by $\{\epsilon_{t-1}, X_{t-1}, \epsilon_{t-2}, X_{t-2}, \ldots\}$ and that the regressors are constant and/or I(0) with $EX_tX_t' = \Sigma_X$ and possibly nonzero mean. For convenience, further assume that ϵ_t is conditionally (on past ϵ_t and X_t) homoskedastic. Also assume that $T^{-1}\sum_{t=1}^{[Ts]} X_tX_t' \xrightarrow{p} s\Sigma_X$ uniformly in s for $s \in [0,1]$. Note in particular that X_t can include lagged dependent variables as long as they are integrated of order zero (I(0)) under the null. A special case of these assumptions is that ϵ_t is i.i.d. and X_{t-1} is strictly exogenous and I(0).

This section examines three different classes of tests: tests for a single break date; tests based on recursive residuals and recursive coefficients; and tests against the time-varying parameter model. These tests are but a few of those available but are among the most relevant for the time series regression problem at hand. For further references on parameter instability and breaks, the reader is referred to Hackl and Westlund (1989, 1991) and Stock (1993).

3.1. Tests For a Single Break Date

The alternative hypothesis of a single break in at least one of the k coefficients is,

$$(3.2) \qquad \beta_t = \beta,\ t \le r, \quad \text{and } \beta_t = \beta + \delta,\ \ t > r$$

where r, $r_0 \le r \le r_1$, is the break date and δ is a k-dimensional vector. When the break date is known, a natural test for a change in β is the Chow (1960) test, which can be implemented in asymptotically equivalent Wald, Lagrange multiplier (LM), and LR forms. In LR form, the test for a break at a fraction r/T through the sample is,

$$(3.3) \qquad F_T(r/T) = T\{SSR_{1,T} - (SSR_{1,r} + SSR_{r+1,T})\}/(SSR_{1,r} + SSR_{r+1,T})$$

where $SSR_{1,r}$ is the sum of squared residuals from the estimating (3.1) with observations 1, . . . , r, etc. For $r/T \to s'$, where s' is fixed, $F_T(r/T)$ has an asymptotic χ_k^2 distribution under the null.

As in the case of a possible break in the mean, when the break date is unknown, the situation is more complicated. However, the QLR statistic (generalized by Davies (1977) to general models with parameters which are unidentified under the null) provides a natural test in this case, and is,

$$\text{QLR} = \max_{r=r_0,\ldots,r_1} F_T(r/T). \tag{3.4}$$

As in section 2, the QLR statistic does not have an asymptotic χ^2 distribution, but the FCLT can be used to obtain a limiting representation of the statistic as a functional of Brownian motion. The mathematical argument is broadly similar to that in section 2 but more involved and thus is not given here; for details, see Kim and Siegmund (1989) and, for a general treatment of "sup tests" in nonlinear models, Andrews (1993). The limiting representation of the QLR statistic is,

$$\text{QLR} \Rightarrow \sup_{s_0 \le s \le s_1} F_k^*(s), \text{ where } F_k^*(s) = B_k^\mu(s)' B_k^\mu(s)/(s(1-s)) \tag{3.5}$$

where $r_0/T \to s_0$, $r_1/T \to s_1$, and $B_k^\mu(s)$ is the k-dimensional Brownian bridge, $B_k^\mu(s) = W_k(s) - sW_k(1)$, where W_k is a k-dimensional Brownian motion. For fixed s', $F_k^*(s')$ has a χ_k^2 distribution.

The assumptions on the regressors used to obtain (3.5) hold if X_t contains a constant and/or I(0) regressors, but not if X_t is I(1). A sufficient condition for (3.5) *not* to hold is that the standard Chow test for fixed r/T does not have an asymptotic χ^2 distribution, since F*(s') has a χ^2 distribution for any fixed s'. In general this will occur for I(1) regressors and in these cases the derivations must be modified; see Banerjee, Lumsdaine and Stock (1992), Chu and White (1992) and Hansen (1992) for examples.

In principle the QLR statistic can be extended to more than one break date. A practical difficulty is that the computational demands increase sharply with the number of breaks (all values of the two-break F-statistic need to be computed for break dates (r, s) over the range of r and s), which makes numerical evaluation of the limiting distributions difficult for more than two or three break dates. More importantly, positing multiple breaks suggests that the breaks might better be thought of as continuous rather than discrete. This leads to a formulation in which the parameters change stochastically in each period by random amounts, as in the time-varying parameter model discussed in subsection 3.3.

3.2. Tests Based on Forecast Errors and Recursive Coefficient Estimates

Another approach to the detection of breaks is to examine recursive regression coefficients and/or forecast errors. The stochastic process of recursive regression coefficients is the random function $\tilde{\beta}(\cdot)$, where

$$\tilde{\beta}(r/T) = (\Sigma_{t=2}^{r} X_{t-1}X_{t-1}')(\Sigma_{t=2}^{r} X_{t-1}y_t) \qquad (3.6)$$

for $r_0 \leq r \leq r_1$. Plots of recursive coefficients are provided as a diagnostic option in some commonly used econometrics packages, for example PC-GIVE, and have been widely used in empirical work (e.g. Hendry and Ericsson (1991)). These tests typically have been proposed without reference to a specific alternative, although the most commonly studied alternative is a single structural break.

Because the recursive coefficients are evaluated at each point r, the null distribution of the recursive coefficients differs from the usual distribution of the OLS estimator. Ploberger, Krämer and Kontrus (1989) used the FCLT to obtain the asymptotic distribution of the sequence of recursive coefficients:

$$T^{\frac{1}{2}}(\tilde{\beta} - \beta) \Rightarrow \beta^*, \text{ where } \beta^*(s) = \sigma_\epsilon \Sigma_X^{-\frac{1}{2}} W_k(s)/s \qquad (3.7)$$

For fixed s', $\beta^*(s')$ has the usual asymptotic distribution of the OLS estimator. An important implication of (3.7) is that conventional "95%" confidence intervals plotted as bands around the path of recursive coefficient estimates are inappropriate since those bands fail to handle simultaneous inferences on the full plot of recursive coefficients.

Related to the recursive coefficients test are tests based on cumulative one-step-ahead forecast errors from recursive OLS estimation. This CUSUM test was originally proposed by Brown, Durbin and Evans (1975). An important feature of the CUSUM statistic is that, as shown by Krämer, Ploberger and Alt (1988), it has local asymptotic power only in the direction of the mean regressors: coefficient breaks of order $T^{-1/2}$ on mean-zero stationary regressors will not be detected. This has an intuitive explanation: the cumulation of the mean-zero regressor will remain mean zero (and will obey a FCLT) whether or not its true coefficient changes, while the nonzero mean of the cumulation of the constant implies that breaks in the intercept will result in systematically biased forecast errors. This is both a limitation and an advantage, for rejection suggests a particular alternative, namely instability in the intercept or the direction of the mean regressors.

Because the CUSUM test requires the construction of recursive residuals and thus running O(T) regressions, here we instead focus on a computationally simpler but similarly motivated test proposed by Ploberger and Krämer (1992) which uses full-sample residuals, $\{\hat{e}_t\}$. The Ploberger-Krämer (1992) (PK) test statistic is based on the cumulative OLS residuals, $\{\hat{e}_t\}$:

$$R_T(s) = \hat{\sigma}^{-1} T^{-1/2} \sum_{t=1}^{[Ts]} \hat{e}_t \qquad (3.8)$$

where $\hat{\sigma}$ is the usual standard error of the full-sample regression. The asymptotic distribution of R_T (treated as a random function) is readily obtained using the FCLT. To simplify the argument, suppose that the only regressor is a constant. Then under the null $\hat{e}_t = \epsilon_t - (\hat{\beta}-\beta) = \epsilon_t - (\bar{y}-Ey) = \epsilon_t - \bar{\epsilon}$, where $\bar{\epsilon} = T^{-1}\sum_{t=1}^{T}\epsilon_t$. Thus $R_T(s) = \hat{\sigma}^{-1}\{T^{-\frac{1}{2}}\sum_{t=1}^{[Ts]}\epsilon_t - sT^{-\frac{1}{2}}\sum_{t=1}^{T}\epsilon_t\} = (\sigma/\hat{\sigma})\{\xi_T(s)-s\xi_T(1)\}$. Now $\hat{\sigma} \overset{p}{\to} \sigma$, so by applying the FCLT we have,

(3.9) $$R_T \Rightarrow W\text{-}sW(1) = B_1^{\mu}.$$

Ploberger and Krämer (1992) show that this same result holds when additional I(0) regressors X_t are included in the regression.

The result (3.9) suggests that parameter stability can be rejected when the regression produces too large a cumulative deviation of forecast errors in one direction. This reasoning suggests two test statistics, the maximum and a mean square of R_T, which we respectively refer to as the Ploberger-Krämer max and Ploberger-Krämer meansq statistics. These statistics and their limiting distributions, obtained from (3.9) and the continuous mapping theorem, are:

(3.10) $$\text{PK-max} = \max_{1\leq r\leq T}|R_T(r/T)| \Rightarrow \sup_{0\leq s\leq 1}|B_1^{\mu}(s)|$$

(3.11) $$\text{PK-meansq} = T^{-1}\sum_{r=1}^{T}R_T(r/T)^2 \Rightarrow \int_{s=0}^{1}B_1^{\mu}(s)^2 ds.$$

If one knew *a-priori* that the break must have occured after some date r_0, the lower limit 1 of the statistics in (3.10) and (3.11) could be replaced by r_0 and the lower limit of the asymptotic representations would be $s_0 = \lim_T r_0/T$. The empirical work in the next section uses r_0=1, that is, the statistics in (3.10) and (3.11).

3.3. Tests Against the Time-Varying-Parameter Model

A flexible extension of the standard regression model is to suppose that the regression coefficients evolve over time, specifically, (3.12) $$y_t = \beta_t' X_{t-1} + \epsilon_t, \quad \beta_t = \beta_{t-1} + \nu_t, \quad \mathrm{var}(\nu_t) = \tau^2$$

where ϵ_t and ν_t are uncorrelated and ν_t is serially uncorrelated. The formulation (3.12) nests the standard linear regression model by letting $\tau^2=0$. By setting $\nu_t=\delta$, t=r+1 and $\nu_t=0$, $t \neq r+1$, (3.12) nests the single-break model (3.2). The alternative of specific interest here, however, is when ν_t is i.i.d. $N(0, \tau^2 G)$ (where G is assumed to be known) so that the coefficient β_t follows a multivariate random walk and thus evolves smoothly but randomly over the sample period. When combined with the additional assumption that ϵ_t is i.i.d. $N(0, \sigma_\epsilon^2)$, this is referred to as the "time-varying parameter" (TVP) model (see Cooley and Prescott (1976) and the reviews by Chow (1984) and Nicholls and Pagan (1985)).

Maximum likelihood estimation of the TVP model is a direct application of the Kalman filter (β_t is the unobserved state vector, $\beta_t=\beta_{t-1}+\nu_t$ is the state equation, and $y_t = \beta_t' X_{t-1} + \epsilon_t$ is the measurement equation) and the estimation of β_t and its standard error under the alternative is well understood; see Harvey (1989) and Hamilton (1993). Here, we therefore focus on the problem of testing the null that $\tau^2=0$.

Starting with Nyblom and Mäkeläinen (1983), several authors have studied the properties of locally most powerful tests of $\tau^2=0$ against $\tau^2>0$ in (3.12) or in models where only some of the coefficients are assumed to evolve over time (that is, where G has reduced rank). Nyblom (1989) derived the locally most powerful test of $\tau^2 = 0$ vs. $\tau^2 > 0$. In general this test depends on G. To obtain a simple expression Nyblom (1989) suggests setting $G = (T^{-1}\sum_{t=1}^{T} X_{t-1}X_{t-1}')^{-1}$ and accordingly obtains the test which rejects for large values of the statistic,

(3.13) $$L = T^{-1}\sum_{s=1}^{T} S_T(s/T)'(\hat{\sigma}^2 T^{-1}\sum_{t=1}^{T} X_{t-1}X_{t-1}')^{-1} S_T(s/T)$$

where $S_T(s/T) = T^{-½}\sum_{t=s+1}^{T}\hat{e}_t X_{t-1}$, where $\{\hat{e}_t\}$ are the OLS residuals from the full-sample estimation of (3.1) under the null. Conditional on $\{X_t\}$ the TVP model induces a heteroskedastic random walk into the error term, and the statistic L detects this using the cumulated product of the OLS residuals and the regressors.

Nyblom (1989) derived the statistic (3.13) by applying local arguments to a likelihood for generally nonlinear, nonnormal models, and his general statistic simplifies to (3.13) in the Gaussian linear regression model. If $X_t = 1$, (3.13) reduces to the locally most powerful invariant test of the null that y_t is i.i.d. Gaussian against the alternative that y_t is the sum of independent Gaussian i.i.d. and random walk components.

The asymptotics of the Nyblom statistic L also follow from the FCLT and the continuous mapping theorem. Under weak conditions, $\epsilon_t X_{t-1}$ is a martingale difference sequence, so by the FCLT and algebraic manipulations, under the null hypothesis,

(3.14) $$L \Rightarrow \int_{s=0}^{1} B_k^{\mu}(s)'B_k^{\mu}(s)ds.$$

4. Structural Breaks in the Money-Income Relation

A cornerstone of conventional arguments for the use of monetary aggregates as the instruments of monetary policy is that there is an exploitable link between movements in the monetary aggregates and output. There is now a vast empirical literature which examines this link. Since the work of Sims (1972, 1980), a leading tool for studying this link has been Granger causality tests and vector autoregressions; see for example Feldstein and Stock (1993), Friedman and Kuttner (1992) and Konishi, Ramey and Granger (1992). For there to be an exploitable link from money to output requires this relationship to be stable. This section therefore uses the tests for parameter stability discussed above to reexamine the link between two monetary aggregates, M1 and M2, and real GDP in the United States. Most of the literature on money Granger causality tests has focused on the money-real output relation, and this is the relation studied here. For results on the money-nominal output relation, see Feldstein and Stock (1993).

The data are quarterly for the United States, 1959:1 - 1992:4. In addition to M1, M2, and real GDP, some specifications include two short-term interest rates: the 90-day U.S. Treasury bill rate and the 6-month commercial paper rate. The commercial paper rate, and in particular the spread between the commercial paper rate and the Treasury bill rate, has been highlighted elsewhere as an historically useful predictor of output, and so is included here (see Stock and Watson (1989a) and Friedman and Kuttner (1993)). (Note however that these previous studies used the matched maturity 6-month commercial paper and Treasury bill rates, whereas the 90-day Treasury bill rate is used here.) Quarterly data on M1, M2, and the interest rates were obtained by averaging the monthly data over the quarter. All data were taken from the Citibase data base.

4.1. Preliminary Unit Root and Cointegration Analysis

Numerous researchers have examined the unit root and cointegration properties of these data and rather than reproduce those results we summarize them here briefly. Most recent studies of the money-income relation which use postwar U.S. data have modeled the monetary aggregates and real GDP as each having a unit root in logarithms, so that the stationary relationships are specified in growth rates. In addition, these studies tend to conclude that interest rates are integrated of order one so that they are I(0) either in first differences or in growth rates. This finding is generally based on the application of augmented Dickey-Fuller (1979) (ADF) pretests;

see in particular Miller (1991), Friedman and Kuttner (1992), Konishi, Ramey and Granger (1993), and Hoffman and Rasche (1991). Other researchers have focused on longer time series of money and output and have found complementary results, in particular Hafer and Jansen (1991) (using data from 1915-1988) and Stock and Watson (1993a) (using data from 1900-1988) concluded that the unit root model provided a good approximation to the univariate properties of these series over the twentieth century. These conclusions are of course the same as Nelson and Plosser's (1982) seminal findings that the unit root could not be rejected in these long annual time series using ADF tests.

Other research has, however, suggested that the unit root model might inadequately describe these data. Stock (1991) computed asymptotic confidence intervals for the largest autoregressive roots in the Nelson-Plosser (1982) series and found that, while the unit root could not be rejected, neither could a wide range of stationary roots. For example, for real GNP, the 90% confidence interval for the largest autoregressive root is (.78, 1.07), based on 62 annual observations. Using postwar quarterly data, Stock and Watson (1989b) suggested that money is better described as I(1) with a quadratic trend, although this result is not robust to the addition of data from the late 1980's and early 1990's. Also using postwar quarterly data, Christiano and Ljungqvist (1988) suggested that money is best modeled as having a mildly explosive root (although one might conjecture that this too is not robust to including the last five years of data). DeJong and Whiteman (1991) used Bayesian techniques to reexamine the Nelson-Plosser (1982) data and concluded that real output in particular was trend stationary rather than difference stationary. However, Bayesian inference about unit roots has been shown to be sensitive to the choice of priors, which is difficult to resolve on objective scientific grounds because of the peculiar nature of the distributions when the largest root is nearly one; see Phillips (1991) and the associated discussion in the special issue of the *Journal of Applied Econometrics* on Bayesian unit root analysis.

While this brief survey indicates that uncertainty remains about the size of the largest autoregressive roots in money, real output and interest rates, it suggests that most of the evidence is that the unit root model cannot be rejected. In any event, the unit root framework (and the resulting estimation of Granger causality relations in growth rates) has become standard in this literature. Therefore this paper adopts the assumption that the growth rates of real GDP (Δy), real M1 (deflated by the GDP deflator; $\Delta m1$) and real M2 (GDP deflator; $\Delta m2$) are I(0), and that the first differences of the level of the Treasury bill rate (ΔR_{TB}) and the commercial paper rate (ΔR_{CP}) are I(0).

The final step in preliminary data analysis is to ascertain whether there is cointegration among any of these variables. A conventional approach to this is to use multivariate cointegration pretests, such as the Johansen (1988) or Stock-Watson (1988) tests, to estimate the number of unit roots and cointegrating vectors in the system. However, Monte Carlo evidence suggests that these procedures can exhibit very large size distortions in finite samples (see for example Haug (1993)). When combined with the low size-adjusted power of these tests, this suggests that the general multivariate cointegration tests are unlikely to yield reliable results, especially with only three decades of data. We therefore instead rely on economic theory and evidence from longer data sets to guide inference about the relevant orders of cointegration.

Economic theory suggests two potential cointegrating vectors among the five-variable system: a long-run money demand relation and the stationarity of the risk premium on commercial paper over Treasury bills:

$$m_t = \beta_y y_t + \beta_R R_{TB,t} + z_{MDt} \tag{4.1}$$

$$R_{CP,t} = \gamma_{CP\text{-}TB} R_{TB,t} + z_{Rt} \tag{4.2}$$

where z_{MDt} and z_{Rt} are stationary disturbances. If the cointegrating relation among the interest rates (4.2) holds, then the long-run money demand relation could equivalently be expressed using the commercial paper rate, but that specification would be redundant in the sense that it spans the same space of cointegrating vectors and is therefore ignored henceforth. Economic theory also suggests some of the coefficients in these cointegrating relations. Theories of money demand often emphasize unit income elasticities, so a natural candidate for β_y is unity. If the tax and institutional structure of the commercial paper and Treasury bill markets is constant over the period, then it is plausible that the risk premium $R_{CP,t}$-$R_{TB,t}$ is I(0), that is, $\gamma_{CP\text{-}TB}$ is unity.

The potential cointegrating relations (4.1) and (4.2) are examined empirically in table 1. The first row reports ADF tests of the hypothesis that velocity has a unit root, and the second row reports the ADF-GLS test proposed by Elliott, Rothenberg and Stock (1992), which they showed to be nearly asymptotically efficient. and substantially more powerful than standard ADF. Because no cointegrating coefficients were estimated to obtain velocity, the usual univariate critical values for these two tests can be used. The number of lagged first differences to include when computing the test statistic was obtained by sequentially testing the hypothesis that the highest-order coefficient is zero, as suggested by Ng and Perron (1993). This procedure is an asymptotically justifiable data-dependent method for lag length selection. Here, the largest number of lags considered was 8, the smallest 1; the lag lengths actually chosen ranged between 1 and 7. We interpret the theory of a stable long-run money demand function as excluding the possibility of a deterministic trend, so the tests are against the alternative that the series at hand is stationary around a constant mean. The statistics in table 1 indicate that the hypothesis of noncointegration of real money and output cannot be rejected at the 10% level for M1. Although the ADF tests fails to reject noncointegration for M2 at the 10% level, the more powerful ADF-GLS test rejects the hypothesis that M2 velocity is I(1) at the 5% level.

Economic theory suggests interest sensitivity of money demand, particularly for M1 which bears no interest. To circumvent the finite-sample reduction of power and size distortions associated with estimation of the interest semielasticity using the postwar data set, the next two rows of table 1 test for the stationarity of a money demand residual, constructed by imposing unit interest elasticity and by using an estimate of the interest semielasticity from 1903-1945 obtained from Stock and Watson (1993a, table 5). Because the semielasticity is based on data outside the current sample, the unit root hypothesis can be tested here using conventional Dickey-Fuller tests and critical values. Here, the results are the opposite as for velocity: the long-run money demand residual for M1 is stationary, but not for M2.

The next row provides Engle-Granger (1987) ADF tests for cointegration based on an in-sample OLS estimate of the interest semielasticity. Because the OLS residual is used, the critical values are taken from MacKinnon (1991). Interestingly, for neither M1 nor M2 is the hypothesis of non-cointegration rejected. This need not, however, be inconsistent with the rejection of noncointegration found in the previous tests, for the inclusion of the additional step of estimating a cointegrating vector reduces power and the previous two lines used *a-priori* restrictions to avoid this.

Estimates of the semielasticity appear in the final two rows of panel A. The estimates are based on, respectively, the OLS regression of log velocity on a constant and the level of the 90-day Treasury bill rate, and the Phillips-Loretan (1991)/Saikkonen (1991)/Stock-Watson (1993a) dynamic OLS (DOLS) estimator.[4] The latter estimator is asymptotically efficient in the sense made precise by Saikkonen (1991) and inference is χ^2, so t-statistics and confidence intervals can be computed as usual using the reported standard errors. The DOLS estimates are consistent with the ADF and ADF-GLS test results in the first four rows: for M1, the point estimate is not statistically different from the prewar estimate of -.104, while for M2 the point estimate is much smaller although it is statistically different from zero. In summary, the evidence of panel A points to M1, real output and the interest rate being cointegrated, with a large interest semielasticity (-.068 ± .048). M2, real output and the interest rate also appear to be cointegrated, but with a smaller interest elasticity. The smaller estimated semielasticity for M2 makes economic sense, for M2 includes many interest-bearing instruments such as money market mutual funds while M1 bears essentially no interest.

Panel B of table 1 reports similar statistics for the relation between R_{TB} and R_{CP}. Based on the ADF and ADF-GLS statistics, the spread appears to be integrated of order zero. The DOLS estimator of the cointegrating vector is very nearly one and a 95% two-sided confidence interval for the cointegrating coefficient is (0.966, 1.086). This provides support for the subsequent imposition of cointegration between the two interest rates with a unit cointegrating coefficient.

4.2. Tests for Predictive Content and Parameter Stability

We now turn to tests of predictive content and parameter stability in regressions using M1 or M2 to predict quarterly real GDP growth. Each regression includes a constant and lags of real GDP growth as well as lags of the candidate predictors, so the F-tests of the exclusion restrictions are Granger causality tests. When the 90-day Treasury bill and the commercial paper rate are both included, ΔR_{TB} and the paper-bill spread are used. The cointegrating residual z_{MD} is computed as $m_t - y_t - \hat{\beta}_{R,DOLS} R_{TB,t}$, where $\hat{\beta}_{R,DOLS}$ is the DOLS estimate of the interest semielasticity reported in the last row of table 1, panel A, with the M1 estimate of β_R used in the M1 equations and the M2 estimate of β_R used in the M2 equations. Four lags are used for all variables except the M2 money demand residual z_{MD}, for which only the first lag is included to avoid a singlular regressor matrix. When z_{MD} is included, the Granger causality test statistic for money is augmented to include the restriction that the coefficient on z_{MD} is zero.

Granger causality test statistics and p-values are reported in table 2. The patterns are typical of those found elsewhere in the money-income causality literature. With the exception of the bivariate regression 1, the marginal predictive content of M1 in the various regressions is statistically insignificant. However, interest rates and the paper-bill spread are significant. The results for M2 differ, with money growth being statistically significant at the 5% level in the regressions with the Treasury bill rate. While M2 is insignificant in regressions 4 and 5 with the commercial paper rate, when the cointegrating residual is included M2 is again significant at the 10% level even with the paper-bill spread included (regression 6). Including the paper-bill spread weakens but does not eliminate the predictive content of M2 for real output.

Table 3 presents tests of the stability of the coefficients in the regressions in table 2. Asymptotic critical values were computed by simulation of the null limiting distributions with T=500 and 8,000 Monte Carlo replications. The QLR test was computed with 15% symmetric trimming, that is, r_0=[.15T] and r_1=[.85T]. Because M1 is not significant in regressions 2-6 of panel A, the tests should be interpreted not of the stability of the inconsequential M1-GDP relationship but rather of the stability of the interest rate-output relations. There is evidence of instability in the Treasury-bill specifications, with each specification having two tests which reject at the 10% level or less. The evidence against stability is strongest in specifications 5 and 6, which include the paper-bill spread.

The results for M2 are qualitatively different. In the first four specifications, stability is not rejected at the 10% level by any of the statistics. While M1 is not a useful stable predictor (nor is the Treasury bill rate along with M1), the M2/interest rate specifications, in particular the error correction model 3, does appear to be stable. However, when the paper-bill spread is included, the stability of the relationship breaks down, with the QLR test rejecting stability at the 5% level in both specifications 5 and 6.

The instability of the specifications with the paper-bill spread accords with other evidence that its predictive content for real output was greatest in the 1970's and early 1980's and has declined subsequently (Bernanke (1990), Stock and Watson (1993b)). Judged by both predictive content and parameter stability, we are thus left with the M2 error correction specification, which includes the Treasury bill rate and the long-run M2 demand residual (equation 3 in panel B). The QLR, Ploberger-Krämer, and Nyblom tests do not detect parameter instability in this specification, and money growth and the error correction term are jointly statistically significant with a p-value of 2%.

5. Implications for Empirical Work and Discussion

The empirical work in section 4 used preliminary data analysis (unit root and cointegration analysis) to determine the class of empirical specifications. Given these specifications, their stability was then assessed using the tests from section 3. An important part of the preliminary analysis was determining whether there is a long run money demand relation and cointegration between the interest rates. Because of the poor sampling properties of cointegration tests, the analysis emphasized imposing as much economic theory as possible, which led to numerical values for two of the three cointegrating parameters. The tests suggested that cointegrating relations exist for both real M1 and M2 but that the interest semielasticity for M2 is much smaller than for M1. In addition, the spread between the commercial paper and Treasury bill rates appears to be integrated of order zero.

Several caveats should be emphasized. While the break tests examine stability within the sample, they do not of course guarantee stability out of sample. Ideally, the tests would be augmented by out-of-sample evidence on stability. Also, this analysis has been based on asymptotic distributions, and Monte Carlo evidence suggests that there are some size distortions in finite-sample applications of these break point tests. However, these size distortions are small when compared with those found for unit root/cointegration tests (Diebold and Chen

(1992)). Even with these caveats, the empirical evidence in section 4 suggests that the money growth error correction specification is a plausible candidate for forecasting real output. One might speculate further that this stable link between money and real output could be exploited by the monetary authorities to achieve better economic stabilization. However, this further step requires additional controversial assumptions on the ability of the central bank to use these reduced-form relations for macroeconomic control and indeed on its ability to control money growth itself.

Footnotes

1. Quandt (1960) proposed computing the maximum of the likelihood ratio statistics, but (2.4) is instead the maximum of the Wald F-statistics. The two tests are asymptotically equivalent and although the test studied below is the maximal Wald test, we refer to it as the QLR test.

2. An excellent text on the FCLT is Hall and Heyde (1980). Extensions of the FCLT to dependent errors are discussed in Hall and Heyde (1980), Herrndorf (1984), and Phillips and Solo (1992).

3. The continuous mapping theorem states that if $f(\cdot)$ is a continuous function and if $Z_T \Rightarrow Z$, then $f(Z_T) \Rightarrow f(Z)$ (e.g. Hall and Heyde (1980)).

4. The DOLS estimator was computed using 2 leads and lags of ΔR_{TB} and its standard errors were computed by modeling the regression error as an AR(4); see Stock and Watson (1993a) for details.

References

Anderson, T.W. and D. Darling (1952), "Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Processes," *Annals of Mathematical Statistics*, 23, 193-212.

Andrews, D.W.K. (1993), "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 821-856.

Banerjee, A., Lumsdaine, R.L., and Stock, J.H. (1992), "Recursive and Sequential Tests of the Unit Root and Trend Break Hypotheses: Theory and International Evidence," *Journal of Business and Economic Statistics*, 10, 271-288.

Bernanke, B. (1990), "On the Predictive Power of Interest Rates and Interst Rate Spreads," *New England Economic Review*, November/December, 51-68.

Brown, R.L., Durbin, J. and Evans, J.M. (1975), "Techniques for Testing the Constancy of Regression Relationships over Time with Comments," *Journal of the Royal Statistical Society, Series B*, 37, 149-192.

Chow, G.C. (1960), "Tests of Equality Between Sets of Coefficients in Two Linear Regressions," *Econometrica*, 28, 591-605.

Chow, G.C. (1984), "Random and Changing Coefficient Models," in Z. Griliches and M. Intriligator (eds.), *Handbook of Econometrics, Vol. 2.* Amsterdam: North Holland.

Christiano, L.J. and L. Ljungqvist (1988), "Money Does Granger-Cause Output in the Bivariate Money-Output Relation," *Journal of Monetary Economics*, 22, 217-236.

Chu, C-S. J., and H. White (1992), "A Direct Test for Changing Trend," *Journal of Business and Economic Statistics*, 10, 289-299.

Cooley, T.F. and E.C. Prescott (1976), "Estimation in the Presence of Stochastic Parameter Variation," *Econometrica*, 44, 167-184.

Davies, R.B. (1977), "Hypothesis Testing When a Nuisance Parameter is Present only Under the Alternative," *Biometrika*, 64, 247-254.

Diebold, F.X. and C. Chen (1992), "Testing Structural Stability with Endogenous Break Point: A Size Comparison of Analytic and Bootstrap Procedures," manuscript, Department of Economics, University of Pennsylvania.

DeJong, D.N and C.H. Whiteman, (1991), "Reconsidering 'Trends and Random Walks in Macroeconomic Time Series'," *Journal of Monetary Economics*, 28, no. 2, 221-254.

Dickey, D.A. and Fuller, W.A. (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," *Journal of the American Statistical Association*, 74, no. 366, 427-431.

Elliott, G., T.J. Rothenberg, and J.H. Stock (1992), "Efficient Tests for an Autoregressive Unit Root," NBER Technical Working Paper no. 130.

Engle, R.F. (1983), "Hypothesis Tests in Econometrics," in Z. Griliches and M. Intrilligator (eds.), *Handbook of Econometrics, Vol. 2.* Amsterdam: North Holland.

Engle, R. F. and C.W.J. Granger (1987), "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251-276.

Feldstein, M. and J.H. Stock (1993), "The Use of a Monetary Aggregate to Target Nominal GDP," NBER Working Paper no. 4304.

Friedman, B.M. and K.N. Kuttner (1992), "Money, Income, Prices and Interest Rates," *American Economic Review*, 82, 472-492.

Friedman, B.M. and K.N. Kuttner (1993), "Why Does the Paper-Bill Spread Predict Real Economic Activity," in J.H. Stock and M.W. Watson, eds., *New Research on Business Cycles, Indicators and Forecasting*, University of Chicago Press for the NBER.

Hackl, P. and A. Westlund (1989), "Statistical Analysis of 'Structural Change': An Annotated Bibliography," *Empirical Economics*, 143, 167-192.

Hackl, P. and A.H. Westlund (eds) (1991), *Economic Structural Change: Analysis and Forecasting*, Springer-Verlag, Berlin, 95-119.

Hall, P. and C.C. Heyde (1980), *Martingale Limit Theory and its Applications.* New York: Academic Press.

Hamilton, J. (1993), "State Space Models," forthcoming, *Handbook of Econometrics, Vol. IV.* Amsterdam: Elsevier.

Hansen, B.E. (1992), "Tests for Parameter Instability in Regressions with I(1) Processes," *Journal of Business and Economic Statistics*, 10, 321-336.

Harvey, A.C. (1989), *Forecasting, Structural Models and the Kalman Filter*, Cambridge, U.K.: Cambridge University Press.

Haug, A.A. (1993), "Tests for Cointegration: A Monte Carlo Comparison," manuscript, Department of Economics, York University, North York, Ontario, Canada.

Hendry, D.F. and Ericsson, N.R. (1991), "An Econometric Analysis of U.K. Money Demand in *Monetary Trends in the United States and the United Kingdom* by Milton Friedman and Anna J. Schwartz," *American Economic Review*, 81, 8-38.

Herrndorf, N. A. (1984), "A Functional Central Limit Theorem for Weakly Dependent Sequences of Random Variables," *Annals of Probability* 12, 141-153.

Hoffman, D. and R.H. Rasche (1991), "Long-Run Income and Interest Elasticities of Money Demand in the United States," *Review of Economics and Statistics* 73, 665-674.

Johansen, S. (1988), "Statistical Analysis of Cointegrating Vectors," *Journal of Economic Dynamics and Control*, 12, 231-54.

Kim, H-J. and D. Siegmund (1989), "The likelihood Ratio Test for a Change-Point in Simple Linear Regression," *Biometrika*, 76, 3, 409-23.

Krämer, W., Ploberger, W. and Alt, R. (1988), "Testing for Structural Change in Dynamic Models," *Econometrica*, 56, 1355-1370.

Konishi, T., V.A. Ramey and C.W.J. Granger (1992), "Stochastic Trends and Short-Run Relationships Between Financial Variables and Real Activity," manuscript, Department of Economics University of California – San Diego.

MacKinnon, J. (1991), "Critical Values for Cointegration Tests," ch. 12 in R.F. Engle and C.W.J. Granger (eds.), *Long-Run Economic Relationships*. Oxford: Oxford University Press.

Nelson, C.R., and Plosser, C.I. (1982), "Trends and Random Walks in Macro-economic Time Series: Some Evidence and Implications," *Journal of Monetary Economics* 10, 139-162.

Ng, S. and P. Perron (1993), "Unit Root Tests in ARMA Models With Data-Dependent Methods for the Selection of the Truncation Lag," manuscript, C.R.D.E., University of Montreal, Quebec.

Nicholls, S. and A. Pagan (1985), "Varying Coefficient Regression," *Handbook of Statistics*, 413-449.

Nyblom, J. (1989), "Testing for the Constancy of Parameters Over Time," *Journal of the American Statistical Association*, 84, 223-30.

Nyblom, J. and T. Mäkeläinen (1983), "Comparisons of Tests for the Presence of Random Walk Coefficients in a Simple Linear Model," *Journal of the American Statistical Association*, 78, 856-864.

Phillips, P.C.B. (1991), "To Criticize the Critics: An Objective Bayesian Analysis of Stochastic Trends," *Journal of Applied Econometrics* 6, 333-364.

Phillips, P.C.B. and M. Loretan (1991), "Estimating Long Run Economic Equilibria," *Review of Economic Studies*, 58, 407-436.

Phillips, P.C.B. and V. Solo (1992), "Asymptotics for Linear Processes", *Annals of Statistics*, 20, no. 2, 971-1001.

Ploberger, W., W. Krämer, and K. Kontrus (1989), "A New Test for Structural Stability in the Linear Regression Model," *Journal of Econometrics*, 40, 307-318.

Ploberger, W. and W. Krämer (1992), "The CUSUM test with OLS Residuals," *Econometrica*, 60, 271 - 286.

Quandt, R.E. (1960), "Tests of the Hypothesis that a Linear Regression System Obeys Two Separate Regimes," *Journal of the American Statistical Association*, 55, 324-330.

Saikkonen, P. (1991), "Asymptotically Efficient Estimation of Cointegrating Regressions," *Econometric Theory*, 7, 1-21.

Sims, C.A. (1972), "Money, Income and Causality," *American Economic Review*, 62, 540-552.

Sims, C.A. (1980), "Macroeconomics and Reality," *Econometrica*, 48, 1-48.

Stock, J.H. and M.W. Watson (1988), "Testing for Common Trends," *Journal of the American Statistical Association*, 83, 1097-1107.

Stock, J.H. and M.W. Watson (1989a), "New Indexes of Leading and Coincident Economic Indicators," *NBER Macroeconomics Annual 1989*, 351-394.

Stock, J.H. and M.W. Watson (1989b), "Interpreting the Evidence on Money-Income Causality," *Journal of Econometrics*, 40, no. 1, 161-182.

Stock, J.H. and M.W. Watson (1993a), "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems", *Econometrica*, 61, 783-820.

Stock, J.H. and M.W. Watson (1993b), "A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience," in J.H. Stock and M.W. Watson (eds.), *Business Cycles, Indicators and Forecasting*, University of Chicago Press for the NBER.

Stock, J.H. (1991), "Confidence Intervals for the Largest Autoregressive Root in U.S. Economic Time Series," *Journal of Monetary Economics*, 28, no. 3, 435-460.

Stock, J.H. (1993), "Unit Roots, Structural Breaks and Trends," forthcoming in R. Engle and D. McFadden (eds.), *Handbook of Econometrics, Volume IV*. Amsterdam: Amsterdam.

Table 1
Cointegration Tests: Velocity and Long-Run Money Demand

1959:1 - 1992:2

Variable	M1	M2
A. Money Demand and Velocity		
v		
ADF	-2.34	-2.46
ADF-GLS	-0.22	-2.46**
$v-.104R_{TB}$		
ADF	-3.44***	-2.00
ADF-GLS	-3.42***	-1.14
$v+\hat{\beta}_{R,OLS}R_{TB}$		
ADF	-1.47	-1.95
$\hat{\beta}_{R,OLS}$	-.060	-.0060
$\hat{\beta}_{R,DOLS}$	-.068	-.0064
(SE)	(.024)	(.0023)
B. Commercial Paper Risk Premium		
$R_{CP}-R_{TB}$		
ADF	-4.43***	
ADF-GLS	-3.87***	
$\hat{\gamma}_{CP-TB,OLS}$	1.041	
$\hat{\gamma}_{CP-TB,DOLS}$	1.032	
(SE)	(.033)	

Notes: The lag length for the augmented Dickey-Fuller (1979) (ADF) tests was chosen by sequential 10% downward likelihood ratio tests, with a maximum of 8 lags and a minimum of 1 lag. The tests in the first two rows of panel A and in panel B do not involve in-sample estimation of cointegrating vectors so critical values are obtained from Dickey and Fuller (1979). The test in the third row of panel A entails estimation of a single cointegrating vector and critical values are obtained from MacKinnon (1991). The cointegrating coefficient estimators are discussed in the text. A constant but no time trend was included in all regressions. The ADF regressions were run over 1961:2-1992:2 with earlier observations used for initial conditions. The OLS and DOLS cointegrating regressions were run over all available data (allowing for leads and lags in the DOLS estimators).

Significant at the *10%, **5%, ***1% level.

Table 2
Predictive Content of M1 and M2

Dependent Variable: Real GDP Growth
Estimation period: quarterly, 61:2 to 92:2

Eq.	Regressors	F-tests (p-values) on lags of: Δm	ΔR_{TB}	ΔR_{CP}	R_{TB}-R_{CP}
		A. Tests of M1			
1	Δy, $\Delta m1$	2.83 (0.028)			
2	Δy, $\Delta m1$, ΔR_{TB}	1.11 (0.356)	5.19 (0.001)		
3	Δy, $\Delta m1$, ΔR_{TB}, z_{md}	0.88 (0.497)	4.88 (0.001)		
4	Δy, $\Delta m1$, ΔR_{CP}	0.34 (0.853)		5.99 (0.000)	
5	Δy, $\Delta m1$, ΔR_{TB}, R_{CP}-R_{TB}	0.58 (0.677)	3.87 (0.006)		3.20 (0.016)
6	Δy, $\Delta m1$, ΔR_{TB}, R_{CP}-R_{TB}, z_{md}	0.49 (0.784)	3.80 (0.006)		3.21 (0.016)
		B. Tests of M2			
1	Δy, $\Delta m2$	8.11 (0.000)			
2	Δy, $\Delta m2$, ΔR_{TB}	2.81 (0.029)	2.13 (0.082)		
3	Δy, $\Delta m2$, ΔR_{TB}, z_{md}	2.80 (0.020)	2.05 (0.092)		
4	Δy, $\Delta m2$, ΔR_{CP}	1.58 (0.185)		2.43 (0.052)	
5	Δy, $\Delta m2$, ΔR_{TB}, R_{CP}-R_{TB}	1.73 (0.149)	1.97 (0.105)		2.73 (0.033)
6	Δy, $\Delta m2$, ΔR_{TB}, R_{CP}-R_{TB}, z_{md}	2.05 (0.078)	1.87 (0.120)		2.88 (0.026)

Notes: In the cases that z_{md} is included, the test for the significance of money includes the restriction that the coefficient on the error correction term z_{md} is zero. All regressions include a constant and the first through fourth lag of the indicated regressors, except for z_{md}, of which only the first lag is included. P-values (given in parentheses) are computed using the usual F distribution. Earlier observations are used as initial conditions.

Table 3

Tests for Structural Breaks and Time-Varying Parameters with M1 and M2

Dependent Variable: Real GDP Growth

Estimation period: quarterly, 61:2 to 92:2

Eq.	Regressors	QLR	P-K max	P-K meansq	Nyblom L
	A. *Tests of M1*				
1	Δy, $\Delta m1$	25.45**	1.16	0.50**	1.53
2	Δy, $\Delta m1$, ΔR_{TB}	32.08**	0.90	0.38*	1.48
3	Δy, $\Delta m1$, ΔR_{TB}, z_{md}	34.93**	0.87	0.35*	1.73
4	Δy, $\Delta m1$, ΔR_{CP}	30.37*	1.01	0.42*	1.55
5	Δy, $\Delta m1$, ΔR_{TB}, $R_{CP}-R_{TB}$	45.52***	0.98	0.45*	1.93
6	Δy, $\Delta m1$, ΔR_{TB}, $R_{CP}-R_{TB}$, z_{md}	46.84***	0.92	0.36*	2.08
	B. *Tests of M2*				
1	Δy, $\Delta m2$	19.58	0.60	0.08	1.11
2	Δy, $\Delta m2$, ΔR_{TB}	24.42	0.65	0.12	1.32
3	Δy, $\Delta m2$, ΔR_{TB}, z_{md}	27.23	0.85	0.26	1.69
4	Δy, $\Delta m2$, ΔR_{CP}	27.52	0.79	0.17	1.46
5	Δy, $\Delta m2$, ΔR_{TB}, $R_{CP}-R_{TB}$	38.99**	0.72	0.16	1.66
6	Δy, $\Delta m2$, ΔR_{TB}, $R_{CP}-R_{TB}$, z_{md}	42.83**	1.02	0.35*	2.05

Notes: The regressions under the null of parameter stability are the same as in table 2 and the test statistics are described in the text. See the notes to table 2. Significant at the *10%, **5%, ***1% level.