Testing Long-Run Monetary Neutrality Propositions: Lessons from the Recent Research

James Bullard

Monetary economists long have thought that government injections of money into a macroeconomy have a certain neutral effect. The main idea is that changes in the money stock eventually change nominal prices and nominal wages, ultimately leaving important real variables, like real output, real consumption expenditures, real wages, and real interest rates, unaffected. Since economic decision making is based on real factors, the long-run effect of injecting money into the macroeconomy is often described as *neutral*—in the end, real variables do not change and so economic decision making is also unchanged. How long such a process takes, and what might happen in the meantime, are hotly debated questions. But relatively few economists debate the merits of *long-run neutrality*. Indeed, long-run neutrality is instead taken as a given, almost an axiom, a logical consequence of suppositions made in economic theory.

Curiously, during most of the postwar period the empirical evidence on long-run monetary neutrality has been in a state of flux. No doubt this is in part because it is difficult to look at the data generated by the world's economies and come to any

firm conclusion about whether monetary injections had important real effects, in the short run or in the long run. In addition, many of the empirical tests that were devised ran into important criticisms that seemed to invalidate their conclusions. These criticisms were based, at least in part, on questionable handling or interpretation of the time-series properties of the data. In recent years, however, economists have devised new tests of long-run monetary neutrality, as well as related neutrality-type propositions. A fair amount of literature has been written on the subject, and the purpose of this paper is to review this literature.[1]

The next section provides more detail concerning the background behind the current empirical tests of neutrality propositions. In the following sections, some of the recent research using the newer set of tests is reviewed, and a few related papers are discussed along with the results authors have found using somewhat different methodologies. The final section offers some comments about directions for future research.

SOME BACKGROUND

What is Long-Run Neutrality?

In discussing long-run monetary neutrality, economists typically refer to a specific, hypothetical experiment that normally is not observed directly in actual economies. The experiment is a one-time, permanent, unexpected change in the level of the money stock. If, for instance, the money stock was $5 billion one day, and had been $5 billion for a long time, then what would the effect be of suddenly changing it to $6 billion and keeping it there for a long time? According to the quantity theory of money, prices should rise eventually in proportion to the increase in the money stock, and all real variables, perhaps after some transition

[1] Not all papers dealing with neutrality issues—an enormous amount of literature—can be surveyed here. Instead, attention is restricted to those that use the newer techniques discussed later.

time, would return to their original values and stay there until some further disturbance comes along. This is long-run monetary neutrality.

In the hypothetical experiment, it is important that the new level of the money stock be maintained for some, possibly long, period of time, to allow the transition effects to vanish. Theoretically, the change in the money stock has to be "permanent." In the world's economies, we observe a high degree of persistence in many macroeconomic variables, but it is generally difficult to tell the difference between "highly persistent" and "permanent." In the empirical work surveyed below we will see the use of many tests—*unit root diagnostic tests*—intended to categorize macroeconomic variables into those that have been subject to permanent shocks and those that have not. It is important to bear in mind, however, that these tests may not accurately distinguish between the two cases—statistically speaking, the tests have limited power. The tests are used because they offer the best available method for making the distinction between highly persistent and permanent changes, but they are far from perfect.

In the hypothetical experiment, it is also important that the change be unexpected, because if the economy's participants knew that the money stock was going to increase, and therefore, that prices were about to increase, they might start changing their present behavior. For example, they might buy consumption goods today, before the price increase takes effect. Prices then might begin to rise *in advance* of the money stock change. This complicates the story, and hence, we will think in terms of unanticipated changes in the money stock level. In the discussion below, this will be approximated by the notion of a "permanent shock" to the money supply.

In the world of monetary theory, nearly all models based on standard economic assumptions embody some form of monetary neutrality.[2] Most likely this is because monetary theorists generally think long-run monetary neutrality is sensible, and, therefore, they build it into their models.

Empirical tests that convincingly documented departures from long-run monetary neutrality therefore would be quite surprising (or quite suspect!) to monetary economists.

There is a second hypothetical experiment, related to the first, that more closely resembles the types of monetary policy actions we see in actual economies. This experiment says that the government initially maintains a certain *growth rate* for the money stock for a long period of time. At some date, that growth rate is adjusted unexpectedly to some new rate, say, from 3 percent to 5 percent on an annual basis, and is kept there for another long period of time. What effect should this have on important real variables like the capital-labor ratio, real output, real consumption expenditures, and real interest rates? If the answer is that after a long period of time, nothing would happen to the real variables, we have what is referred commonly to as long-run monetary *superneutrality*. Here again, one might expect an important transition period (commonly known as "the short run") when the economy is adjusting to the new rate of monetary growth. Quite a lot could happen to real variables during this adjustment period. But the neutrality and superneutrality propositions discussed in this paper mainly concern long-run, limiting effects.

Perhaps surprisingly, there are many plausible analyses that suggest that departures from long-run monetary superneutrality might be consistent with standard economic theory. It is, in fact, relatively easy to produce such theories. Moreover, these departures could go either way; that is, a permanently higher rate of monetary growth might eventually either raise or lower the level of economic activity, or change other important real variables in either a positive or negative direction.[3] Accordingly, whereas long-run neutrality is taken almost as an axiom of monetary economics, long-run superneutrality is far more circumspect. An empirical test that convincingly showed departures from long-run superneutrality would not be too surprising, since this result

2 The phrase "standard economic assumptions" means maintaining assumptions that markets clear at all times and that all agents behave rationally.

3 For a description of these departures, see the survey by Orphanides and Solow (1990).

is consistent with a number of existing economic theories.[4]

It is important to note that whether the level of real output rises or falls, or whether other real variables change in a particular direction in response to a permanent increase in the money growth rate, does not have any particular connotations for social welfare. In many theories, inflation distorts a Pareto optimal equilibrium, so that as a long-run proposition the population in the economy generally prefers lower rates of money growth accompanied by lower rates of inflation. Different theories make different predictions in this regard, however, and to sort these out one would have to consider various theories and their underlying assumptions in some detail. Since this would take us too far astray, social welfare will not be addressed in this survey.

There is another side to the superneutrality question. Fischer (1996) suggests that the reason the central banks of the world's industrialized economies have avidly pursued long-run price stability is because in the long run, inflation has distortionary effects that adversely impact a real variable, or a group of real variables, that people care about. If monetary growth causes inflation, and inflation has distortionary effects, then long-run monetary superneutrality should not hold in the data. On the contrary, a permanent shock to the rate of monetary growth should have some long-run effect on the real economy; why else should we worry about it? Care needs to be taken, however, in defining which variables are supposed to be affected and which are not—this is an area of some confusion in the literature.[5] In the current paper we will try to avoid this problem through use of the language "superneutrality with respect to variable *x*."

The above discussion has referred to changes in real variables, meaning changes in the *level* of the variable, especially so with respect to the level of real output. Of course, real output in industrialized economies generally grows over time. A shift in the level would be a one-time movement, say from 100 to 90, whereupon the variable would resume growing at its previous rate.

Thus permanent effects on the level of a variable need not imply permanent effects on the growth rate of that variable. Consequently, a natural question to ask is whether permanent changes in the monetary growth rate affect a country's *rate* of economic growth; that is, is money superneutral with respect to economic growth? Many researchers in recent years have in fact investigated questions of this type (mostly with methodology outside the focus of this survey). There is much less theory concerning this issue, but some of the results I discuss later will have some bearing on this topic.
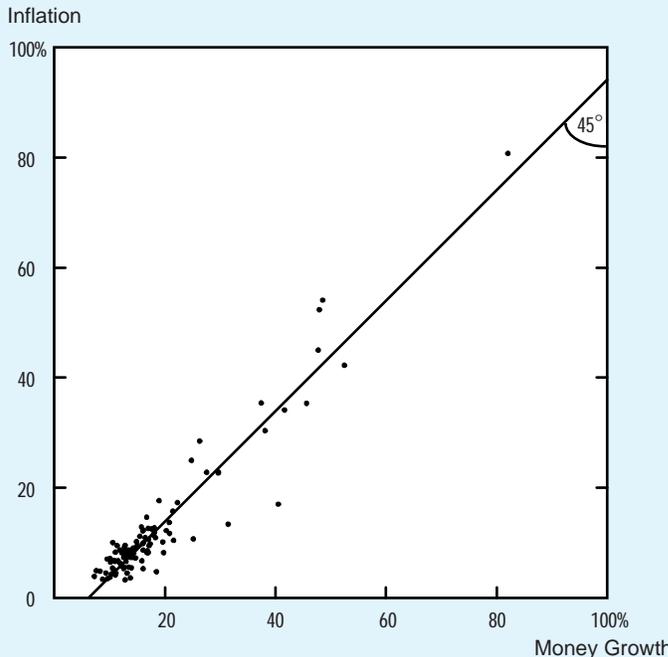
Prima Facie Evidence. In his Nobel Lecture, Lucas (1996) addresses the topic of monetary neutrality, both in the short run and the long run, and discusses theoretical developments that might reconcile the perceived short-run effects of an increase in the money supply with long-run monetary neutrality. Lucas mentions several pieces of evidence as constituting the main reasons that he would like a satisfactory theory of the real effects of monetary policy to address. Among these, he cites Friedman and Schwartz (1963) who argue that all major recessions in the United States between 1867 and 1960 were preceded by substantial contractions in the money supply, suggesting that monetary policy mistakes were a primary contributor to business cycle downturns during this period. Lucas states that severe monetary contraction seemed to play an important role especially during the Great Depression of 1929-33. But he also cites work by Sargent (1986) who argues that huge reductions in the rate of monetary expansion —reductions much larger than anything experienced in the post-Civil War United States—did not lead to any unusually large reduction in real output in the hyperinflationary post-World War I European economies. These reductions were carried out in conjunction with monetary reform. The hyperinflations ended abruptly when credible reform was announced. But these citations are subsidiary to Lucas's (1996, p. 668) main contention, that there is clear

[4] The situation described is summarized by Canova (1994, p. 123), who states, "... there are very few available models which display superneutrality, while most existing models, both in the neoclassical and neoKeynesian tradition, possess neutrality of money...."

[5] See Marty (1994) for a discussion.

Money Growth and Inflation

Inflation

Postwar average rates of money growth versus average inflation rates in 110 countries. Observations near the 45 degree line, which is not fitted to the data, are consistent with the quantity theory. This figure is from McCandless and Weber (1995).

evidence—even "decisive confirmation"—that long-run monetary neutrality holds.

Figure 1 shows the evidence that Lucas (1996) cites. This figure, from McCandless and Weber (1995), plots the average rates of monetary growth against average rates of inflation for 110 countries. The averages are taken over 30 years, 1960-90. Monetary growth is measured as the annual growth rate of M2 for a country, and inflation is measured as the annual rate of increase in the consumer price index for a country. The 45-degree line is not fit to the data, but instead represents a theoretical presumption based on the quantity theory, that the rate of inflation should correspond to the rate of money growth (adjusted for the real output growth rate in a particular economy). McCandless and Weber report a simple correlation of .95 between money growth and inflation based on this data. Lucas (1996, p. 666) asks "... how many specific economic

theories can claim empirical success at the level exhibited in figure 1? ... The kind of monetary neutrality shown in this figure needs to be a central feature of any monetary or macroeconomic theory that claims empirical seriousness."

While Figure 1 is impressive, one should be careful to note that these results are different from the stories about long-run monetary neutrality and superneutrality outlined above. Evidently, the average rate of money growth is correlated highly with the average rate of inflation in a country. But the story about long-run monetary neutrality is about a permanent, unexpected change in the level of the money stock in a single country, and the ultimate impact of such a change. And, the story about superneutrality concerns the long-run effect of a permanent, unexpected change in the rate of monetary expansion. Taking averages over long periods of time, while informative at some level, masks the information about such events, to the extent they might have occurred in the data. To study long-run neutrality more directly, the time-series evidence on inflation and monetary growth for individual countries needs to be considered. Can we isolate permanent, or at least highly persistent, changes in the money stock (or the monetary growth rate), which are then correlated with persistent changes in the price level (or the rate of inflation) and simultaneously are uncorrelated with permanent movements in important real variables? That is the challenge of testing monetary neutrality propositions.

Time-Series Evidence. Some tests of long-run monetary neutrality during the 1960s simply regressed the level of real output on a distributed lag of observations on the money stock. In reaction to this practice, Sargent (1971) and Lucas (1972) argued that such evidence was circumspect for two related reasons. One is that Sargent and Lucas built simple and plausible reduced-form models of the macroeconomy in which long-run monetary neutrality held by construction, but which also would produce data such that, if the standard

Review
NOVEMBER/DECEMBER 1999

practice was applied, the researcher would conclude that long-run monetary neutrality failed. Thus, any evidence based on the (then) standard methodology was difficult to interpret.

The second reason—the one that is at the heart of the methods used in the recent research—was that the story of monetary neutrality involves permanent changes in the level of the money stock, and that one cannot effectively test such a theory without evidence that the actual money stock has been subject to a permanent change. The idea of permanent changes in economic variables is statistically modeled as a unit root in the autoregressive representation of a time series; a time series with a unit root has quite different properties from a stationary series.[6] During the early 1970s when Lucas and Sargent first wrote about this topic, the implications of unit roots in economic time series were only beginning to be appreciated. Later, in an influential paper, Nelson and Plosser (1982) argued that many U.S. macroeconomic time series were best characterized by a unit root in their univariate, autoregressive representations. Their results brought the issue of how to handle these nonstationary time series to the fore in macroeconometrics, and led to econometric methodologies that respected the potential for nonstationarity in important macroeconomic variables.

The nonstationarity in economic variables was viewed as something of a headache for much of macroeconometrics. But in a remarkable turn of events, it actually was a boon to testing neutrality propositions. As Lucas and Sargent had argued, one needs permanent changes in the money stock as part of the historical record to test the proposition of long-run neutrality in a time-series setting. But permanent shocks are exactly what macroeconomic time series provide.

This was exactly the line pursued by Fisher (1988) and Fisher and Seater (1989, 1993), and also in a series of papers by King and Watson (1992, 1994, 1997). These authors provided new tests of long-run neutrality propositions that respected

the Lucas-Sargent critique and required little macroeconomic structure.

TESTING NEUTRALITY PROPOSITIONS

Recent Tests Based (Mostly) on U.S. Data

Fisher and Seater (1993) work in terms of bivariate systems, with a measure of money as one of the variables. Adopting their notation, let m be the natural logarithm of the nominal money stock M. Let y be a second variable, expressed in either real or nominal terms, which is the logarithm of a variable like the price level or real output, and where the variable itself is Y.[7] Denote the order of integration of a variable by $\langle x \rangle$, so that if x is integrated of order 1, we write $\langle x \rangle = 1$. Sometimes we also will use the phrase "x is $I(1)$" to describe the order of integration. Denote a difference operator by Δ, so that Δy indicates the approximate growth rate of the variable Y. Fisher and Seater study the following system

$$(1) \qquad a(L)\Delta^{\langle m \rangle}m_t = b(L)\Delta^{\langle y \rangle}y_t + u_t$$

$$(2) \qquad d(L)\Delta^{\langle y \rangle}y_t = c(L)\Delta^{\langle m \rangle}m_t + w_t$$

where $a(L)$, $b(L)$, $c(L)$ and $d(L)$ are lag polynomials, and $a_0 = d_0 = 1$ and b_0 and c_0 are unrestricted. The error vector $(u_t, w_t)'$ is *iid* with zero mean and covariance Σ. Now let $x_t \equiv \Delta^i m_t$ and $z_t \equiv \Delta^j y_t$, with $i, j = 0$ or 1. Fisher and Seater define a certain *long-run derivative* (LRD) that is central to their findings. The LRD is a change in z with respect to a permanent change in x, given by

$$(3) \qquad LRD_{z,x} \equiv \lim_{k \to \infty} \frac{\partial z_{t+k}/\partial u_t}{\partial x_{t+k}/\partial u_t},$$

provided $\lim_{k \to \infty} \partial x_{t+k}/\partial \mu_t \neq 0,$ otherwise the LRD is undefined. Fisher and Seater then define long-run neutrality and long-run

[6] One could use other methods to statistically model a permanent shift in the level or growth rate of a monetary variable. One could, for instance, posit a discrete shift in the mean of the variable at a given date, T, and one could then check to see how other variables responded to such a permanent movement. Nothing here is ruling out such an approach, but the literature surveyed in this paper focuses on unit-root characterizations of variables of interest as measures of whether these series have permanent components or not.

[7] To simplify the discussion in this section, interest rates are left out here, even though they are included in Fisher and Seater's (1993) framework.

FEDERAL RESERVE BANK OF ST. LOUIS

61

superneutrality in this framework, and for each, discuss four cases that depend on the order of integration of the variables.[8]

First of all, money is long-run neutral with respect to y if $LRD_{y,m}=1$ when y is a nominal variable, or if $LRD_{y,m}=0$ when y is a real variable. The four cases are: 1) $\langle m \rangle < 1$. Here the LRD is not defined because there have been no permanent shocks to the level of the money stock, and the data are uninformative concerning long-run monetary neutrality. 2) $\langle m \rangle \geq \langle y \rangle + 1 \geq 1$. Here the LRD is zero because while there have been permanent shocks to the level of the money stock, there have been none to y. If y is a nominal variable, long-run neutrality is violated, otherwise it holds. 3) $\langle m \rangle = \langle y \rangle \geq 1$. This case admits tests of long-run neutrality, in an effort to find out if the permanent shocks to the level of the money stock are correlated with the permanent shocks to the variable y. 4) $\langle m \rangle = \langle y \rangle -1 \geq 1$. This case is more complicated. A necessary condition for long-run neutrality is that the permanent shock to money does not change the growth rate of y.

Secondly, money is long-run superneutral with respect to y if $LRD_{y,\Delta m}=0$. The cases are 1) $\langle \Delta m \rangle < 1$. Here the LRD is not defined because there have been no permanent shocks to the growth rate of the money stock, and the data are uninformative concerning long-run monetary superneutrality. 2) $\langle \Delta m \rangle \geq \langle y \rangle + 1 \geq 1$. The LRD is zero because while there have been permanent shocks to the growth rate of the money stock, there have been none to y. Long-run superneutrality holds. 3) $\langle \Delta m \rangle = \langle y \rangle \geq 1$. This case admits tests of long-run superneutrality, in an effort to find out if the permanent shocks to the level of the money stock are correlated with the permanent shocks to the variable y. 4) $\langle \Delta m \rangle = \langle y \rangle -1 \geq 1$. Here $LRD_{\Delta y, \Delta m}=0$ is testable; that is, one can determine whether a permanent change in the growth rate of money is associated with a permanent change in the growth rate of y.

Fisher and Seater (1993) use these results to analyze previous research efforts testing long-run neutrality propositions,

efforts that, because of the time they were written, did not take such explicit account of the time-series properties of the data. They interpret the evidence in Anderson and Karnovsky (1972), Kormendi and Meguire (1984), Lucas (1980), and Geweke (1986) mostly as consistent with long-run neutrality and not very informative about long-run superneutrality. They also provide some evidence of their own. They use the Friedman and Schwartz (1982) data on money, prices, nominal income, and real income from 1867 to 1975 in the United States. All variables are viewed as $I(1)$, making tests of long-run neutrality possible. With respect to nominal income and prices, long-run monetary neutrality holds in this data, but with respect to real output, long-run monetary neutrality fails.

As mentioned earlier, evidence of the failure of long-run monetary neutrality is either surprising or suspect among monetary theorists; the Fisher and Seater finding was no exception. In a note, Boschen and Otrok (1994) re-estimate the systems studied by Fisher and Seater, again using the Friedman and Schwartz (1982) data, but now updating the time series through 1992. They split the data set into two subsamples, 1869-1929 and 1940-92. They find that long-run neutrality holds in both of the subsamples using the Fisher and Seater methodology. They conclude that there may have been something special about the financial disruption during the Great Depression era that causes the test to fail when that period is included.

Haug and Lucas (1997) comment further on these findings. They reason that, since Canada did not experience bank failures during the Great Depression, the evidence on long-run neutrality using Canadian data might provide further evidence that something unusual happened in the United States during this period. Their data set includes real national income and the M2 money supply from 1914-94. They argue that pre-1914 data is inappropriate for this purpose because changes in the money supply were not exogenous in Canada at that time. They

conclude, based on augmented Dickey-Fuller (ADF) tests, that both time series are $I(1)$. And, according to the Fisher and Seater (1993) methodology, long-run monetary neutrality with respect to real output cannot be rejected using the entire Canadian sample period. Haug and Lucas interpret this finding as independent support for the arguments of Boschen and Otrok (1994).

Olekalns (1996) similarly explored an alternative data set, 94 years of annual Australian data. The downturn of the 1930s was less severe in Australia. Olekalns uses the Fisher and Seater methodology, and two measures of money, M1 and M3, along with real gross domestic product. All variables are reasonably described as $I(1)$ according to ADF tests. Olekalns finds that long-run monetary neutrality cannot be rejected using the narrower money measure. However, using the broader money measure, long-run neutrality can be rejected for this data set, and the rejection carries even when dummy variables are used to control for the Depression period as well as the WWII period. Olekalns concludes that results can be sensitive to the measure of money used.

A recent paper by Coe and Nason (1999) also contributes to this literature. They use the Fisher and Seater (1993) test for long-run neutrality, and they employ the same U.S. data as Fisher and Seater, except that they update the data through 1997. When Coe and Nason use a broad measure of the money stock (as Fisher and Seater did), they replicate the Fisher and Seater rejection of long-run monetary neutrality with respect to real output. But when they replace the broad money measure with the monetary base, they can no longer reject long-run neutrality. They also consider about a century of data from the United Kingdom, and fail to reject long-run neutrality using either broad or narrow measures of money. Coe and Nason conclude that the Fisher and Seater rejection of long-run neutrality is not robust to a change in either the measure of money or the country of study.[9]

An important work in this literature, King and Watson (1997) also use bivariate systems, and they also take careful note of the order of integration of the variables involved when devising tests of neutrality propositions. They study a "final form" model

$$(4) \quad \Delta y_t = \mu_y + \theta_{y\eta}(L)\varepsilon_t^\eta + \theta_{ym}(L)\varepsilon_t^m$$

$$(5) \quad \Delta m_t = \mu_m + \theta_{m\eta}(L)\varepsilon_t^\eta + \theta_{mm}(L)\varepsilon_t^m$$

where y is the logarithm of real output, the θ coefficients are lag polynomials, ε_t^m is a serially independent, zero mean shock to money, and ε_t^η is a vector of nonmonetary shocks that affect output. King and Watson show that

$$(6) \qquad \gamma_{ym} = \frac{\theta_{ym}(1)}{\theta_{mm}(1)}$$

is the long-run elasticity of real output with respect to permanent shocks to the money stock. Thus, long-run neutrality here is analogous to the Fisher and Seater definition: $\gamma_{ym} = 0$. Again, long-run neutrality can be investigated only if there have been permanent shocks to the money stock.

Importantly, King and Watson (1997) emphasize identification issues. They analyze long-run neutrality propositions across a range of possible identifications of their bivariate system, in an effort to understand the robustness of various conclusions to differing assumptions. They rewrite the equations (4) and (5) as

$$(7) \qquad \Delta y_t = \lambda_{ym}\Delta m_t + \sum_{j=1}^{p} \alpha_{j,yy}\Delta y_{t-j}$$
$$+ \sum_{j=1}^{p} \alpha_{j,ym}\Delta m_{t-j} + \varepsilon_t^\eta$$

$$(8) \qquad \Delta m_t = \lambda_{my}\Delta y_t + \sum_{j=1}^{p} \alpha_{j,my}\Delta y_{t-j}$$
$$+ \sum_{j=1}^{p} \alpha_{j,mm}\Delta m_{t-j} + \varepsilon_t^m$$
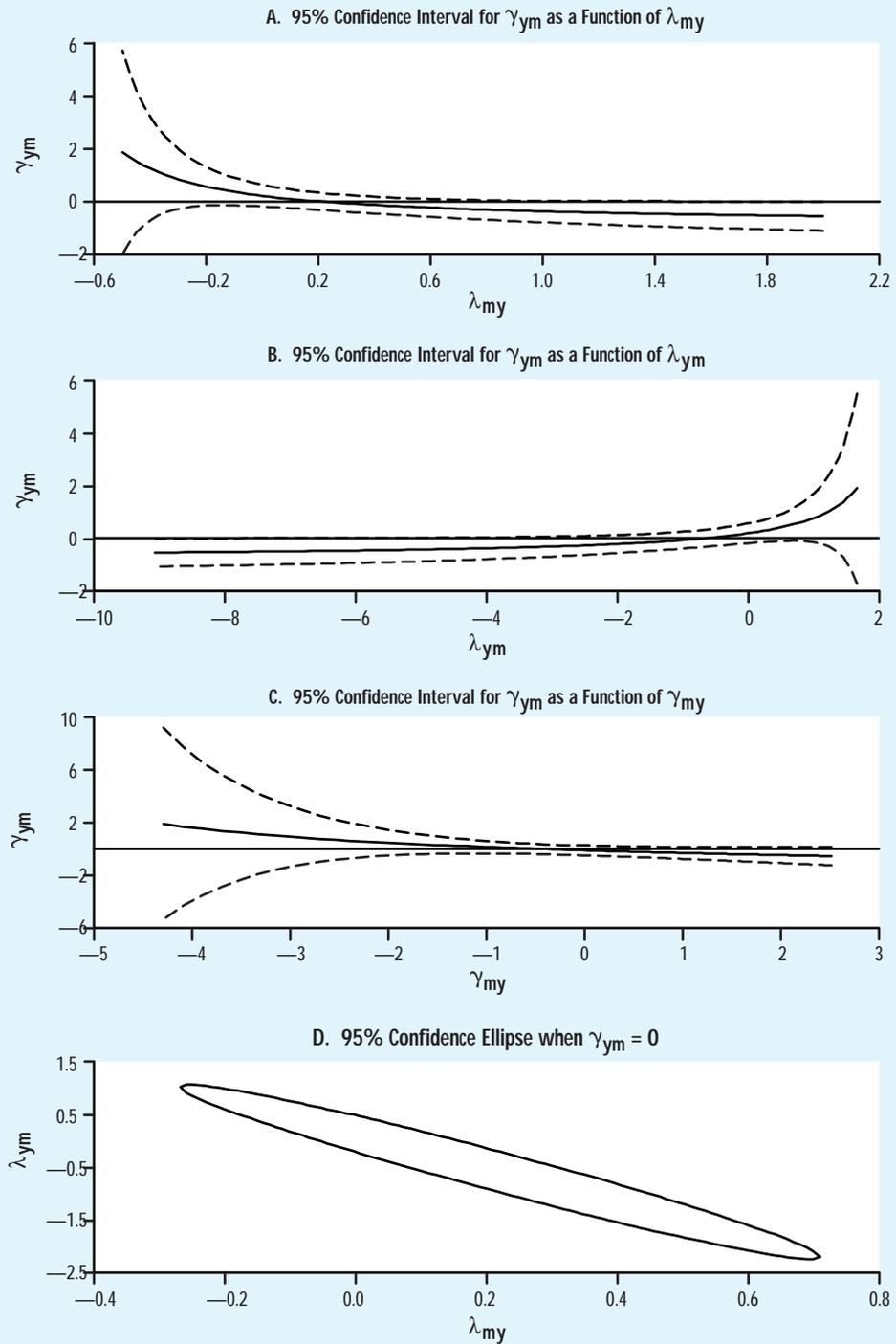
and they assume that
$$\text{cov}(\varepsilon_t^m, \varepsilon_t^\eta) = 0.$$

They note that there are several plausible ways to complete their identification of the system. One could assume $\lambda_{ym} = 0$, or that

[9] Coe and Nason also study the asymptotic power properties of the Fisher and Seater long-horizon regression test, and they conclude that the test has low power against alternative hypotheses of monetary non-neutrality. For small samples, Monte Carlo experiments reveal poor size-adjusted power especially at longer horizons. Coe and Nason conclude based on this portion of their analysis that the Fisher and Seater approach to testing long-run monetary neutrality may not be informative.

Figure 2

Inflation and Unemployment

A. 95% Confidence Interval for γ_{ym} as a Function of λ_{my}

B. 95% Confidence Interval for γ_{ym} as a Function of λ_{ym}

C. 95% Confidence Interval for γ_{ym} as a Function of γ_{my}

D. 95% Confidence Ellipse when $\gamma_{ym} = 0$

The evidence on long-run monetary neutrality according to King and Watson (1997). The panels show how the point estimate of γ_{ym} changes under the differing identifying restrictions, with the dotted lines indicating 95% confidence intervals. The bottom panel displays a 95% confidence ellipse for λ_{ym} and λ_{my} under the identifying restriction $\gamma_{ym} = 0$.

$\lambda_{my} = 0$; this means that the impact elasticity of one variable on the other is zero. Alternatively, one could simply assume long-run monetary neutrality by imposing $\gamma_{ym} = 0$. And finally, one could assume that $\gamma_{ym} = 1$ where γ_{my} is the long-run elasticity of money with respect to a permanent shock to real output. King and Watson (1997, p.77) argue that this last assumption is consistent with stable prices in an economy with constant velocity.

Because King and Watson wish to investigate the robustness of neutrality results to alternative identifying assumptions, they use all four of these possibilities. Furthermore, they allow a wide variety of values for each elasticity, not just the zeroes and ones of the previous paragraph. Thus, the identifying assumptions are that either one of the impact elasticities is known to be a certain value, or that one of the long-run elasticities is known to be a certain value. They then turn to estimation and report results considering a number of neutrality propositions. The quarterly data are for the United States and cover the sample period from 1949:1 to 1990:4; except for systems with unemployment, in which case the sample period is from 1950:1 to 1990:4. The lag length p is set to six, although they experiment with values of four and eight at some points. Based on unit-root diagnostic tests, King and Watson conclude that all the series involved can reasonably be viewed as $I(1)$, so that tests of neutrality propositions can be executed.

King and Watson (1997) first investigate the long-run neutrality of money in the context of a bivariate system using real output and money (M2). They begin by estimating a value for γ_{ym} using the identifying assumption that λ_{my} is known. They find that a 95-percent confidence interval for γ_{ym} contains zero (and so supports long-run monetary neutrality) so long as $\lambda_{my} > 1.4$. If we interpret λ_{my} the parameter as a short-run elasticity of money demand, a reasonable range is $.1 \le \lambda_{my} \le .6$, so that the evidence is consistent with long-run neutrality. King and Watson complete similar calculations for identifying assumptions

involving λ_{ym} and γ_{my}. They also estimate 95-percent confidence intervals for λ_{my}, λ_{ym}, and γ_{my} using the identifying assumption that long-run neutrality holds, $\gamma_{ym} = 0$, in order to see if the confidence intervals produced contain the most reasonable values for these parameters. All of this evidence comes down in favor of long-run neutrality, which is consistent with the findings of Fisher and Seater (1993) and Boschen and Otrok (1994), because the sample period here covers the postwar United States.[10] This evidence is summarized in Figure 2.

The superneutrality of money is investigated using a bivariate system with money growth (replacing the level of the money stock) and real output, and the hypothesis is that the long-run elasticity of the level of output to a permanent change in the growth rate of money, $\gamma_{y, \Delta m}$, is zero. The evidence on this question turns out to be mixed, in that for some identification schemes that King and Watson consider reasonable, the hypothesis that $\gamma_{y, \Delta m} = 0$ can be rejected at the 5-percent level. Moreover, the effect can go either way: a permanently higher rate of money growth tending to permanently increase the level of real output, or to decrease the level of real output. For instance, if the identifying assumption is $\lambda_{\Delta m, y} = 0$ (which King and Watson again interpret as a short-run money demand elasticity), then the estimated value of $\gamma_{y, \Delta m}$ is positive and statistically significant, while if the identifying assumption is that $\lambda_{\Delta m, y} = .6$, then the estimated value of $\gamma_{y, \Delta m}$ is negative and statistically significant. As mentioned earlier, theories exist that are consistent with both possibilities.

King and Watson go on to investigate a neutrality proposition associated with the early 20th century economist Irving Fisher. The proposition is that nominal interest rates move one-for-one with permanent changes in inflation, leaving the real interest rate unaffected. Using a system with consumer price index inflation, π, playing the role of the money variable, and the nominal interest rate on three-month Treasury bills, R, playing the

[10] Jefferson (1997) also investigates monetary neutrality questions using the King and Watson (1997) methodology, except that he considers measures of both inside money (defined as nominal checkable deposits, or M2 less currency) and outside money (defined as the monetary base). He uses nearly a century of data from the United States and finds some departures (under some identifying restrictions) from long-run neutrality when inside money is used.

role of the output variable, King and Watson (1997) investigate the hypothesis that $\gamma_{R\pi} = 1$; that is, the long-run elasticity of the nominal interest rate with respect to a permanent inflation shock is one. The evidence here again turns out to depend on the identification scheme. When statistically significant differences from the standard Fisher relation occur, they occur in a negative direction, with nominal interest rates rising less than one-for-one with perma-nent shocks to inflation. In other words, real interest rates are lowered permanently by permanent, positive shocks to the inflation rate. King and Watson find that identifying the model by assuming $\gamma_{R\pi} = 1$, and then estimating 95-percent confidence intervals for the remaining parameters, leads to the conclusion that there are reasonable configurations of parameters that are consistent with the Fisher hypothesis. Nevertheless, the main conclusion is that nominal interest rates do not adjust fully to permanent inflation shocks, and this conclusion holds across a large set of identification schemes.

Finally, King and Watson turn to estimating the slope of a long-run Phillips curve; that is, the long-run response of unemployment to permanent shocks in the inflation rate. This particular test is discussed in more detail in another paper, King and Watson (1994). The bivariate system now includes the CPI inflation rate in the role of the nominal variable, and the unemployment rate in the role of the real variable. The hypothesis is that the long-run Phillips curve is vertical, which means $\gamma_{u\pi} = 0$ in this framework. King and Watson report that a statistically significant (negative) slope for the long-run Phillips curve can be obtained only if the identifying assumption is that $\gamma_{\pi u} > 2.3$, or alternatively that $\gamma_{u\pi} < -0.7$. In particular, if either of these two impact elasticities are zero, then a vertical long-run Phillips curve cannot be rejected. King and Watson conclude that a reasonable estimate of the long-run Phillips curve based on this data is either vertical or at least "very steep."

While King and Watson's strong suit is that they can investigate the robustness of

results on neutrality for a wide variety of identification schemes, they do so only for bivariate systems, and they note the possibilities for exposure to omitted variable bias. One of the few multivariate studies available using techniques related to those of Fisher and Seater (1993) for testing neutrality is by Boschen and Mills (1995). They use the notion of permanent shocks to the level of the money stock to test long-run monetary neutrality in the U.S. data. They use a relatively high dimensional system, and they organize their research around the idea that, if long-run neutrality does not hold, there would be a nonstationary component of real output that is determined by long-term movements in the money stock. They study a vector error correction model (VECM) representation

$$(9) \quad \Delta X_t = \mu + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Pi X_{t-k} + \varepsilon_t,$$

where $X' \equiv (y, m, \upsilon)'$, and y is aggregate output, υ is a vector of real shocks, m is a vector of monetary shocks, and ε is normally distributed, *iid*, and has mean zero. Interest centers on the long-run impact coefficient matrix Π that describes the long-run relationships in the model. For each cointegrating relationship, this matrix will have a nonzero row. If there is a cointegrating relationship between the monetary variables and output, then these variables contribute to the trend shifts in output, and long-run neutrality is violated.[11]

Boschen and Mills (1995) use quarterly U.S. data from 1951:4 to 1990:4. They include variables describing productivity, real oil prices, weighted foreign real GDP of major U.S. trading partners, real government purchases, taxes, labor supply, the M1 money stock, the M2 money stock, and the nominal three-month Treasury bill rate. They use augmented Dickey-Fuller tests as diagnostics for the presence of nonstationarity in these data; they found (sometimes weak) evidence of a unit root in all the series. They test for cointegrating relationships among the blocks of real and nominal variables, and then between the nominal and real variables, as a means of testing for long-run monetary neutrality.

Based on these tests, Boschen and Mills conclude that long-run monetary neutrality holds during the postwar period in the U.S. This result confirms the findings that Fisher and Seater (1993), Boschen and Otrok (1994), and King and Watson (1997) reported regarding long-run monetary neutrality in the postwar U.S. data. This result also provides the best available evidence that omitted variable bias did not contaminate the previous results on this question.

Recent Tests Using International Data

So far, we have results that conform to the suggestions of King and Watson and Fisher and Seater only for U.S. data—certainly a natural place to start but not the true extent of the available evidence. As a first effort at generalization, Weber (1994) explicitly set out to apply the King and Watson testing procedures to G7 economies: Canada, France, Germany, Italy, Japan, the United Kingdom, and the United States. The data is quarterly, from the postwar era, but the particular years vary across countries.

Weber begins with a battery of unit-root diagnostics—using a much more elaborate procedure than the papers discussed so far—in an effort to make careful statements about the evidence for the presence of a unit root in the time series. For each country, he uses several different measures of the money stock, in part to confront the question of whether the results are sensitive to how money is defined. The combination of several diagnostic tests and many different time series produces a plethora of results that are not all the same. As a general rule, however, narrower monetary aggregates tended to be $I(1)$, while broader aggregates tended to be $I(2)$. Strictly speaking, according to the methodology outlined above, if money is $I(2)$ then superneutrality can be tested, whereas neutrality cannot. In response to this situation, Weber takes two approaches: In some cases, he performs neutrality tests anyway and warns the reader to interpret the results with caution, while in other cases he uses the $I(1)$ aggregates to test for neutrality and the $I(2)$ aggregates to test for superneutrality. The remaining series on output, inflation, interest rates, and unemployment rates, for the most part, can be reasonably interpreted as $I(1)$.
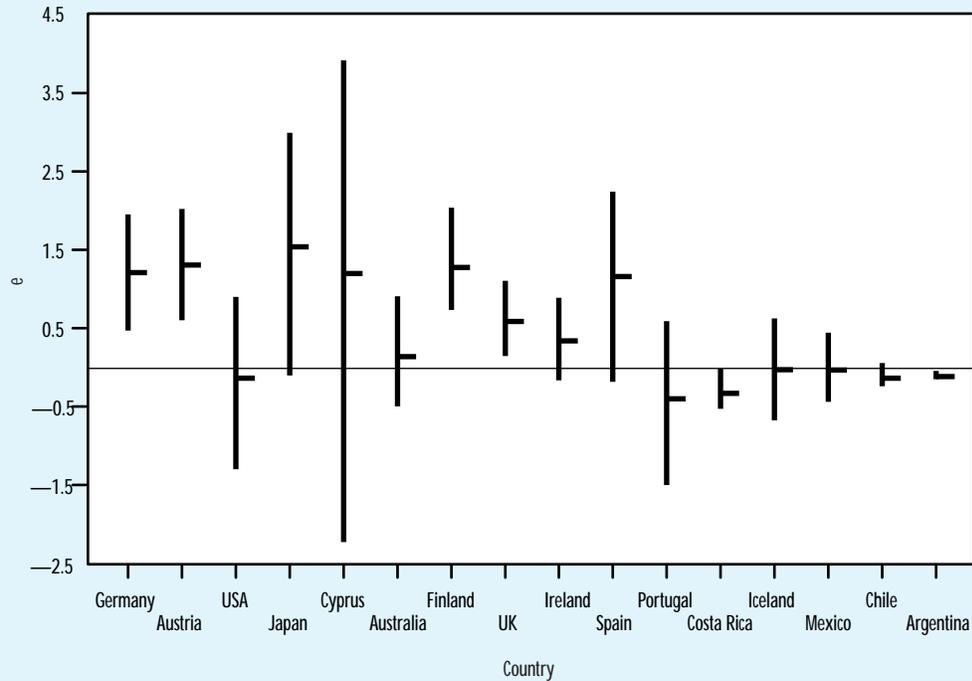
Weber then turns to tests of long-run monetary neutrality in these countries, using the same wide variety of possible identifying assumptions that King and Watson used. His general finding is that for broader monetary aggregates, such as M2 or M3, a wide variety of identifying restrictions are consistent with (fail to reject) long-run monetary neutrality in the G7 economies during the postwar era. For narrower measures of money, the range of identifying restrictions consistent with long-run monetary neutrality is much smaller. Confidence ellipses for λ_{ym} and λ_{my} under the identifying assumption that money is long-run neutral, $\gamma_{ym}=0$, include the plausible region of the space where $\lambda_{ym} < 0$, (the short-run impact of money on output is positive) and $\lambda_{my} > 0$ (money reacts countercyclically to output in the short run). This is true across the G7 economies for both narrow and broad measures of money.

Superneutrality is examined using a bivariate VAR in differenced money growth and differenced real output. In general, long-run superneutrality with respect to the level of real output is rejected for a wide variety of identifying restrictions across the G7 economies.

Considering the question of whether the long-run Phillips curve is vertical, Weber proceeds using the changes in inflation and unemployment in his bivariate VAR. For six of the seven G7 economies, the hypothesis that $\gamma_{u\pi}=0$ cannot be rejected except in cases of rather extreme identifying assumptions. The exception is Italy, where this hypothesis can be rejected readily. Weber also considers a "reverse" hypothesis, with causality running from unemployment to inflation. In this case the hypothesis $\gamma_{\pi u}=0$ can be rejected easily across the non-Italian economies. For Italy this hypothesis is rejected only for extreme identification schemes. Weber

Figure 3

**Long-Run Response of the Level of Output
To a Permanent Increase in Inflation**

Countries ordered from lowest to highest average inflation in sample.
Horizontal lines represent point estimates. Vertical lines represent 90 percent confidence bounds.

The point estimate of the long-run response of the level of real output to a permanent inflation shock is generally positive for the low inflation countries but zero or negative for high inflation countries. This figure is reproduced from Bullard and Keating (1995) and is reprinted with permission by the *Journal of Monetary Economics.*

speculates that wage indexation in Italy during much of this period accounts for the differences between Italy and the other industrialized economies.

Finally, Weber goes on to test the Fisher relation for the G7 economies using a bivariate VAR in differenced inflation and differenced nominal interest rates. He finds that for Germany, a Fisher relation can be rejected for a wide variety of identifying restrictions, although some important benchmark restrictions do not lead to rejection. For the United States, Weber confirms the findings of King and Watson (1997) that nominal interest rates do not adjust one-for-one with permanent shocks to inflation. Even stronger evidence in this direction is found for the United Kingdom. But for Japan, Canada, Italy, and France,

the evidence is much more favorable for a Fisher relation, $\gamma_{r\pi}=1$, to hold. The general finding in the previous literature (see for instance, Lothian, 1985, for 20 OECD countries) is that nominal interest rates adjust less than one-for-one with inflation.

While Weber considered G7 economies, Bullard and Keating (1995) consider virtually all of the countries in the world where enough data existed to formulate tests of long-run neutrality propositions in the spirit of Fisher and Seater (1993) and King and Watson (1997). In working with a large number of countries, data availability and quality impinge significantly on the analysis. Accordingly, Bullard and Keating restrict attention to countries that produce at least moderately high quality data (according to published

assessments) and have at least 25 years of consecutive annual observations (quarterly data often is not available). This leaves them with 58 countries. Bullard and Keating focus their analysis on one particular version of a neutrality proposition: the effect of a permanent shock to inflation on the level of output. If money is long-run neutral—and the evidence reported in this survey suggests that this is a reasonable assumption—then this can be viewed as a test of monetary superneutrality with respect to the level of real output. Moreover, problems with the definition of money within and across countries are avoided.

Bullard and Keating also begin with a battery of unit-root diagnostic tests for the real output and GDP deflator time series they use. They divide countries into groups based on the results of these tests, according to whether a country can be characterized as having experienced permanent shocks to inflation or not, and similarly for the level of real output. Countries that experienced permanent shocks to inflation and also to the level of output are candidates for a test based on a bivariate VAR; there were 16 countries in this group, dubbed Group A. There were also nine Group B countries for which evidence of a unit root in inflation was found, but evidence of a unit root in output was lacking. A large number of countries, 31, showed no evidence of permanent shocks in the inflation series, and were put into Group C. The two remaining countries were special cases.

Bullard and Keating (1995) then ran a two-variable VAR for the Group A countries, in differenced inflation and differenced output. They committed to the long-run identifying restriction that money is long-run neutral, $\gamma_{\pi y}=0$ in the King and Watson (1997) notation, and did not attempt to search over alternative identifying restrictions. They used the techniques of Blanchard and Quah (1989) to decompose shocks into permanent and transitory components, and consequently they considered the impulse-response functions of reactions of the two variables to both permanent and transitory shocks.

The main results for the Group A countries are as follows: The long-run response of the level of output to a permanent inflation shock was positive and statistically significant for four countries, negative and statistically significant for one country, and not statistically different from zero for the remainder. The point estimate of this long-run response generally declined as the in-sample average inflation rate increased, as shown in Figure 3.

The Group B countries, which possess permanent inflation shocks but no permanent output shocks, provide *prima facie* evidence of superneutrality. The Group C countries are uninformative because they do not possess permanent inflation shocks. Altogether, the results appear to be consistent with superneutrality for most of the countries that are informative.

However, as Figure 3 indicates, and as is borne out by the associated impulse-response functions, low-inflation countries appear to react very differently to permanent inflation shocks than high-inflation countries.[12] In particular, for low-inflation countries the point estimate of the long-run response is generally positive, while for high-inflation countries it is zero or negative. This suggests that averaging results from low and high-inflation countries may be misleading.

Bullard and Keating also comment on the prospects for permanent inflation shocks to permanently alter *rates of growth* of real output in this sample. According to ADF and other diagnostic tests for unit roots, real output growth rates are stationary in nearly all countries that experienced permanent shocks to inflation. This is direct evidence for superneutrality with respect to output growth rates. This result does not seem like one that is likely to change with data sets or countries, since the stationarity of real output growth is likely to remain under most conceivable criteria.[13]

All of the Bullard and Keating data are for postwar economies. Serletis and Krause (1996) use the Backus and Kehoe (1992) data set, which includes more than 100 years of annual observations on real output, prices, and money for Australia, Canada,

[12] Other aspects of the impulse-response functions had natural interpretations according to conventional wisdom, and also were generally consistent across countries.

[13] This result conflicts with other evidence from the cross-country growth regression literature, such as Barro (1996).

Denmark, Germany, Italy, Japan, Norway, Sweden, the United Kingdom, and the United States.[14] They test for unit roots using the procedures of Zivot and Andrews (1992), and they conclude that money is reasonably described as $I(1)$ except in Germany and Japan where it is $I(0)$; these latter two countries are therefore uninformative on neutrality questions in this data set. Serletis and Krause (1996) find that output is $I(0)$ for Australia, Canada, Denmark, Italy, the United Kingdom, and the United States. These countries, therefore, provide direct evidence in favor of long-run neutrality with respect to output. Serletis and Krause use the Fisher and Seater (1993) regression test to produce estimates for the remaining money-price or money-output combinations. These results generally support a hypothesis of long-run monetary neutrality.

The same data set is used by Serletis and Koustas (1998), who apply the King and Watson methodology to study long-run neutrality and superneutrality issues over a range of plausible identifying restrictions. They use only the money and real output series, and apply a battery of tests to determine the integration properties of the data. Except for the money series for Italy, which is $I(2)$, they conclude that all series are $I(1)$ and hence provide a reasonable dataset with which to test long-run monetary neutrality (superneutrality for Italy).[15] The results state that it is generally difficult to reject long-run monetary neutrality in this dataset under plausible identifying restrictions. An exception is the United Kingdom, when the identifying restriction is that $0 \le \lambda_{my} \le .6$. Superneutrality of money with respect to real output in the Italian data can be rejected under plausible identifying restrictions.

The Serletis and Krause (1996) and Serletis and Koustas (1998) results may appear to impinge on the Fisher and Seater (1993) and Boschen and Otrok (1994) findings for the United States, namely that the results for long-run monetary neutrality in the United States over the last century depend critically on inclusion or removal of the Great Depression years from the sample.

Both Serletis and Krause (1996) and Serletis and Koustas (1998) fail to reject long-run neutrality even when this period is included (under a range of plausible identification schemes in the latter case). However, Serletis and Koustas (1998) in fact reject long-run neutrality under the Fisher and Seater (1993) identifying restriction ($\gamma_{my} = 0$), but they do not reject under other, possibly more plausible, identifying restrictions. Of course, differences in results could also be due to differences in the data sets employed. Similar comments can be made concerning the results of Olekalns (1996) using a near-century of Australian data.

Crosby and Otto (1999) move away from the money-inflation-output nexus discussed in many of the papers so far, in order to analyze the long-run connection between inflation and the capital stock using the methods of Fisher and Seater (1993) and King and Watson (1997). Crosby and Otto consider a bivariate VAR with inflation playing the role of the nominal variable, and the capital stock playing the role of the real variable. They use the long-run identifying restriction that shocks to the capital stock do not have permanent effects on the rate of inflation, which is similar to the long-run restriction sometimes employed in the papers discussed earlier. They construct an annual capital stock series for 64 countries using postwar data, with differing sample periods for different countries. Their unit-root diagnostics (ADF tests) indicate that 34 of these countries have both permanent shocks to inflation and to the capital stock. For these countries they test superneutrality with respect to the capital stock using their bivariate VAR. The Crosby and Otto estimates indicate that a permanent inflation shock has no statistically significant long-run impact on the capital stock for a large majority of the countries. Departures from this result are generally on the positive side, with a permanent inflation shock tending to raise the stock of capital in a country. Crosby and Otto argue that these results are robust to a number of changes in their analysis, including an alternative identifying restriction.

[14] Some data are missing, notably 1914-24 and 1939-49 for Germany and 1941-51 for Japan. Also missing are 1915-20 for Denmark, and 1940-45 for Norway.

[15] These results on orders of integration are somewhat different from those of the previous paragraph, even though the data set is the same, because Serletis and Koustas (1998) use different (and more standard) procedures to test for the presence of unit roots than Serletis and Krause (1996). In section four of their paper, Serletis and Koustas (1998) discuss the differences when the Zivot and Andrews (1992) methodology is used.

A real variable of interest to many economists is productivity. In an attempt to understand the long-run relationship between inflation and productivity, Sbordone and Kuttner (1994) devote a portion of their analysis to bivariate VAR methodology similar to that used by King and Watson (1997). They use data from the postwar United States, and they conclude that both series are reasonably characterized as $I(1)$. Sbordone and Kuttner use the King and Watson (1997) approach to identification, setting impact multipliers and long-run multipliers to various values in an effort to learn about the sensitivity of the results to alternative identification schemes. Under many of these schemes, the long-run impact of a permanent, positive shock to inflation on productivity is negative. If the identification scheme is the monetarist one—the long-run impact of a permanent shock to productivity on inflation is zero—then the estimated effect of a permanent inflation shock on productivity is negative but is not statistically different from zero.

Koustas and Serletis (1999) use the King and Watson (1997) methodology of searching across alternative identification schemes to study the Fisher relation between nominal interest rates and inflation rates. They employ data from the postwar period for 11 industrialized countries: Belgium, Canada, Denmark, France, Germany, Greece, Ireland, Japan, the Netherlands, the United Kingdom, and the United States. According to the authors' unit root diagnostic tests, all of these countries except two (Denmark and Japan) can reasonably be interpreted as possessing the nonstationarity in interest rates and inflation rates required to use the King and Watson techniques. The basic finding is that the long-run Fisher relation can be rejected across countries for a wide range of plausible identification assumptions. The authors also argue that taking tax effects into account accentuates this finding. The Koustas and Serletis results are more consistent across countries on this question than those of Weber (1994), who found more mixed results for a similar set of countries.

Rapach (1999a) is the first author to consider a trivariate VAR in this literature. His variables are the inflation rate, the nominal interest rate, and the level of real output. The data is from the postwar period for 14 industrialized (OECD) countries, where continuous observations on all three variables are available starting from the 1960s and extending to the mid-1990s.[16] Rapach (1999a) uses long-run identifying restrictions following Blanchard and Quah (1989); he needs three for the trivariate system. Rapach first extends the often-used monetarist restriction that permanent shocks to interest rates and output cannot have permanent effects on the inflation rate. Rapach's third restriction, also motivated by theoretical considerations, is that permanent shocks to output ("permanent technology shocks") leave the real interest rate unchanged. Since the long-run response of inflation to a permanent technology shock is already set to zero, this last restriction is accomplished by making the long-run response of the nominal interest rate to a permanent technology shock equal to zero.

Rapach uses unit-root diagnostic tests to conclude that the variables are reasonably described as $I(1)$ for these countries, and runs the trivariate VAR in an effort to estimate, primarily, the long-run responses of the level of real output to a permanent inflation shock, and of the real interest rate to a permanent inflation shock (the difference between two estimated long-run responses in this system). For all countries, the point estimates indicate that real interest rates fall in response to a permanent inflation shock. Moreover, these effects are generally statistically significant (or very close) at conventional significance levels. The point estimates also indicate that the response of the level of real output to a permanent inflation shock is positive for 11 of 14 countries, and four of these are statistically significant, or nearly so, at conventional significance levels. These latter results are generally consistent with the findings of Bullard and Keating for low inflation countries in a bivariate VAR framework. The results on real interest

[16] The countries are Australia, Austria, Belgium, Canada, Denmark, France, Ireland, Italy, Japan, the Netherlands, New Zealand, Sweden, the United Kingdom, and the United States.

rates are more strikingly in favor of nonsuperneutrality than the findings of Weber for G7 economies in bivariate systems. Weber searched over identification schemes while Rapach commits to a particular scheme, but Rapach studies interactions between three variables instead of two and analyzes more countries.

Related Methodology Using U.S. Data

Ahmed and Rogers (1996, 1998) work on empirical issues related to the long-run impact of permanent inflation shocks on real variables, but with methods somewhat different from those discussed earlier. In particular, Ahmed and Rogers construct a theoretical model economy and use this economy to motivate restrictions imposed in their empirical work. The model consists of an infinitely-lived representative agent who might hold money either because it enters the utility function or because of a cash-in-advance constraint. The technology for production of private sector real output is Cobb-Douglas, multiplied by a technology shift variable and also by a function of government size. Special cases of this framework (restrictions on theoretical parameters) deliver the standard results from the theoretical literature on monetary superneutrality surveyed by Orphanides and Solow (1990).

Ahmed and Rogers (1998) use annual U.S. data from 1889 to 1995 covering inflation, real output, real consumption expenditures, real investment, and the ratio of government spending to output. Much of the data is from Kendrick (1961). Based on diagnostic testing, Ahmed and Rogers conclude that a reasonable description of the data is that these series are $I(1)$, with the exception of the size of government, which they sometimes treat as $I(0)$. The authors then estimate cointegrating relationships for two specifications of the model. Based on these estimates, a permanent, positive shock to inflation is associated with a permanent drop in the consumption-output ratio and a permanent increase in the investment-output ratio.

These effects are large and statistically significant. The Fisher effect does not hold, as Ahmed and Rogers infer that real interest rates decline in the face of permanent, positive inflation shocks according to these estimates.

Ahmed and Rogers then turn to estimation of a VECM using the cointegrating relationships implied by their theoretical model, in an effort to find out what happens to the levels of the real variables following a permanent inflation shock. For two different specifications, the estimates indicate that a permanent shock to inflation increases the level of output, consumption, and investment. Ahmed and Rogers also consider variance decompositions and note that the inflation shock only accounts for a small fraction of the forecast error variance in consumption, investment, and output.[17] They interpret these results as follows: Permanent inflation shocks do not occur very often, but when they do, they have a significant impact on the economy. Accordingly, when looking at the data historically, one might reasonably abstract from inflation in building a model, but when contemplating significant changes in inflation rates, one should not assume the effects will be negligible.[18]

Bernanke and Mihov (1998a) test long-run monetary neutrality, and, like Ahmed and Rogers (1998), they depart from the methodology described in the main portion of this survey. In particular, Bernanke and Mihov use their own, larger VAR model of short-run monetary policy which is described in more detail in another paper (Bernanke and Mihov, 1998b) as a starting point. This model uses monthly data for the United States during the postwar era, and has the following variables: total bank reserves and nonborrowed reserves, both measured as deviations from a trend and the federal funds rate (collectively the *policy variables*); interpolated monthly real GDP and interpolated monthly GDP deflator inflation, an index of spot commodity prices and real balances (with money measured as M2). They use a semi-structural approach to derive identification restrictions based on relationships between the policy

[17] Rapach (1999a) also computes variance decompositions and concludes that inflation shocks do not explain a significant fraction of output forecast error variance.

[18] Ahmed and Rogers (1998) also consider subsamples. They find that the effects of inflation on real variables move in the same direction, but are much weaker, during the postwar period as opposed to the pre-WWI or the interwar period.

and nonpolicy variables, and among the policy variables, and they estimate the VAR. They do not examine the temporary-permanent dichotomy of the shocks to the variables in their system; the focus instead is on isolating an action that can reasonably be termed "a shock to monetary policy." Bernanke and Mihov's evidence in favor of long-run neutrality is based on the impulse-response functions of their estimated VAR: These functions show that the long-run (120-month) response of output to the policy shock is not significantly different from zero, although positive. At the same time, they find short-term impacts of the policy shock, such as a liquidity effect, that are in accord with conventional wisdom.

Bernanke and Mihov (1998a) then turn to an analysis of the robustness of their results, by considering alternative identification schemes in a manner analo-gous to the King and Watson (1997) methodology. They find that the evidence on long-run neutrality is in a sense stronger when one is willing to accept an identifica-tion that produces a smaller liquidity effect. They also find that imposing long-run neu-trality as an identifying restriction does not imply that one can reject their specification. Bernanke and Mihov conclude that these findings inspire confidence in their VAR model of monetary policy, since it is con-sistent with both a liquidity effect and long-run monetary neutrality.

CONCLUSIONS

This survey has covered a fair amount of territory. To avoid confusion about what the results actually say, this section includes a summary of the main findings organized by the nature of the proposition.

Long-Run Monetary Neutrality. In this survey, we did not find much evidence against the long-run neutrality of money. Fisher and Seater (1993) usefully reinter-preted some of the major time series studies on neutrality published in the 1970s and 1980s as consistent with long-run monetary neutrality, and uninformative regarding long-run monetary superneutrality

(Geweke's 1986 title notwithstanding!), once one takes proper account of the time series properties of the data. While Fisher and Seater (1993) found evidence against long-run neutrality with respect to real output for the United States during the last century, Boschen and Otrok (1994) pointed out that such a result did not hold once the Great Depression years were excluded from the sample. In another comment on this question, Haug and Lucas (1997) could not reject long-run neutrality in a century of Canadian data. Olekalns (1996) did find some evidence against long-run neutrality in a near-century of Australian data using a broad money measure, but the neutrality hypothesis could not be rejected using a narrower money measure. Coe and Nason (1999) find that long-run neutrality cannot be rejected for a century of U.S. data when the monetary base is the monetary variable, nor could they reject long-run neutrality for a century of U.K. data.[19]

Long-run neutrality received support from the studies focused exclusively on the postwar U.S. data. King and Watson (1997) searched over a wide range of identification schemes and found little evidence against long-run neutrality. Boschen and Mills (1995), studying a larger system of vari-ables with cointegration techniques, but without the extensive robustness checking, also found little reason to doubt long-run neutrality. Bernanke and Mihov (1998a) argue that their model is consistent with long-run neutrality using the postwar U.S. data; like King and Watson (1997), they explore the robustness of their findings to an extensive range of alternative identifica-tion schemes.

The studies that used data from more than one country also found general support for the long-run monetary neutrality pro-position. For instance, Weber (1994), using techniques similar to those of King and Watson (1997), generally supports long-run neutrality for the G7 economies during the postwar period across a wide variety of identification schemes, especially when money is measured using broader monetary aggregates. Weber's results also confirm the findings of King and Watson

[19] Coe and Nason also raise important questions concerning the statistical properties of the Fisher and Seater long-horizon regression test.

(1997) and Boschen and Mills (1995) for the postwar U.S. data. Serletis and Krause (1996) use the Backus and Kehoe (1992) data set for 10 industrialized countries, including the United States and Australia, covering more than a century. They found general support for long-run monetary neutrality, even in the United States and Australia, where Fisher and Seater (1993) —for the United States—and Olekalns (1996)—for Australia—had cast doubt. In a more extensive study, Serletis and Koustas (1998), use the same data set and apply the King and Watson (1997) technology of searching over plausible identification schemes. Here, wide support for long-run monetary neutrality is found across industrialized countries and plausible identification schemes, even though the data set is a very long-time series of the type that sometimes displayed evidence against long-run monetary neutrality in previous studies.

Long-Run Monetary Superneutrality. This survey has also shown that the evidence in favor of long-run monetary superneutrality is far more mixed. This is perhaps not too surprising since, as was stressed in the introduction, it is a relatively simple matter to write down neoclassical, market clearing, rational expectations theories in which superneutrality does not hold. In addition, since inflation is generally regarded as a distortionary force in macroeconomic systems, we might reasonably expect real variables to be altered in the face of permanent shocks to money growth and inflation.

Analyzing postwar U.S. data, King and Watson (1997) find that rejection of long-run monetary superneutrality with respect to real output is possible for a range of identification schemes they consider reasonable. Weber (1994) confirms this result using similar methodology across the G7 economies during the postwar period. Bullard and Keating (1995) analyze data from a number of countries worldwide during the postwar period. They consider permanent inflation shocks and the subsequent reaction of the level of real output. The results generally support superneutrality, but Bullard and Keating note that in the low inflation countries (such as the G7), point estimates tend to be positive and are sometimes statistically significant. Serletis and Koustas (1998) reject long-run monetary superneutrality for Italy over the last century, in a bivariate system with money and real output, over a range of identifying restrictions. Crosby and Otto (1998) generally find that permanent inflation shocks have little or no permanent effect on the level of the capital stock in a large sample of countries during the postwar period. When they do find statistically significant effects, permanently higher inflation is associated with a permanently higher capital stock. In Rapach's (1999a) study of a trivariate VAR using postwar data from 14 OECD countries, permanent inflation shocks generally were associated with permanently higher levels of real output and, more strikingly, with permanently lower real interest rates. Ahmed and Rogers (1998), using methodology that departs somewhat from the other studies in the survey, consider a century of U.S. data and conclude that permanent inflation shocks have permanent, positive effects on important real variables, including output, consumption, and investment. They also stress that these shocks do not explain a large portion of the forecast error variance in the data.

While the overall evidence on these questions is mixed, considering only lower inflation countries leads to the conclusion that permanently higher money growth or inflation is associated with permanently higher output and permanently lower real interest rates. As Ahmed and Rogers (1998) stress, this result is inconsistent with many —almost all?—current quantitative business cycle models, which generally predict that permanently higher inflation permanently lowers consumption and output. There is little support for such a prediction in the studies surveyed here. This is an important empirical puzzle that stands as a challenge for future research.

Related Propositions. We have also seen in this survey a smattering of evidence on other, related neutrality propositions.

King and Watson (1994, 1997) analyze the slope of a long-run Phillips curve in the postwar U.S. data, and find that a vertical curve is a reasonable approximation. Weber (1994) reports generally similar findings for the G7 economies.

King and Watson (1997) also study Fisher relations between interest rates and inflation, and conclude that nominal interest rates do not adjust one-for-one with permanent inflation shocks under a wide range of plausible identification schemes. Weber (1994) finds more mixed results for the G7 economies, but evidence presented by Rapach (1999a) and Koustas and Serletis (1999) is squarely on the side of less than one-for-one adjustment across industrialized nations.

Bullard and Keating (1995) comment on the effect of permanent inflation shocks on long-run economic growth. Because growth rates are generally stationary according to diagnostic tests, and inflation rates often are not, the methodology of Fisher and Seater (1993) and King and Watson (1997) suggests that permanent inflation shocks have no permanent effect on economic growth. Ahmed and Rogers (1998) include a comment in a similar vein.[20]

Areas for Further Research. Canova (1994), commenting on Weber (1994), stressed that the nature of the methodology of Fisher and Seater (1993) and King and Watson (1997) —however correct it may be from a logical point of view—places heavy reliance on the existence of (and on the number of) unit roots in the time series being studied. Canova comments that these tests of neutrality propositions depend in an important way on getting the inference on the unit root correct, and yet, tests for unit roots are known to have low power. Most authors, including Weber (1994), are well aware of this issue, and many use a battery of tests for a unit root in a series or other measures in an effort to be conservative about their conclusions in this regard. But Canova (1994, p. 121) notes, nevertheless, that this procedure "... conditions the results of economic hypotheses on shaky statistical ground. ..."

One problem is in the nature of the unit-root diagnostic tests. Since time series characterized by a unit root have such different properties from stationary time series, the researcher is forced into a declaration of a unit root or not. Once this declaration is made, the researcher can proceed with further analysis. This brings to mind a possible role for fractional integration in testing neutrality propositions. In fact, this possibility has been explored recently in a study by Bae and Jensen (1998). More work in this area may be fruitful in the future.

Canova (1994) also comments that Weber's (1994) results are based on bivariate VAR systems, as are many others reported here. He worries that the results may not be the same when larger systems are explored. Some papers surveyed here have taken steps in that direction, including Boschen and Mills (1995), Ahmed and Rogers (1998), Rapach (1999a), and Bernanke and Mihov (1998a). These studies generally have supported results from the bivariate VARs. However, much more could be done in multivariate systems than has been completed to date.

Even without turning to multivariate systems, one notes that much of the work surveyed has focused on real output, and that less work has been done on the long-run bivariate relationship between money or inflation and other important real variables. One exception is Crosby and Otto (1999), who take a step in this direction, using the capital stock instead of real output as their primary real variable.[21] A good deal more could be done by simply investigating the long-run bivariate relationships more systematically for variables other than real output.

All of the analyses surveyed consider one country at a time when testing neutrality propositions. One would like to know if a panel approach, implemented for a group of similar countries like the G7, would produce results similar to the ones reported in the studies surveyed here, or if important interactions between the countries are being left out. A simpler line would be to study multivariate systems for a single country that attempt to account for cross-border effects by including an international

[20] This idea is also related to work by Jones (1995).

[21] For another exception see Rapach (1999b).

variable. This would seem to be particularly important for some smaller, open economies sometimes included in these studies.[22]

While problems certainly remain, it seems that the 1990s have been quite fruitful in this area of empirical macroeconomics. Tests of neutrality propositions not subject to the critique of Lucas (1972) and Sargent (1971)—tests that have eluded economists during much of the postwar era—have been devised and executed for a variety of times and places. This body of work gives us economists what is perhaps our first glimpse at the evidence on long-run monetary neutrality and superneutrality, and allows assessment of the merits of these propositions separate from the logical force of theoretical arguments.

REFERENCES

Ahmed, Shaghil, and John H. Rogers. "Long-Term Evidence on the Tobin and Fisher Effects: A New Approach," International Finance Discussion Paper #566, Board of Governors of the Federal Reserve System (September 1996).

_____, and _____ . "Inflation and the Great Ratios: Long-Term Evidence From the U.S.," International Finance Discussion Paper #628, Board of Governors of the Federal Reserve System (October 1998).

Andersen, Leonall, and Denis Karnovsky. "The Appropriate Time Frame For Controlling Monetary Aggregates: The St. Louis Evidence," in Controlling Monetary Aggregates II: The Implementation, Federal Reserve Bank of Boston, 1972.

Backus, David K., and Patrick J. Kehoe. "International Evidence on the Historical Properties of Business Cycles," American Economic Review (September 1992), pp. 864-88.

Bae, Sangkun, and Mark J. Jensen. "Long-Run Neutrality in a Long-Memory Model," working paper, University of Missouri at Columbia (December 1998).

Barro, Robert J. "Inflation and Growth," this Review (May/June 1996), pp. 153-69.

Bernanke, Ben S., and Ilian Mihov. "The Liquidity Effect and Long-Run Neutrality," National Bureau of Economic Research Working Paper #6608 (June 1998a).

_____, and _____ . "Measuring Monetary Policy," Quarterly Journal of Economics (August 1998b), pp. 869-902.

Blanchard, Olivier J., and Danny Quah. "The Dynamic Effects of Aggregate Demand and Supply Disturbances," American Economic Review (September 1989), pp. 655-73.

Boschen, John F., and Leonard O. Mills. "Tests of Long-Run Neutrality Using Permanent Monetary and Real Shocks," Journal of Monetary Economics (February 1995), pp. 25-44.

_____, and Christopher M. Otrok. "Long-Run Neutrality and Superneutrality in an ARIMA Framework: Comment," American Economic Review (December 1994), pp. 1470-73.

Bullard, James, and John W. Keating. "The Long-Run Relationship Between Inflation and Output in Postwar Economies," Journal of Monetary Economics (December 1995), pp. 477-96.

Canova, F. "Testing Long-Run Neutrality: Empirical Evidence For G7 Countries With Special Emphasis on Germany," Carnegie-Rochester Conference Series on Public Policy (1994), pp. 119-25.

Coe, Patrick J., and James M. Nason. "Long-Run Monetary Neutrality in Three Samples: The United Kingdom, the United States, and the Small," manuscript, University of Calgary and University of British Columbia (May 1999).

Crosby, Mark, and Glen Otto. "Inflation and the Capital Stock," Journal of Money, Credit, and Banking, forthcoming (1999).

Fischer, Stanley. "Why are Central Banks Pursuing Long-Run Price Stability?" in Achieving Price Stability, Federal Reserve Bank of Kansas City Symposium, 1996.

Fisher, Mark E. "The Time Series Implications of the Long-Run Neutrality and Superneutrality of Money," Ph.D. Dissertation, University of Chicago (1988).

_____, and John J. Seater. "Neutralities of Money," manuscript, Board of Governors of the Federal Reserve System and North Carolina State University (1989).

_____, and _____. "Long-Run Neutrality and Superneutrality in an ARIMA Framework," American Economic Review (June 1993), pp. 402-15.

Friedman, Milton, and Anna Schwartz. A Monetary History of the United States, 1867-1960, Princeton University Press, 1963.

_____, and _____. Monetary Trends in the United States and the United Kingdom, University of Chicago Press, 1982.

Geweke, John F. "The Superneutrality of Money in the United States: An Interpretation of the Evidence," Econometrica (January 1986), pp. 1-21.

Haug, Alfred A., and Robert F. Lucas. "Long-Run Neutrality and Superneutrality in an ARIMA Framework: Comment," American Economic Review (September 1997), pp. 756-59.

Hoffman, Dennis L., and Robert H. Rasche. Aggregate Money Demand Functions, Kluner Academic Publishers, 1996.

Jefferson, Philip N. "On the Neutrality of Inside and Outside Money," Economica (November 1997), pp. 567-86.

Jones, Charles I. "Time Series Tests of Endogenous Growth Models," Quarterly Journal of Economics (May 1995), pp. 495-525.

Kendrick, J. Productivity Trends in the United States, Princeton University Press, 1961.

[22] Nason and Rogers (1999) use methodology related to that discussed in this survey to study investment and current account dynamics for Canada.

King, Robert G., and Mark W. Watson. "Testing Long-Run Neutrality," working paper #4156, National Bureau of Economic Research (September 1992).

_____, and _____. "The Postwar U.S. Phillips Curve: A Revisionist Econometric History," *Carnegie-Rochester Conference Series on Public Policy* (1994), pp. 157-219.

_____, and _____. "Testing Long-Run Neutrality," Federal Reserve Bank of Richmond *Economic Quarterly* (Summer 1997), pp. 69-101.

Kormendi, Roger C., and Phillp G. Meguire. "Cross-Regime Evidence of Macroeconomic Rationality," *Journal of Political Economy* (October 1984), pp. 875-908.

Koustas, Zisimos. "Canadian Evidence on Long-Run Neutrality Propositions," *Journal of Macroeconomics* (Spring 1998), pp. 397-411.

_____, and Apostolos Serletis. "On the Fisher Effect," *Journal of Monetary Economics* (August 1999), pp. 105-30.

Lothian, James R. "Equilibrium Relationships Between Money and Other Economic Variables," *American Economic Review* (September 1985), pp. 828-35.

Lucas, Robert E., Jr. "Econometric Testing of the Natural Rate Hypothesis," in *The Econometrics of Price Determination,* Board of Governors of the Federal Reserve System, 1972.

_____. "Two Illustrations of the Quantity Theory of Money," *American Economic Review* (December 1980), pp. 1005-14.

_____. "Nobel Lecture: Monetary Neutrality," *Journal of Political Economy* (August 1996), pp. 661-82.

Marty, Alvin L. "What is the Neutrality of Money?" *Economics Letters* (April 1994), pp. 407-09.

McCandless, George T., and Warren E. Weber. "Some Monetary Facts," Federal Reserve Bank of Minneapolis *Quarterly Review* (Summer 1995), pp. 2-11.

Nason, James M., and John H. Rogers. "Investment and the Current Account in the Short Run and the Long Run," manuscript, University of British Columbia (September 1999).

Nelson, Charles R., and Charles I. Plosser. "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics* (September 1982), pp. 139-62.

Olekalns, Nilss. "Some Further Evidence on the Long-Run Neutrality of Money," *Economics Letters* (March 1996), pp. 393-98.

Orphanides, Athanasios, and Robert M. Solow. "Money, Inflation and Growth," in *Handbook of Monetary Economics,* North-Holland, 1990.

Rapach, David E. "International Evidence on the Long-Run Superneutrality of Money," Working Paper, Trinity College (1999a).

_____. "Monetary Shocks and Real Exchange Rate Hysteresis: Evidence From the G-7 Countries," *Review of International Economics,* (April 1999b).

Sargent, Thomas J. "A Note on the 'Accelerationist' Controversy," *Journal of Money, Credit, and Banking* (August 1971), pp. 721-25.

_____. "The End of Four Big Inflations," in *Rational Expectations and Inflation,* Harper and Row, 1986.

Sbordone, Argia M., and Kenneth Kuttner. "Does Inflation Reduce Productivity?" Federal Reserve Bank of Chicago *Economic Perspectives* (November/December 1994), pp. 2-14.

Serletis, Apostolos, and Zisimos Koustas. "International Evidence on the Neutrality of Money," *Journal of Money, Credit, and Banking* (February 1998), pp. 1-25.

_____, and David P. Krause. "Empirical Evidence on the Long-Run Neutrality Hypothesis Using Low-Frequency International Data," *Economics Letters* (March 1996), 323-27.

Weber, A. "Testing Long-Run Neutrality: Empirical Evidence For G7 Countries With Special Emphasis on Germany," *Carnegie-Rochester Conference Series on Public Policy* (1994), pp. 67-117.

Zivot, Eric, and Donald W. K. Andrews. "Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Journal of Business and Economic Statistics* (July 1992), pp. 251-70.