2 MONETARIST INTERPRETATIONS OF THE GREAT DEPRESSION: *An Evaluation and Critique*

Robert J. Gordon and James A. Wilcox

Explanations which run in terms of one single cause have been more and more discredited and should be regarded with suspicion.

—Haberler (1958, p. 5)

Between the early 1960s and mid-1970s the Great Depression received surprisingly little attention from economists. This fascinating period, the original combat zone that pitted monetarists against nonmonetarists, seemed until recently a neglected orphan, too young to be worthy of serious study by economic historians but too old to possess the easily accessible Commerce Department quarterly national-income data that today's macroeconometricians view as qualifying an era for detailed scrutiny. Only within the past few years has the

This paper is dedicated to the memory of the late R. A. Gordon. The research was supported by the National Science Foundation. The authors are grateful to James Glassman for his help. They benefited from the helpful suggestions of Clarence L. Barber, Frank Brechling, Louis Cain, Steve Easton, Robert Eisner, R. A. Gordon, James Lothian, Robert Lucas, Thomas Mayer, Donald McCloskey, Allan Meltzer, Frederic Mishkin, Joel Mokyr, George R. Neumann, and Steve Sheffrin.

K. Brunner (ed.), *The Great Depression Revisited*

orphan grown up sufficiently to attract the attention of a prominent economic historian, Peter Temin, whose attack (1976) on the Friedman-Schwartz (1963*a*) monetary explanation of the depression has helped to open up a new round of controversy, including the recent contributions of Meltzer (1976), Mayer (1978*a*, 1978*b*), and Schwartz (1981).

A limitation of the Temin book and the subsequent debate has been its relatively narrow focus on the first two years of the contraction (1929–31) and on the relation between money and income. As Mayer (1978*b*) points out, a study that emphasizes conditions in 1929 and 1930 cannot effectively criticize the main thrust of the Friedman-Schwartz analysis, which pays scarcely any attention to the first year of the contraction and concentrates on the period subsequent to the first wave of bank failures in October 1930. In a sense, monetarists and their opponents are like two knights in a jousting match who ride by each other without ever making contact. Monetarists consider virtually the only interesting question to be the source of the unique depth and severity of the depression and naturally concentrate on the 1931–33 phase when the contraction exhibited an unprecedented acceleration. The nonmonetarist opponents tend to concentrate on the initial decline in private spending, which, they claim, brought the bank failures and monetary collapse in its wake.

This paper rejects the proposition that there is only a single interesting question to ask about the decade of the 1930s. It is concerned not only with the role of money in the 1929–33 contraction but also with the relative role of monetary and nonmonetary factors in the recession of 1937–38 and subsequent recovery and, in addition, with the division of nominal-income change between prices and real output.[1] New empirical evidence bearing on each of these issues is provided.

The results suggest that both extreme monetarist and nonmonetarist interpretations of the decade of the 1930s are unsatisfactory and leave interesting features of the data unexplained. Arguing against acceptance of an extreme monetarist interpretation are (1) the inability of changes in the money supply alone to explain the severity of the initial collapse in income between 1929 and the fall of 1931, (2) the steady weakening of the correlation between changes in nominal income and money as the 1930s progressed, (3) the failure of monetary factors to explain the nature and timing of the 1938–41 recovery, and (4) the apparent absence of any tendency for the mechanism of price flexibility to provide strong self-correcting forces as required by an approach that stresses monetary rules and opposes policy activism. Arguing against acceptance of an extreme nonmonetarist interpretation are (1) the close association between the collapse in income and the lagged effect of monetary changes after the fall of 1931, (2) the milder contraction and earlier recoveries associated with the

more expansive monetary policies pursued in Europe, (3) the close association between money and income in the 1937-38 recession, and (4) the failure of the price-change data to adhere to the expectational Phillips curve approach imbedded in many postwar econometric models constructed by nonmonetarists.[2]

The debate surrounding monetarist interpretations of the Great Depression does not center on the potency of monetary changes as a cause of income variation. Although some economists in the early 1960s treated the quantity theory and the Keynesian income-expenditure theory as mutually exclusive analytical frameworks, from today's vantage point the 1965 "battle of the radio stations" regarding whether only money matters or money never matters seems quaintly anachronistic.[3] Recently the monetarist controversy has been reoriented, as a result of an emerging consensus on both sides that both monetary and nonmonetary factors "matter" for the determination of income (Stein 1976). Instead, the central issues separating the monetarists and their opponents include the merits and potential benefits and costs of government policy activism, both monetary and fiscal, and the stability and inherent self-correcting properties of the private economy.

This new perspective can be summarized by constructing a "monetarist platform," which brings together in four "planks" the monetarist position on the remaining areas of disagreement:[4]

Plank 1: Without the interference of demand shocks introduced by erratic government policy, private spending would be stable, because people base their consumption plans on a relatively stable "permanent" concept of income.

Plank 2: Even if private planned spending is not completely stable, flexible prices create a natural tendency for it to come back on course.

Plank 3: Even if private planned spending is not completely stable and prices are not completely flexible, an activist monetary and fiscal policy to counteract private demand swings is likely to do more harm than good.

Plank 4: Even if prices are not completely flexible, so that the economy can wander away from equilibrium in the short run, there can be no dispute regarding the increased flexibility of prices, the longer the period of time allowed for adjustment.

From this orientation, a modern monetarist would not be required to devote excessive attention to showing that money played a major causal role in the Great Depression, because the potency of money is no longer a matter for debate.[5] He would be more interested in denying that autonomous swings

in private spending, not explainable by movements in government policy or in permanent income, played a major role in the contraction of 1929–33 or in the subsequent recovery. And he would be particularly concerned with the issue of price behavior in the 1930s. Did the economy display strong self-correcting forces in the form of flexible prices that would have tended to bring the economy back to its natural unemployment rate without the need for government intervention?

This paper is divided into two main sections. The first evaluates the relative contributions to nominal-income behavior of private spending behavior and government actions. The central focus is the same question that concerns both Temin (1976) and Schwartz (1981)—whether money played no role in the first two years of the contraction (the Temin position), or whether autonomous private spending movements played no role (the Schwartz position). But the scope of our analysis is broader than an evaluation of the Temin-Schwartz debate regarding 1929–31, and our purview extends to the whole decade of the 1930s.

The final section of the paper investigates the potency of the economy's self-correcting mechanism of price flexibility—a pivotal question in the monetarist controversy but one that is given no attention at all by Temin, Schwartz, or most other recent writers. Monetarists not only tend to give greater credence to price flexibility as a source of self-correction in the private economy but also tend to adopt an analytic framework that differs from that of nonmonetarists.

Monetarists tend to view deviations of output from equilibrium ("natural output") as being a voluntary response of firms and workers to deviations of actual prices from their expected level. This "price surprises cause output changes" framework is evident both in theoretical writings and in empirical research.[6] Nonmonetarists, on the other hand, tend to discuss the same problems in terms of a disequilibrium-adjustment framework.[7] Empirical nonmonetarist explanations of wage and price change tend to place deviations between actual and expected inflation on the left-hand side of the equation and measures of commodity-market or labor-market disequilibrium on the right-hand side.[8]

The most dramatic recent contribution tending to support the monetarist belief in self-correction is Darby's (1976*b*) attempt to remeasure unemployment during the Great Depression and show that in the late 1930s unemployment was rapidly returning to its natural level as agents adjusted the deviation between actual and expected prices. In this paper we present new evidence on the relation between prices, expected prices, unemployment, and output, in an attempt to reassess the potency of the economy's self-correcting mechanism of price flexibility.

MONETARY AND OTHER EXPLANATIONS OF NOMINAL-INCOME CHANGE

Distinguishing Hypotheses

Temin's entire book (1976) is devoted to an examination of two views: the "money hypothesis" and the "spending hypothesis." In order to clarify the positions held by the various protagonists and to judge their consistency with the data, we will distinguish a broader spectrum of four views, ranging from hard-line monetarism to hard-line antimonetarism.

(a) "Hard-line monetarism." The 1929–33 contraction was both initiated and aggravated by monetary factors, and nonmonetary factors played no role. The prime exponent of this view is Schwartz (1981), who has departed from her earlier advocacy in Friedman and Schwartz (1963*a*) of view (*b*), which admits the possible role of nonmonetary forces in initiating the contraction.[9]

(b) "Soft-line monetarism," the Friedman-Schwartz position. Any combination of factors, both monetary and nonmonetary, could have caused the initial stage of the contraction through the first wave of bank failures in late 1930. But from that point, bank failures played a crucial role in converting a serious recession into a deep depression. The decline in the stock of money, while itself aggravated by the severity of the contraction, did not play a purely passive role but instead worsened the decline in income. As a result, aggressive open-market purchases by the Federal Reserve could have lessened the severity and duration of the depression. This view differs from the more extreme position (*a*) in its explicit admission that the initial phase of the contraction could have been due to nonmonetary factors, that the money supply is at least partly endogenous, and that at least part of the 1929–33 decline in the supply of money could therefore have been caused by nonmonetary factors.[10] Following Hicks (1974), we may identify this version of soft-line monetarism as the theory of the "double slump," in which a first phase of a severe depression was followed, not by a recovery, but by a second, more severe phase caused by monetary factors.[11]

(c) "Soft-line nonmonetarism." This position emphasizes nonmonetary factors as sources of the 1929–33 contraction, while not denying the possible role of money in aggravating the slump. The behavior of housing construction and international factors are most often emphasized. Bolch and Pilgrim's (1973) study linking the housing slump to a decline in household formation is an example of this genre and is classified under category (*c*) because of the explicit inclusion of monetary factors in individual equations in the model. R. A. Gordon's work (1951, 1974) emphasizes overinvestment in both housing and other industries but does not deny a role for monetary factors.[12]

(*d*) *"Hard-line nonmonetarism."* Temin's recent work is the most notable example of this extreme view, which was predominant in the 1940s and 1950s but which has become increasingly rare since the early 1960s. Temin limits his advocacy of this extreme view to the interval between October 1929 and September 1931, but within this two-year period his sweeping claim is unguarded: "There is no evidence of *any* effective deflationary pressure from the banking system between the stock-market crash in October, 1929, and the British abandonment of the gold standard in September, 1931" (1976, p. 169, emphasis added).

Since the views labeled (*b*) and (*c*) differ only in emphasis, it is impossible to distinguish their validity with any degree of precision. Although their emphasis is very different, Friedman and Schwartz and R. A. Gordon would probably agree that both bank failures and other nonmonetary factors played at least *some* role in the 1929–33 contraction. Since interactions between money and spending may dominate the effect of either force taken by itself, any attempt to split up the contraction into the share due to money and the share due to a particular nonmonetary factor—for example, housing—is an unproductive scientific enterprise that is bound to satisfy no one. Instead, the real question is whether either extreme view (*a*) or (*d*) can be excluded.

The Temin Claim That Money Did Not Matter at All

The data show that the money-supply concept M_2 declined by 2.5 percent during the first four quarters of the contraction and by another 7.9 percent during the second four quarters.[13] For Temin to hold the extreme position (*d*), he must deny that this decline, whatever its source, had any effect at all on the level of nominal income. His position is surprising, since it conflicts with almost all econometric work on postwar data, ranging from the St. Louis model of Andersen and Jordan (1968), to the reduced-form money-income equations of Sims (1972, 1977), to the large-scale structural models best represented by MPS (Ando and Modigliani 1976).

Temin's case rests on two propositions. First, for the decline in real output to have been caused by monetary stringency, interest rates should have been observed to increase. In terms of the classroom *IS–LM* model, if it is claimed that *IS* movements (autonomous shifts in investment and consumption spending) were unimportant, then the decline in output could only be explained by a leftward shift in the *LM* curve, which would have caused interest rates to increase unless the *IS* curve were horizontal. But short-term interest rates on risk-free securities actually exhibited a sharp decline throughout the 1929:3–1931:3 period. Second, Temin adds, the position of the *LM* curve de-

pends on the level of real balances and thus could not have shifted leftward in the light of the increase in real balances that actually occurred through 1931:3

Figure 1 plots the level of real balances (M_2/P) and exhibits the increase observed by Temin during the interval 1930:2–1931:2.[14] Temin's defense of view (*d*) collapses, however, if we can show that the 1929–31 decline in interest rates and increase in M_2/P are logically consistent with a model in which nominal spending depends positively on nominal money.

The situation described by solid lines in figure 2 describes an initial *IS–LM* equilibrium. The positive slope of the *LM* curve reflects a nonzero interest elasticity of the demand for money, and its position depends on the level of real balances (M/P). The negative slope of *IS* reflects a nonzero interest elasticity of investment and/or consumption spending, and its position depends on the level of "autonomous spending" ($\bar{A}$—exports, government spending, and the autonomous components of consumption and investment, which in turn depend partly on tax rates). When the *LM* and *IS* curves have the designated slopes, the aggregate demand curve *DD* in the bottom frame in P, Q space has a negative slope and a position that depends on autonomous spending and the *nominal* money supply. *DD* traces the locus of all intersections of *IS* and *LM* for given $\bar{A}$ and M. So this is a model in which a shift in nominal money shifts the *DD* curve and nominal income and thus is consistent with the positive effects of money on spending found in postwar econometric results.[15]

But the model in figure 2 can also easily explain the decline in short-term interest rates and increase in real balances on which Temin rests his argument. The necessary ingredient is a drop in the level of autonomous spending from $\bar{A}_0$ to $\bar{A}_1$. If we initially hold constant the level of nominal money at M_0, the *IS* curve shifts left from IS_0 to IS_1, and the aggregate demand curve shifts left from DD_0 to DD_1. The price level drops from P_0 to P_1, output falls from Q_0 to Q_1, and the interest rate drops from r_0 to r_1.

So the movements in the variables all go in the direction noted by Temin; nevertheless, nominal money *does matter*. Let nominal money drop from M_0 to M_1, and both output and prices will drop further to Q_2 and P_2. Because the aggregate supply curve is positively sloped, rather than vertical, the price level must fall by less than the money supply, and so M/P must fall and the interest rate must rise in situation E_2 as compared to E_1. Because the price level is altered by a change in nominal money, one cannot deduce monetary impotence from movements in real balances or interest rates.[16]

Because the argument in figure 2 relies on a shift in autonomous spending from $\bar{A}_0$ to $\bar{A}_1$, it is incompatible with the extreme hard-line monetarist view (*a*). In principle the economy could reach point E_2 by a different process. The argument presented in figure 2 assumes a zero expected rate of deflation. If in fact the negative 1929–31 rate of change of prices was rapidly incorporated

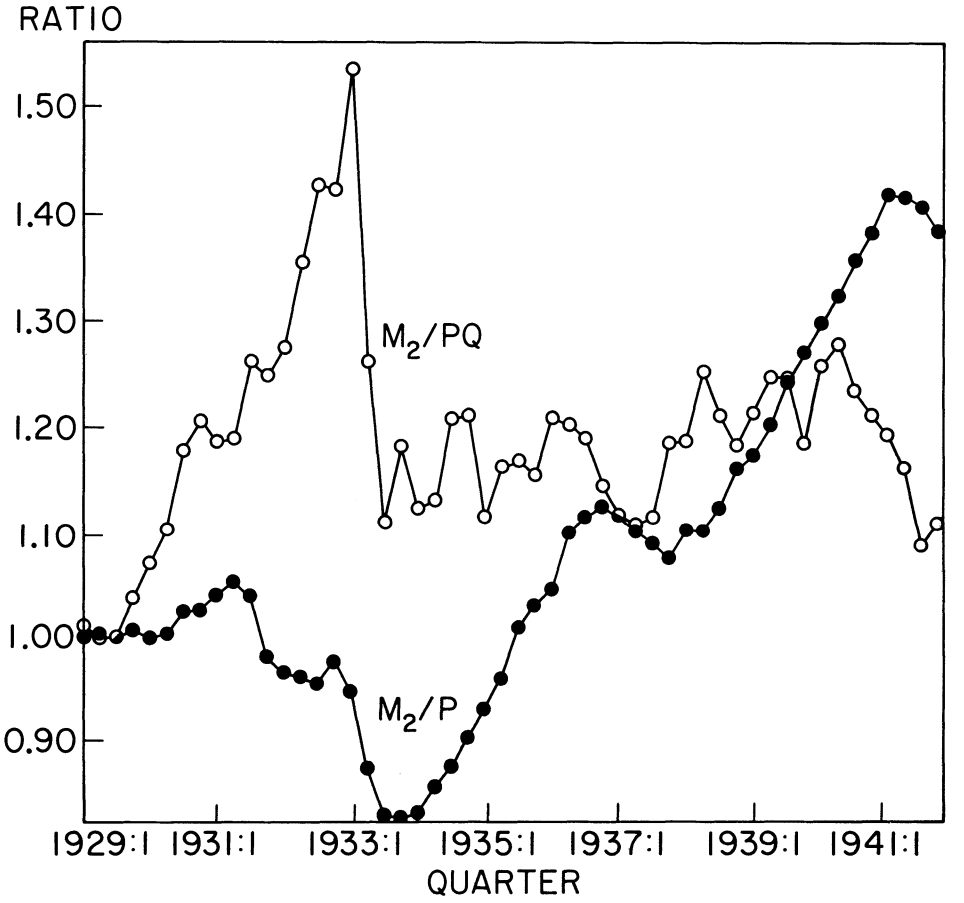

Figure 1. The Real Money Supply (M_2/P) and the Inverse of Velocity (M_2/PQ), 1929:1–1941:4 (1929:3 = 1.0)

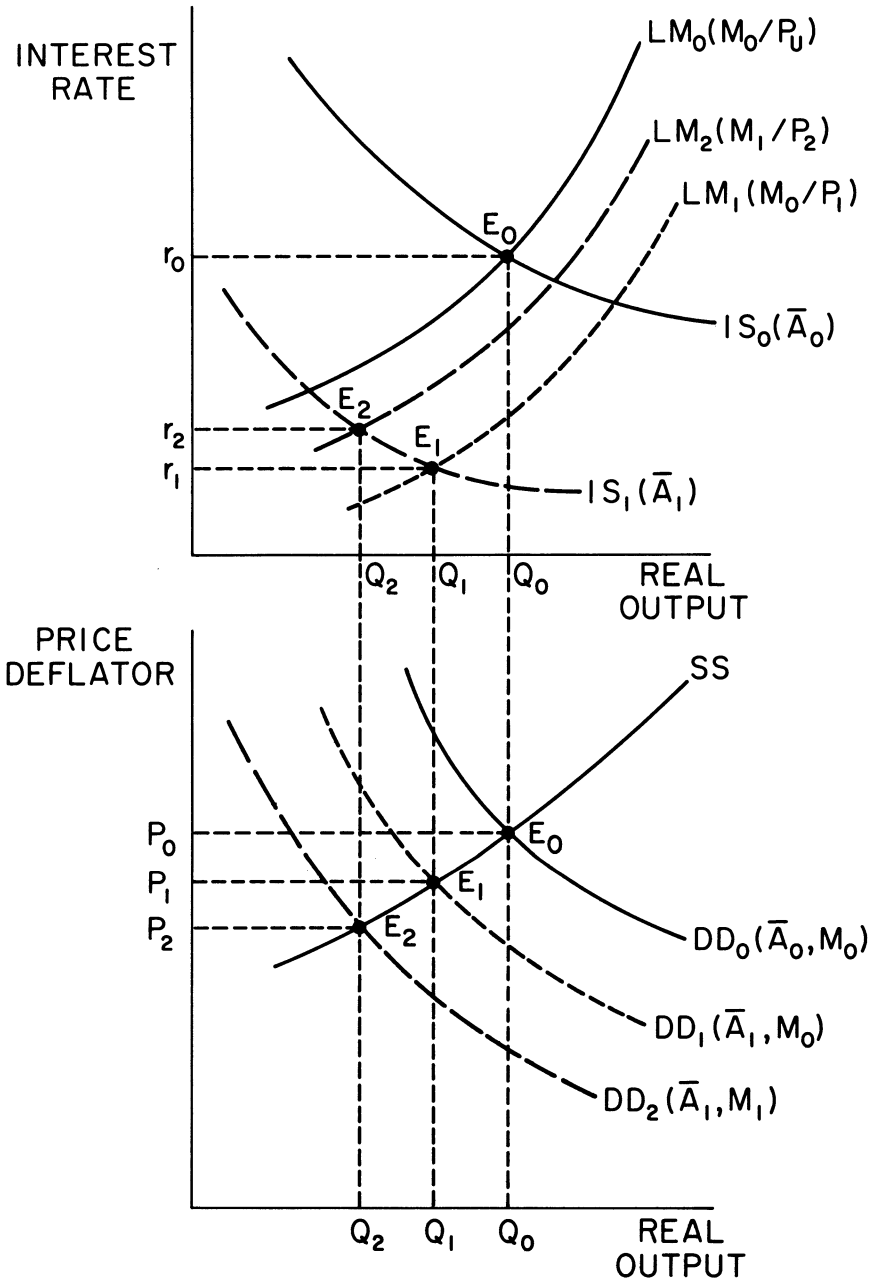

Figure 2. Effects of Nominal Money and Autonomous Spending Changes

into expectations, then the real interest rate would lie above the nominal interest rate. Because the *LM* curve is defined for the nominal rate (on which the demand for real balances depends) and the *IS* curve is defined for the real rate, it would be necessary to draw in a second *IS* curve in terms of the nominal interest rate. This would be displaced vertically below IS_0 by the rate of expected deflation. There is no reason why this lower curve, call it IS_i, could not yield the same intersection point E_2 in figure 2. Thus all the earlier statements about position E_2 would hold, even though autonomous spending had remained completely constant.[17]

Yet those who would rely completely on price deflation caused by a declining money supply to explain the first year of the 1929–33 contraction—leaving no room at all for autonomous spending to play a role—surely strain credulity. Consider the situation in 1930:2. M_2 had fallen only 1.8 percent from its 1929:3 peak. The GNP deflator had declined by only 2.2 percent. In the entire period between 1921:3 and 1929:3, eight full years, the GNP deflator varied over a range of only 4.4 percentage points, and the 1929:3 observation was almost exactly in the middle of the range. Why should economic agents in the spring of 1930 suddenly have started to expect a deflation substantial enough to explain the observed decline in nominal interest rates, when actual price behavior still remained within the range of an eight-year period that had been characterized by remarkable price stability?

Despite the very small declines in M_2 and P over this first three-quarter period, real output declined by 9.6 percent. Velocity declined by 9.9 percent.[18] Without a sudden and inexplicable shift from stable-price expectations to expectations of deflation, the first three quarters of the contraction must be explained by a leftward shift in the *IS* curve due to a decline in autonomous spending. This conclusion is consistent with the more formal simulation results presented later in figure 3 and table 3.

The expected-deflation argument becomes increasingly plausible after the summer of 1930. In 1930:3 the GNP deflator broke out of the range observed during the 1920s. By 1931:3 it had declined 13.6 percent below the 1929:3 peak and 11 percent below the lowest value observed in the 1920s. It is not implausible that expectations of deflation began in late 1930 to shift the *IS* curve downward, although in figure 1 it appears that the decline in velocity (increase in $1/V$) was interrupted between 1930:3 and 1931:2. Thus a scenario that appears consistent with the ratios in figure 1 would have the initial three quarters of the contraction explained by a sharp leftward shift of *IS* due to a decline in autonomous spending. After 1930:2 the decline in M_2 began in earnest, offsetting the downward pressure on velocity of the continuing *IS* shift. After 1931:2 a deflationary spiral began, in which deflationary expectations shifted down *IS*, while M_2 began falling more rapidly

than prices, thus shifting the *LM* curve to the left as well. And, as Tobin (1975) has recently reminded us, the depressing impact on expenditures of a price deflation can include not only upward pressure on the real interest rate and resulting postponement of spending, but also redistribution toward creditors with low spending propensities from debtors with high spending propensities.[19]

The Granger Test Results and Extreme Monetarism

So far we have rejected Temin's arguments for view *(d)* by showing that the observed facts are consistent with a model in which money influences spending. This does not, however, constitute proof that such a model represents an accurate description of the 1929–33 economy. It is still conceivable that the observed facts could have been generated by an economy in which money had no effects on spending and in which the observed correlation between money and income was caused by an entirely endogenous and contemporaneous response of the money supply to bank failures due in turn to the *IS*-induced weakness of spending.[20]

At present the main argument against the extreme position (*d*) is the consensus among reduced-form and structural econometricians that "money matters" in the postwar economy. But there is no reason why the same techniques applied to postwar data cannot be used to analyze interwar data. In a frequently cited study, Sims (1972) developed a method to test the direction of causation between money and income and found that he could reject a reverse-feedback effect of income on money, while he could not reject an impact of lagged money on income.

A related method introduced by Granger (1969) involves regressing Y_t on a constant, a time trend, its own lagged values, and lagged values of X_t:

$$Y_t = \alpha_0 + \alpha_1 t + \sum_{i=1}^{K} \beta_i Y_{t-i} + \sum_{j=1}^{L} \gamma_j Y_{t-j} + u_t. \qquad (1)$$

Now, Y is exogenous with respect to X if the lagged X's fail to make a significant contribution to the explanation of Y over and above the influence of the serial-correlation process in Y captured by the lagged values of Y.[21]

Table 1 displays the results of the estimation of equation (1) and contains two sections, each with four lines. Within each section the four regressions consist of one pair with nominal GNP as dependent variable and M_1 and M_2 alternatively as independent variables, and another pair with the two money

Table 1. Granger Test Results, Quarterly Data, 1920:2–1941:4 and Subperiods

	Variables		*F-Ratio for Significance of Lagged X's*			*t-Ratio on Current X*		
	Dependent (Y)	*Independent (X)*	*1920–41*	*1920–28*	*1929–41*	*1920–41*	*1920–28*	*1929–41*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
A. Levels	Y	M_1	2.68**	1.75	1.50	4.17**	1.62*	4.60**
	M_1	Y	1.10	0.85	0.38	4.17**	1.62*	4.60**
	Y	M_2	3.95**	1.78	1.43	5.57**	0.78	4.95**
	M_2	Y	0.38	0.76	0.43	5.57**	0.78	4.95**
B. Growth rates	Y	M_1	2.21**	2.81**	1.15	3.97**	0.46	3.25**
	M_1	Y	0.63	1.68	0.10	3.97**	0.46	3.25**
	Y	M_2	2.42**	3.06**	1.69	4.85**	0.18	3.71**
	M_2	Y	0.45	1.35	0.42	4.85**	0.18	3.71**

Note: * indicates significant at 10 percent level. ** indicates significant at 5 percent level.

Table 2. Granger Test Results, Monthly Data, 1929–41

	Variables		*F-Ratio for Significance of*	*t-Ratio*
	Dependent (Y) (1)	*Independent (X)* (2)	*Lagged X's* (3)	*on Current X* (4)
A. Levels	IPC	M_2	3.26**	3.20**
	M_2	IPC	1.12	3.20**
	S	M_2	1.17	3.24**
	M_2	S	0.92	3.24**
B. Growth rates	IPC	M_2	3.26**	3.64**
	M_2	IPC	1.40	3.64**
	S	M_2	0.83	2.98**
	M_2	S	1.04	2.98**

Note: ** indicates significant at 5 percent level. All data are seasonally adjusted; S is an index of department store sales published by the Federal Reserve Board; IPC is the Federal Reserve Board's index of industrial production multiplied by the CPI.

concepts as alternative dependent variables. Section A defines each variable in its level form, while section B defines each variable as a one-quarter rate of change. Table 1 reports the results for quarterly data estimated for the period 1920:2–1941:4 and two subperiods.[22] Table 2 reports analogous results for monthly data using M_2 and two proxies for aggregate nominal activity, nominal industrial production, and an index of nominal department store sales.[23]

Turning first to table 1, the first three columns report F-tests on the significance of the lagged independent variables. Lagged nominal income has no significant feedback effect on either M_1 or M_2, though its impact on both approaches significance in growth-rate form in the 1920–28 subperiod. Thus the endogeneity of money, upon which Temin rests much of his argument, is not evident in the form of an impact of lagged income on money in quarterly data for either the 1929–41 subperiod or the complete 1920–41 period. However, a current effect of income on money, as we shall see, is an important feature of these periods.

Lagged money has an ambiguous effect on income. In the level equations (section A of table 1) there is a very significant impact for the overall period but not for either of the subperiods. In rate-of-growth form the significance levels increase substantially for 1920–28 but fall for 1920–41. There is an insignificant impact on income during the 1929–41 period for both M_1 and M_2, just as in the level form of the equations.

The three right-hand columns of table 1 report t-ratios for current values of the independent variables. By far the most important characteristic of the 1929–41 period is the simultaneity of movements in money and income. Schwartz could claim that within the current quarter, money has a very rapid and powerful positive effect on nominal GNP, while Temin could claim that within the current quarter, money is responding passively to changes in GNP caused by nonmonetary factors.

The next step, in the light of simultaneity for the 1929–41 subperiod, is to look inside the contemporaneous quarter by examining results for monthly data in table 2. Once again we find no influence of the lagged income proxies on money, while lagged M_2 does appear to have a significant influence on industrial production (although not on retail sales). Once again there is a strong contemporaneous relation within the current month that could go either way, and so still it is not possible to reject the hypothesis that there is significant feedback from income to money within the current month. In the light of the impact of lagged money on industrial production and the consequent rejection of view (*d*), however, nothing important depends on our inability to untangle the direction of causation within the current month. Certainly, proponents of the middle-ground views (*b*) and (*c*) can feel comfortable with an instantaneous feedback from income to money, described by Friedman and Schwartz as "the reflex influence of business on money, the existence of which is not in doubt" (1963*b*, p. 49).

The Dynamic Simulations and Extreme Monetarism

It is one thing for us to reject the extreme nonmonetarist claim that money did not matter at all, but it is quite another for an extreme monetarist to argue that "only money matters" and that there are "no unexplained changes in spending that serve as deus ex machina" after accounting for a series of "negative shocks, monetary in origin" (Schwartz 1981, pp. 33–34). Similarly, Darby (1976*a*) asserts that the first stage of the contraction was entirely monetary in origin:

> The contraction began, in fact, during the summer of 1929, as the decline in fluidity due to the initial monetary shock slowed and reversed. This early part of the contraction from 1929 to 1930 was in no way different from the sharp recession that would be expected from a 6 percent decrease in the money-supply growth rate. [P. 239]

A possible method to test the Darby-Schwartz proposition about the monetary origin of the contraction is to use the average statistical relation be-

tween lagged money and income during the 1920–28 interval, during which there were three separate recessions, to establish what might be expected to follow a deceleration in the growth rate of the money supply. Can the first year or two of the contraction be attributed in its entirety or in part to the prior monetary deceleration? In this section we report the results of a dynamic simulation in which equation (1) is estimated for the period 1920:2–1928:4 with income as dependent variable and lagged income and money as right-hand variables, and then the predicted behavior of income is calculated based on the fitted coefficients.

A number of possible variants of the dynamic simulation could be presented, corresponding to the different lines in table 1. The monetary definition could be M_1 or M_2, and the variables could be in the form of levels or rates of change. To economize on space, only one version is presented, based on the level form with the M_2 definition of money, because this simulation is most favorable to the hypothesis (*a*) that monetary shocks were solely responsible for the contraction. The M_2 variant in level form is more favorable both because M_2 fell relatively more than M_1 during the contraction and also because the pattern of coefficients during the sample period for the M_2 level variant yields a greater simulated contraction in income.[24]

In figure 3 and table 3 the actual values of nominal income are compared with the values of $\hat{Y}_t$ calculated as

$$\hat{Y}_t = \hat{\alpha}_0 + \hat{\alpha}_1 t + \sum_{i=1}^{4} \hat{\beta}_i \hat{Y}_{t-i} + \sum_{j=1}^{8} \hat{\gamma}_j X_{t-j}. \qquad (2)$$

The "hatted" coefficients are those estimated from equation (1) for the sample period 1920:2–1928:4. The X_{t-j} are the actual values of lagged M_2, and $\hat{Y}_{t-i}$ are the fitted values of the equation when the time period $t-i$ equals 1928:4 or earlier and are the values calculated in equation (2) after 1928:4.

The differences between actual (Y_t) and simulated ($\hat{Y}_t$) nominal income presented in figure 3 must be interpreted carefully. $\hat{Y}_t$ measures the estimated contribution to the behavior of Y_t of the *actual* behavior of lagged M_2, given the structural relation between lagged M_2 and Y_t present in the 1920–28 data. The values of lagged M_2 fed into the dynamic simulations are the actual historical values. To the extent that money was partly endogenous, and the observed decline in M_2 during the contraction partially reflects the contemporaneous influence of nonmonetary factors on income, $\hat{Y}_t$ would tend to exaggerate the contribution of exogenous monetary factors.

On the other hand, monetarists may object to the limitation of the influence of money to a lagged effect. To the extent that the contemporaneous correlation of money and income represents the money-to-income channel of causa-

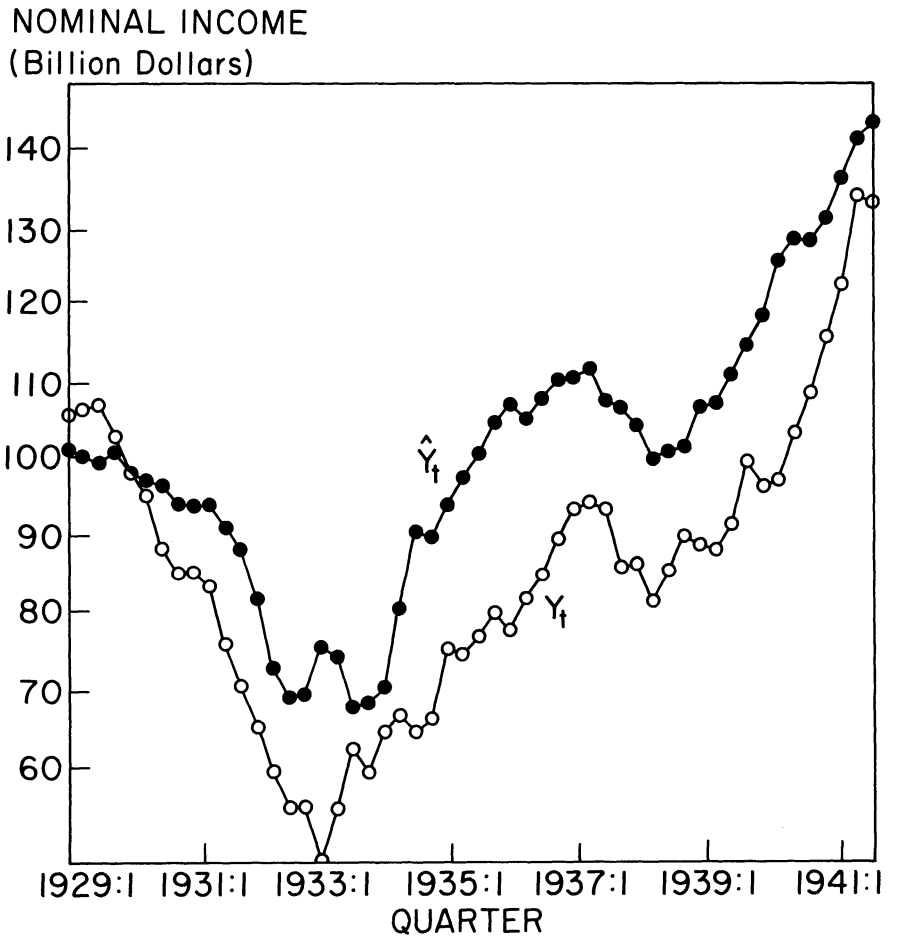

Figure 3. Comparison of Actual and Simulated Nominal Income, 1929–41

Table 3. Summary of Simulation Results on the Role of Lagged Money, 1929–33

	Nominal Income (billion $)							
	Actual Y_t *(1)*	*Simulated with Lagged Money* $\hat{Y}_t$ *(2)*	*Cumulative Change (billion $) from Line A* Y_t *(3)*	$\hat{Y}_t$ *(4)*	*(4)/(3) (5)*	*Marginal Change (billion $) from Line Above* Y_t *(6)*	$\hat{Y}_t$ *(7)*	*(7)/(6) (8)*
A. Peak level (1929:3)	106.0	98.8	——	——	——	——	——	——
B. Half-years								
1. 1929:4–1930:1	100.0	98.8	− 6.0	0.0	0.0	− 6.0	0.0	0.0
2. 1930:2–1930:3	91.3	96.2	−14.7	− 2.6	17.7	− 8.7	− 2.6	+ 29.9
3. 1930:4–1931:1	84.9	93.5	−21.1	− 5.3	25.1	− 6.4	− 2.7	+ 42.2
4. 1931:2–1931:3	79.5	92.0	−26.5	− 6.8	25.7	− 5.4	− 1.5	+ 27.8
5. 1931:4–1932:1	67.9	84.6	−38.1	−14.2	37.3	−11.6	− 7.4	+ 63.8
6. 1932:2–1932:3	57.2	70.9	−48.8	−27.9	57.2	−10.7	−13.7	+128.0
7. 1932:4–1933:1	51.8	72.4	−54.2	−26.4	48.7	− 5.4	+ 1.5	− 27.8
8. 1933:2–1933:3	58.5	71.0	−47.5	−27.8	58.5	+ 6.7	− 1.4	− 20.9

Source: Figure 3.

tion, the exclusion of the current money supply understates the contribution of monetary change. But the addition of current money, somewhat surprisingly, actually dampens the 1929–33 decline in the simulated income series, because the coefficient on current money in the 1920–28 income regression is a small and insignificant *negative* number.

How well does the lagged-money simulation explain the Great Contraction? Figure 3 indicates that $\hat{Y}_t$ consistently lies above the actual value of nominal GNP (Y_t). Dividing up the 1929–33 contraction into two-quarter intervals to facilitate analysis, we can examine the averages presented in table 3. According to the Hicks theory of the "double slump," we should find that the simulation based on lagged money ($\hat{Y}_t$) explains only a portion of the actual slump in Y_t during the first two years of the contraction but that monetary forces then take over and account for most of the decline in Y_t. According to Friedman-Schwartz, the contraction changed its character one year earlier, at the time of the first wave of bank failures during the last quarter of 1930. According to Darby-Schwartz, the $\hat{Y}_t$ series should trace the 1929–31 decline in Y_t quite closely.

Both the contribution of money to the cumulative change in $\hat{Y}_t$ in column (5) and the contribution to the marginal change from one half-year to the next in column (9) are more consistent with the Hicks timing than with the Friedman-Schwartz timing and are not consistent at all with the Darby-Schwartz money-only explanation. Between line A and line B2, money contributes 17.7 percent of the total decline in income; on line B4 the cumulative contribution rises only to 25.7 percent. And the marginal contribution on line B4 is only 27.8 percent. In contrast, there is a dramatic change beginning on line B5, where the marginal contribution of money jumps to 63.8 percent, and to more than 100 percent on line B6. Although a subsequent zigzag causes the simulated Y_t series to miss the timing of the last stage of the contraction in late 1932 and early 1933, the cumulative contribution of $\hat{Y}_t$ to the actual decline in Y_t nevertheless remains in the vicinity of 50 percent in lines B6 through B8.

As we have seen, both Darby and Schwartz have pointed to slow monetary growth in 1928 and early 1929 as the fundamental underlying cause of the first year of the contraction. Indeed, between 1928:1 and 1929:3, M_2 grew by only 0.6 percent at an annual rate in contrast to a rate of 5.2 percent in the preceding five quarters. But even greater decelerations of monetary growth had happened before without causing a drastic drop in nominal income. For instance, while the growth of M_2 slowed from an annual rate of 8.8 percent in the seven quarters preceding 1925:4 to a 0.5 rate in the next four quarters, the subsequent decline in nominal income between peak and trough in the 1927 recession was only 2.8 percent. Thus the simulated value $\hat{Y}_t$, which combines the average relation between lagged money and income observed during the 1920s

with the actual behavior of money in 1929–33, essentially says, "Though monetary growth decelerated in 1928 and 1929, such a monetary slowdown had happened before and can only account for 18 percent of the observed decline in nominal income in the first year of the contraction and 26 percent cumulatively in the first two years."

The 1937–38 Recession and Subsequent Recovery

Monetarist interpretations of the Great Depression are not limited to the 1929–33 contraction phase. In addition, monetarists have long taken the position that the proximate cause of the 1937–38 recession was the three-stage doubling of reserve requirements between August 1936 and May 1937. The same simulation technique can be used to evaluate the validity of this claim. The technique is exactly the same as in the preceding discussion, except that two different simulation results are reported. The first is based on the money-income equation fitted to the 1920–28 period that is used in the simulations in figure 3 and table 3. As is evident in table 4 and figure 4, the value of $\hat{Y}_t$ calculated from the dynamic simulation that starts in 1929:1 remains above the actual value of Y_t throughout the 1937–41 period. Nevertheless, $\hat{Y}_t$ declines between the peak quarter (1937:2) and early 1938 by almost as much as actual income. In short, the simulation based on the 1920–28 coefficients implies that the 1937–38 recession was almost entirely a monetary phenomenon.

A second simulation is based on the same specification extended to the longer 1920:2–1936:4 sample period. The results of the 1937–38 recession confirm the verdict that the simulated $\hat{\hat{Y}}_t$ series explains most of the downturn in Y_t—68 percent in this case as compared to 91 percent for the first simulation.

Although the simulated series $\hat{Y}_t$ and $\hat{\hat{Y}}_t$ indicate that most of the 1937–38 recession can be explained as a consequence of the behavior of lagged money and lagged income, the ability of the two simulated series to track actual income nevertheless deteriorates markedly after early 1938. As indicated in both table 4 and figure 4, the simulated series recover much more markedly than actual Y_t between the first half of 1938 and the first half of 1940. In the latter interval, actual nominal income had exceeded the 1937 peak by only $2.4 billion, or 2.6 percent. But the $\hat{Y}_t$ series had grown by 9.1 percent and $\hat{\hat{Y}}_t$ by 16.8 percent.

After the first half of 1940 the relationship between the actual and simulated series shifted in the direction of rapid actual growth relative to simulated growth. Only about half of the actual growth in nominal income between the first half of 1940 and the last half of 1941 can be explained by the growth of

Table 4. Summary of Simulation Results on the Role of Lagged Money, 1937–41

	Nominal Income (billion $)							
	Actual Y_t *(1)*	*Simulated with Lagged Money* $\hat{Y}_t$, $\hat{\hat{Y}}_t$ *(2)*	*Cumulative Change (billion $) from Peak* Y_t *(3)*	$\hat{Y}_t$, $\hat{\hat{Y}}_t$ *(4)*	*(4)/(3)* *(5)*	*Marginal Change (billion $) from Line Above* Y_t *(6)*	$\hat{Y}_t$, $\hat{\hat{Y}}_t$ *(7)*	*(7)/(6)* *(8)*
A. Simulation based on 1920–28 ($\hat{Y}$)								
1. Peak, 1937:2	93.7	110.5	——	——	——	——	——	——
2. 1938:1–1938:2	83.5	101.2	−10.2	− 9.3	91.1	−10.2	− 9.3	91.1
3. 1940:1–1940:2	96.1	120.6	+ 2.4	+10.1	420.8	+12.6	+19.4	154.0
4. 1941:3–1941:4	131.7	140.3	+38.0	+29.8	78.4	+35.6	+19.7	55.3
B. Simulation based on 1920–36 ($\hat{\hat{Y}}$)								
1. Peak, 1937:2	93.7	88.8	——	——	——	——	——	——
2. 1938:1–1938:2	83.5	81.9	−10.2	− 6.9	67.6	−10.2	− 6.9	67.6
3. 1940:1–1940:2	96.1	103.7	+ 2.4	+14.9	620.8	+12.6	+21.8	173.0
4. 1941:3–1941:4	131.7	121.4	+38.0	+32.6	85.8	+35.6	+17.7	49.7

Source: Figure 4.

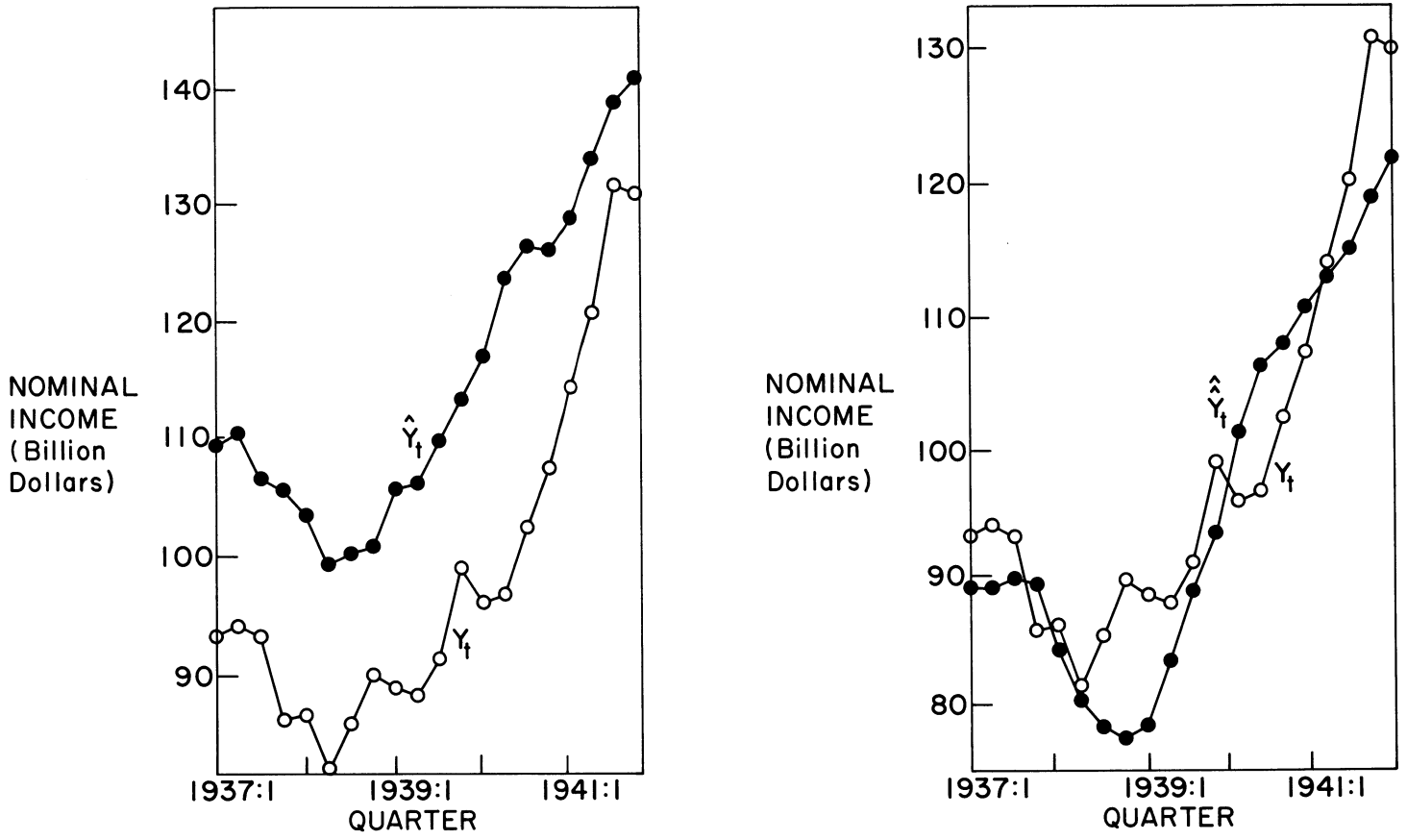

Figure 4. Comparison of Actual and Simulated Nominal Income, 1937–41

lagged money and income. These results appear consistent with a model of income determination in which shifts in private investment and government expenditures play an important role, given the behavior of money. Private investment was sluggish during the interval 1937–40, leading to a weak recovery despite the rapid growth in M_2 that was occurring.[25] Then, after mid-1940, rapid growth in government defense spending shifted the *IS* curve rapidly rightward and caused an accelerated growth in income without any acceleration in the growth rate of money.[26]

Our interpretation is that shifts in the *IS* curve must be relied upon to explain the timing of income growth in the 1938–41 period, just as *IS* shifts appear to have dominated the explanation of income change in the first two years of the Great Contraction, 1929:3–1931:3. This suggests a puzzle. 1Given the weak impetus to spending provided by the monetary acceleration of 1938–40, why should the monetary deceleration of early 1937 have been so potent? One answer is that monetary tightness *per se* was not particularly potent, and instead the 1937–38 recession was due at least partly to nonmonetary factors. One candidate that stands out is the increase between 1936 and 1937 in the full-employment federal surplus equal to fully 3 percent of GNP (equivalent to a $60 billion fiscal swing in today's economy).[27]

Implications of the Regressions and Simulations

Several additional questions can be raised concerning the regression and simulation results. First, do the 1929–33 and 1937–38 downturns in the simulated series $\hat{Y}_t$ and $\hat{\hat{Y}}_t$ reflect just the lagged effect of the decline in money, or is part of the decline contributed by the lagged-income variables? We have examined separate simulations based on regressions in which lagged values of money are excluded, in order to study the postsample predictions based solely on the autoregressive structure of the income variable. There is a minor cycle in the *growth rate* of income in such a dynamic autoregressive simulation but no actual decline in the level of income during the 1929–33 or 1937–38 periods. Thus it appears that all of the decline in the simulated series in figures 3 and 4 is being contributed by the lagged effect of money and none by the lagged-income variables.

Second, is the failure of the simulated series to capture fully the actual 1929–33 decline in income in figure 3 due in any part to the inclusion of a time trend in the original regression equation (1)? The results of alternative simulations based on regressions without time trends can be summarized by showing the contribution of the simulated series to the actual cumulative change by the two alternative methods of estimation:

	With Trend	*Without Trend*
Table 3, peak to 1931:2/1931:3	25.7%	31.3%
Table 3, peak to 1932:4/1933:1	48.7	57.6
Table 4, peak to 1938:1/1938:2 (Line A)	91.1	95.1
Table 4, peak to 1938:1/1938:2 (Line B)	67.6	22.5

Thus the omission of the time-trend variable does increase by a minor amount the contribution of lagged money to an explanation of the 1929–33 and 1937–38 contractions based on the 1920–28 regressions where the time trend is positive. The same omission, however, substantially reduces the contribution of lagged money to an explanation of the 1937–38 contraction based on the 1920–36 regressions where the time trend is negative.

Third, why is there such a difference in the contribution of lagged money to an explanation of the 1937–38 contraction between the two sets of simulations based on the alternative 1920–28 and 1920–36 sample periods? There are very substantial shifts in the coefficients of these reduced-form regression equations when the sample period is altered. Table 5 exhibits the shifts in coefficients on lagged and current money in alternative overlapping eight-year sample periods. There appears to be an inverse correlation between the sum of coefficients on lagged money in column (1), which is greatest in the first three lines, and the coefficient on current money in column (3), which is much larger in the last four lines than in the first three. Thus the relation between money and income appears to have shifted to a mainly contemporaneous one in the 1930s, with a substantial lagged effect of money on income evident only in the earlier periods.

The results in table 5 cast additional doubt on the hypothesis that changes in the money supply were primarily responsible for the behavior of income in the Great Depression. In all of the subperiods in table 5 the t-ratio on the sum of lagged coefficients is extremely small. Although some individual coefficients are significant, they tend to alternate in sign. The dominance of the contemporaneous correlation in the decade of the 1930s adds plausibility to the reverse-feedback hypothesis that the reflex effect of business on money was a primary determinant of shifts in the money supply. Further, it is awkward for monetarists to rely upon an entirely contemporaneous money-to-income effect to support their case, because long lags between policy changes and income changes play an important part in their argument against countercyclical activism in Plank 3 of the monetarist platform.[28]

It is important, however, to distinguish hypothesis (1), that observed movements in the money supply during the 1930s were largely passive and endogenous, from hypothesis (2), that an alternative monetary policy that substituted

Table 5. Effect of Lagged and Current Money on Income in Alternative Eight-Year Sample Periods, Quarterly Data, 1920-40

Sample Periods	Lagged Money		Current Money	
	Sum of Coefficients (1)	*t-Ratio of Sum (2)*	*Coefficient (3)*	*t-Ratio (4)*
1. 1920:2–1928:4	0.936	0.15	−1.118	−0.77
2. 1922:1–1930:4	1.405	0.27	+0.289	+0.21
3. 1924:1–1932:4	1.249	0.19	+0.871	+1.16
4. 1926:1–1934:4	0.262	0.05	+2.357	+2.77
5. 1928:1–1936:4	0.628	0.12	+3.083	+3.47
6. 1930:1–1938:4	0.413	0.07	+3.938	+6.05
7. 1932:1–1940:4	0.595	0.12	+2.748	+3.11

active countercyclical open-market operations could have lessened the severity of the contraction and brought about an earlier and more robust recovery. Both hypotheses (1) and (2) could be correct, but hypothesis (2) cannot be tested on data from the period if hypothesis (1) is correct as well. Coefficients would have shifted, as the work of Robert Lucas (1976) suggests, if an activist monetary policy had been pursued. For this reason, econometric studies of U.S. money and income data are unlikely to settle the debate regarding the potential role of alternative monetary policies, however much they may indicate that the lagged changes in monetary growth that actually did occur are capable of explaining little if any of the fluctuations in income in the decade of the 1930s.

A comparison of the United States with Europe, where both money and income followed quite different paths after the devaluation of sterling in the fall of 1931, helps to overcome the inherent limitations of the U.S. data. In figure 5 the European data exhibit a dramatic divergence from the U.S. behavior of money and income after 1931.[29]

Some of this difference may represent nonmonetary factors that raised income and pulled up the money supply through a reverse-feedback mechanism—for example, the stimulus of the 1931 devaluation in several European countries and the impact on income of activist fiscal stimuli (especially in Germany beginning in 1933). But some of the explanation for the earlier European recovery may rest with activist monetary policy, as described for Sweden by Jonung (1981), lending some credence to Schwartz's (1981) statement that "different policies would have resulted in different behavior."

The comparison between European and U.S. velocity in the bottom frame of figure 5 reveals some interesting similarities and differences. The simultane-

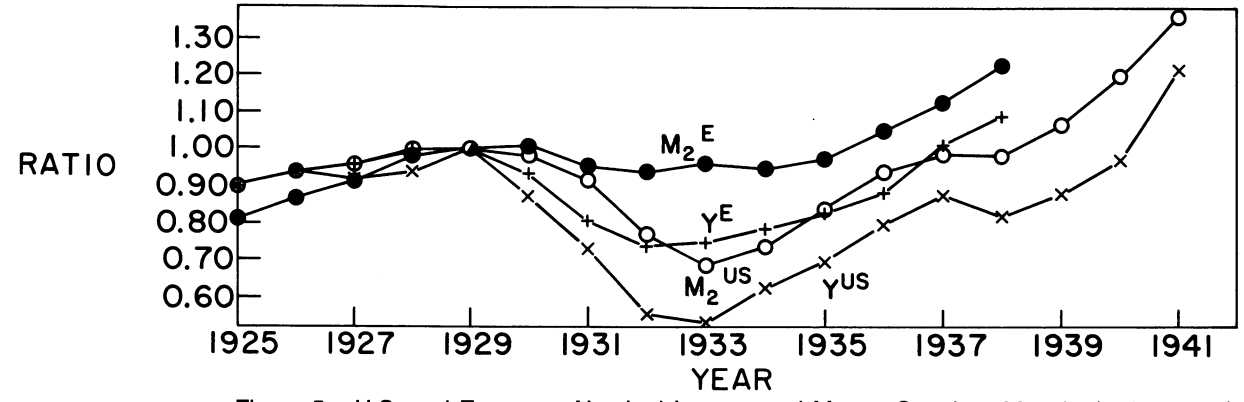

Figure 5*a*. U.S. and European Nominal Income and Money Supply, 1925–41 (1929 = 1.0)

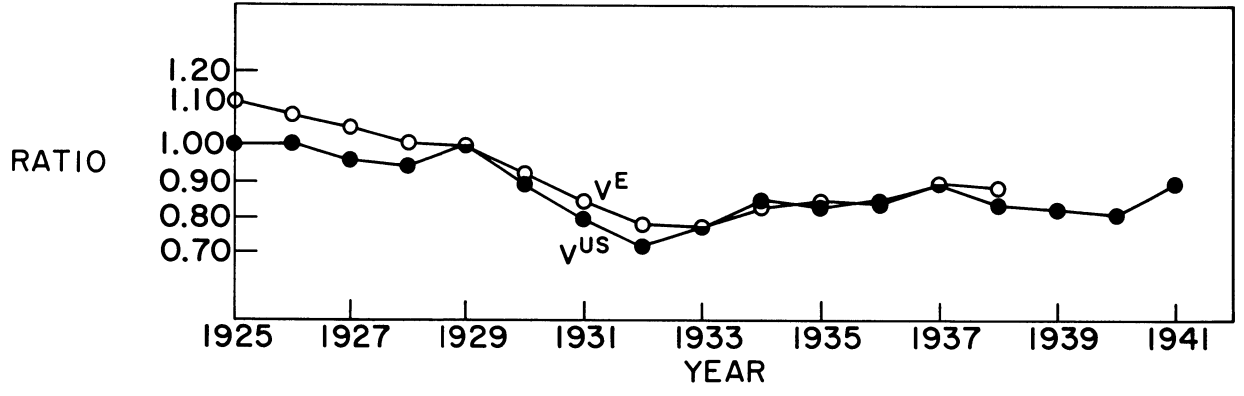

Figure 5*b*. U.S. and European Velocity, 1925–41 (1929 = 1.0)

ous sharp decline in 1929–32 suggests the presence of a common nonmonetary shift factor. The fact that European velocity declined less than that in the United States is not consistent with the predictions of an ordinary *IS–LM* model, given the less restrictive monetary policy pursued in Europe. Finally, the decline in velocity in the United States in 1939 and 1940, and in Europe in 1938, is consistent with the hypothesis that the *LM* curve is positively sloped but does not constitute a demonstration that it is horizontal. In short, the comparison in figure 5 is consistent with the basic themes of this paper that both monetary and nonmonetary factors mattered, that nonmonetary factors were of prime importance in 1929–31, that different monetary policies in the United States after 1931 would have reduced the severity of the contraction, and finally that the stimulus of rapid monetary growth on economic activity in the late 1930s was quite weak.

THE CONTRIBUTION OF NONMONETARY FACTORS

Searching for Nonmonetary Explanations

It is easier to show that nonmonetary factors must have been at work in the first two years of the contraction than to determine what those factors were, much less to assign specific quantitative contributions to each of them. Most recent discussions of nonmonetary factors have suffered from three methodological weaknesses, which we may label "monocausal blinders," the "endogenicity fallacy," and "postwar second-guessing." An analyst wears monocausal blinders when he claims that Factor X could not have possibly been *the* cause of the contraction because it was not quantitatively important enough, or it happened at the wrong time. For instance, "the fall of construction in the late 1920s was deflationary, but too small to precipitate a major depression" (Temin 1976, p. 68). "Construction spending peaked in 1926," some might say, "so if construction spending was the cause, why did the Great Depression not begin in 1927?" This view ignores the possibility that there may have been several causes and timing patterns. Imagine that there were two causes, X_1 (construction), which continuously exerted a downward influence on nominal spending after 1926, and X_2 (say the stock market), which exerted an expansionary influence in 1928 and early 1929 and a contractionary influence after mid-1929. The fact that the economy slumped in late 1929 rather than early 1927 does not deny that cause X_1 made the post-1929 contraction more severe than otherwise. In the same way, table 3 indicates that in the early quarters of 1930, the decline in the money supply probably made the contraction more severe, even if money *alone* can account for only a frac-

tion of the total decline in spending. As the decline in Y_t fed into investment planning, it in turn made investment decline more rapidly than otherwise.

An analyst suffers from the "endogenicity fallacy" when he dismisses X_1 as a possible cause because it is "endogenous and declined because income declined." For instance, Temin writes that "the major part of the fall in construction in the 1930's can best be seen as the result of the fall in income rather than as the result of a change in some alternative variable" (1976, p. 66). Thus, the construction hypothesis is abruptly dismissed. Similarly, money is dismissed as a cause by Temin because the bank failures that were primarily responsible for the decline in the stock of money are viewed as an endogenous response to the decline in income. But this cavalier approach neglects the possibility of interactions among a number of possible causes, an interaction of which Friedman and Schwartz were well aware when they wrote that the endogenicity of money was "part of the partly self-generating mechanism whereby monetary disturbances are transmitted" (1963*b*, p. 50). In the same way, the endogenicity of construction—which is influenced by income, on which the desired stock of structures depends, but which at the same time is part of GNP—is part of the "partly self-generating mechanism whereby *non*monetary disturbances are transmitted."

Finally, "postwar second-guessing" occurs when an analyst claims that X_1 could not have been a cause of the 1929–33 contraction because slumps in X_1 have been observed to occur in the postwar years without resulting in a Great Depression. This ignores possible differences in factors other than X_1 that may have served to insulate the economy from the effects of the X_1 slump. More formally, this point may be made in terms of the national-income identity:

$$S \equiv I + D + F, \tag{3}$$

where S is gross saving, I is gross investment, D is the government deficit, and F is the foreign trade surplus. Dividing both sides by "full-employment" or "natural-employment" output (Q^*), and designating the ratio of gross saving to actual output (Q) as s ($=S/Q$), we have:

$$\frac{sQ}{Q^*} \equiv \frac{I}{Q^*} + \frac{D}{Q^*} + \frac{F}{Q^*}. \tag{4}$$

Imagine, for the sake of argument, that the gross saving ratio, s, is roughly fixed. If there is a decline in the ratio of gross domestic investment to natural output (I/Q^*), then the economy must adjust in some way, either by an offsetting shift in the natural-output government deficit (D/Q^*) or in the natural-output trade surplus (F/Q^*) on the right-hand side of the equation or by a contraction in the output ratio (Q/Q^*) on the left-hand side.

In the postwar period there have been repeated multiyear booms and slumps in the I/Q^* ratio. To some extent these have been offset by the willingness of the federal government to incur budget deficits; nevertheless, their impact has not been entirely offset, and the economy has adjusted by experiencing prolonged periods of persistent above-average or below-average unemployment.[30] The absence of a depression in the postwar period testifies to (1) potent built-in stabilizers and (2) monetary policy, which aggravated cycles in the short run by allowing procyclical slumps in money at the beginning of recessions but nevertheless managed to get money growing again after a quarter or two. It is not inconsistent to say that the actual behavior of investment in 1929-31 would have caused only a 1958- or 1975-size recession had it been accompanied by postwar-size built-in stabilizers and postwar monetary policy and at the same time to say that the behavior of investment caused a much more serious contraction in the 1930s, given the smaller built-in stabilizers and the endogenous procyclical monetary policy conducted by the Fed.[31] Put another way, one can simultaneously claim that the contraction was a nonmonetary phenomenon in origin and that it was monetary in the sense that the actual monetary policy aggravated the slump and an alternative expansionary monetary policy would have moderated it. One can agree with Schwartz that "different policies would have resulted in different behavior" and simultaneously disagree with her statement that "there are no unexplained changes in spending that serve as deus ex machina"(1981, p. 26).

In searching for the nature of the nonmonetary deus ex machina, we do not imply that money did not play an important role, particularly in the 1931–33 phase of the contraction. But we reject the contention that there is only one "main question" to be answered about the contraction—why it was so severe and "why recovery was so slow in coming" (Mayer 1978*b*, p. 130.) An episode as dramatic as the contraction is capable of raising more than a single issue. While monetarists may be content to limit their analysis to a demonstration that inept monetary policy explains the unique magnitude of the contraction, we find equally interesting a search for nonmonetary forces that appear to have been primarily responsible for the 28 percent decline in nominal income in the interval 1929:3–1931:3 and that in turn must have played at least some role in causing the bank failures that the Fed failed to counteract.[32]

Just as there can be more than a single "main question" of interest suggested by the 1929–33 experience, so there may have been more than a single nonmonetary explanation of the severity of the decline in income during the 1929–31 phase. Several possible explanations share the common theme that any excess of spending breeds its own self-correcting contraction. Many authors have constructed business-cycle models based on the interaction of the multiplier and the accelerator. In Goodwin's model the expansion phase is

eventually terminated by supply constraints, which slow the growth of the capital stock and hence the level of net investment. The economy "is always straining to get to the full employment limit, but by the mere fact of being there for a time, it is projected downward again" (1955, p. 209).

Our explanation can be summarized within the flexible-accelerator framework as follows:

1. Net investment in both consumer and producer goods is a function of the deviation between the desired and actual stocks of those goods.
2. A decline in net investment can occur when there is a decline in the desired stock or when something has occurred in the past to raise the current stock too high relative to today's desired stock.
3. Within the framework of the identity (4) above, any such decline in net investment will cause a decline in the output ratio (Q/Q^*) unless offset by a decline in the saving ratio, the natural-employment government deficit, or the natural-employment trade surplus.
4. The major factor that reduced the desired capital stock was the effect of declining population growth on residential housing.
5. The major factors that raised the actual capital stock too high were the overbuilding of residential housing in the mid-1920s and the effect on consumer spending of the overshooting of the stock market during its 1928–29 speculative bubble.

Construction

In a recent paper Hickman (1973) has documented both the effect of the decline in population growth on the desired housing stock and also the extent of overbuilding in the mid-1920s. Hickman's model of the residential housing sector improves on previous work by treating the rate of population growth as endogenous, due to the effect of income on the rate at which individuals in various age groups choose to form households. Hickman is able to decompose the observed decline in the rate of population growth between the early 1920s and mid-1930s into two components—that due to the effect of declining income, and a remaining exogenous decline in "standardized households" due primarily to the decline of immigration.[33]

In order to isolate the effect of the exogenous component of the decline in household formation, Hickman calculated two dynamic simulations of his model, one in which standardized households are assumed to increase steadily at the 1924–25 rate of growth, and another in which income and other economic variables are identical but in which standardized households follow their

actual declining growth path after 1925. The impact of the actual demographic slump gradually becomes more important as the 1930s progress, accounting for a decline in housing starts between the two simulations of 28.3 percent for the year 1933 and 39.1 percent for the year 1940.[34]

It has been suggested that the effect of declining immigration on the desired capital stock of residential housing could not be a contributing factor to a worldwide depression. Such a change in immigration patterns, the argument runs, would reduce the demand for housing in the United States but raise the demand for housing in the former source countries, for example, Italy and Poland. But this position is flawed for several reasons. First, many of the immigrants came from rural areas where their departure led to housing abandonment. A lower immigration flow would reduce the demand for housing in the United States but would to a large extent reduce the rate of abandonment in Italy and Poland rather than stimulate new construction. Second, the marginal product of U.S. immigrants instantly increased upon arrival as compared to their previous situation because of the much greater amount of physical capital available in the United States. Third, there is a long oral tradition in labor economics which claims that increases in immigration led to an expansion in the demand for all types of reproducible capital goods, not just residential housing. Immigrants initially hold a proportionately greater share of their nonhuman wealth in liquid capital, particularly gold and jewelry, but after some period of adjustment to their new environment, this liquid wealth is converted into physical capital.[35]

But the deflationary impact of demography was only the first of the two important causes of the housing problem. The second was the extent of overbuilding in the mid-1920s. For six years (1923–28) real residential construction achieved a level more than double the average of the entire decade before World War I. In four successive years (1924–27) the ratio of real residential construction to real GNP reached by far its highest level of the twentieth century.[36] Hickman's simulations dramatize the extent to which housing starts had risen in 1925 to a rate higher than was consistent with current income, prices, and the rate of household formation. In the most optimistic of his simulations—that which assumes that standardized household growth *continues* at its 1924–25 rate, rather than declining, and that there is no decline in income—predicted housing starts still fall by 35 percent between 1925 and 1930.

Combining the two effects, how much could housing have contributed to the decline in income in the Great Contraction? Hickman's simulation that holds income constant but allows standardized households to follow their actual growth path generates a decline in housing starts between 1925 and 1930 of 49 percent, amounting to about 4 percent of 1925 GNP. The impact of this deflationary force on the economy was delayed by the buoyant behavior of consumption and inventory accumulation in 1929, but when these components

Table 6. Ratios of Real Spending Components to Natural Real Output in 1926, 1929, and 1930

	1926	*1929*	*1930*	*Change, 1926–30*
Consumption expenditures	66.4	68.1	61.8	− 4.6
Nonresidential fixed investment	13.0	12.9	10.3	− 2.7
Residential fixed investment	8.6	5.1	3.0	− 5.6
Other	12.3	13.2	11.9	− 0.4
Total	100.3	99.3	87.0	−13.3

Sources: The natural output series (Q^*) is from Gordon 1978, Appendix B. The 1926 spending components are from Hickman and Coen 1976, table A.2, p. 222. The 1929 and 1930 spending data are in 1958 dollars, to retain comparability with the Hickman and Coen data, from the *Economic Report of the President* 1968.

of spending collapsed in 1930, the downward pressure on income from the housing sector interacted to aggravate the severity of the contraction. Table 6 displays the ratios to real natural output of the major components of real spending in 1926, 1929, and 1930.

Consumption Expenditures

Table 5 indicates that several components of spending declined sharply between 1929 and 1930, with the decline in consumption contributing the most to the decline in real GNP. The behavior of consumption spending partly represents an endogenous reaction to the decline in other components of spending, but in addition some portion of the consumption decline may reflect the influence of the stock market crash or may be an unexplained autonomous puzzle. Unfortunately, the recent debate between Temin (1976) and Mayer (1978*a*) does little to elucidate the role of the stock market in explaining consumption. By focusing on the significance in 1930 of *residuals* from consumption equations, both Temin and Mayer neglect to calculate the contribution of changes in stock market wealth to the *fitted* value of consumption. Yet the timing of the stock market boom and crash must partly explain why the level of consumption spending was so high in 1929 and so low in 1930.

Taking the position that data inadequacies preclude estimation of an interwar macroeconometric model, Mishkin (1978) has used coefficients from a postwar model to assess the impact on consumption expenditures and residential housing of changes in the household balance sheet in the 1930–41 period. Wealth effects are potent enough to explain 45 percent of the decline in these spending components in 1929–30. This is probably an overstatement of the true impact of the exogenous component of the financial developments, because the endogenous response of the financial variables to the decline in income is neglected. When the Hickman housing simulations and Mishkin calculations are combined, we emerge with an explanation of several crucial features of the depression.

First, the housing collapse helps to explain both why the contraction was so severe and why it lasted so long. As late as 1940 the ratio of housing to natural output had not regained even half of its level of the mid-1920s. Real GNP in 1940 was able to exceed its absolute 1929 level through the contribution of government pump-priming that filled in the gap left by the missing investment.[37] Given the fact that M_2 had risen 18 percent in 1940 relative to 1929 and that M_1 had risen by 49 percent, a purely monetary approach cannot provide an explanation of the duration of the depression.

Second, the relation between the stock market and consumption spending helps to explain why the initial 1927–29 collapse of construction did not initiate the depression earlier; the 1928–29 stock market bubble induced a consumption boom that postponed the impact of the housing slump. The stock market collapse precipitated a drastic decline in consumption spending that interacted with and further aggravated the continuing decline in residential construction. This interpretation makes the behavior of consumption at least partly a monetary phenomenon, to the extent that easy money helped boost stock prices and that tight money helped bring on the crash. But no one has ever claimed that the tripling of stock prices between the business-cycle peaks of 1923 and 1929 could be more than partly explained by the 27 percent increase in M_2 over the same interval; a large residual portion of the behavior of stock market prices must be classified as due to a speculative bubble that at some point had to burst. In this sense the behavior of the stock market and its impact on consumption can be termed both autonomous and essentially nonmonetary in origin, even if the precise timing of the stock market crash may depend in part on the timing of monetary policy.

International Interactions

Meltzer has argued that American adherence to the rules of the gold-exchange standard was a factor contributing to the initial decline in spending in 1929.

"A recession can be induced by the changes in [international] relative prices that occurred in 1928 and 1929. A recession induced by changes of this kind is a response to monetary policy if we include in monetary policy a commitment to operate under the rules of the gold standard" (1976, p. 458).

To the extent that the money stock is endogenous and responds negatively to relative output advances, Meltzer has identified a little-noticed monetary influence. Consider a domestic monetary expansion. As domestic output and prices advance relative to output and prices abroad, net exports decline, tempering and possibly reversing the rise in output. Under the gold-exchange-standard rules, the decline in net exports would also result in a gold outflow and subsequent fall in the money supply.

The evidence in favor of adherence to the principles of gold standard during this period is extremely weak, however. Actually, U.S. policy through most of the 1920s was to sterilize gold flows. "From 1923 on, gold movements were largely offset by movements in Federal Reserve credit so that there was essentially *no* relation between the movements in gold and in the total of high powered money; the fairly irregular dips and rises in the gold stock were transformed into a horizontal movement in total high powered money" (Friedman and Schwartz 1963*a*, p. 382).

This policy of sterilization eliminates the link between gold and the money supply central to Meltzer's hypothesis. Further refutation of Meltzer's thesis that monetary policy should have been expected to lead to a recession after the 1927–28 recovery lies in the fact that the ratio of export to import prices in the past had not always risen when U.S. output advanced relative to that of its trading partners. Though an increase in U.S. relative to world income was accompanied by an increase in U.S. relative prices in 1928–29, this was not the case in 1922–23, when the price of U.S. exports fell over 9 percent relative to the price of imports and U.S. output advanced relatively.[38]

Even with the deterioration of U.S. relative international prices at the end of the 1920s, a demonstration of the impact of these price changes on net exports is problematical. Hickman and Coen (1976) attempt to capture the effect of relative price on imports over this era but cannot uncover any significant effect when income and other factors are allowed for. More recently, Artus and Sosa attempt to estimate these price elasticities for the 1963–74 period, concluding that these elasticities "are not extremely large and are felt rather slowly" (1978, p. 46). In addition, real net exports barely changed between 1929 and 1930. Exports and imports declined together. If a relative price change were responsible for causing the United States to export less and import more, a deterioration in the trade balance should be observed. The absence of any change in the real trade balance is an indication that some other factor or combination of factors, both monetary and nonmonetary, was responsible for the simultaneous reduction in income of the United States and its

trading partners, which in turn caused both exports and imports to decline together. The fact that European nominal income fell less than that in the United States, as indicated in figure 5, is consistent with the hypothesis that the depression spread from the United States to Europe but does not support any particular hypothesis about the effect of relative prices on the trade balance.[39]

The most important qualification of all to Meltzer's hypothesis is that a change in relative prices of a particular nation should have caused expenditure *switching,* not a worldwide depression. European output should have been stimulated and U.S. output depressed, with aggregate world output left unaffected.[40] The data show that income and output on both sides of the Atlantic fell together, a pattern consistent with causation from another factor.

Meltzer is on firmer ground when he blames another international factor, the Hawley-Smoot Tariff Act of June 1930, as responsible for converting "a sizeable recession into a severe depression" (p. 469). The tariff was responsible for an increase of almost 50 percent in the effective rate of duties paid on imports between 1929 and 1932. This aggravated the contraction through three main channels:

1. Directly, without any retaliation, the resulting increase in the price of U.S. imports and close domestic substitutes altered the division of the nominal-income decline between output and prices in 1930–32, so that output fell more than otherwise and prices fell less.
2. Foreign retaliation reduced the demand for U.S. exports, which aggravated the contraction through the standard Keynesian multiplier mechanism.
3. Foreign retaliation against U.S. exports of food products, which dropped 66 percent between 1929 and 1932 (Meltzer 1976, p. 460), aggravated the decline in U.S. farm prices, which was an important cause of rural bank failures and in turn of the decline in the supply of money due to currency hoarding.

Whether the impact on output and unemployment of the Hawley-Smoot tariff was more or less important than that of housing and the stock market is probably impossible to determine. The important point is that there was more than one source, not just the behavior of the money supply but also several nonmonetary factors, and that their effects interacted and amplified the severity of the contraction. The role of the tariff, while not explicitly involving the money supply, is nonetheless a factor that is consistent with Plank 3 of the monetarist platform with its emphasis on the harmful effect of government intervention, of which the tariff is a classic example.

Other international factors caused differences in the timing and magnitude of the contraction in individual countries. The end of capital outflows from the

United States to Germany in 1928 helped cause an early downturn in that country. In addition, reparations "greatly intensified the German depression" (Haberler 1976, p. 29). Then devaluations by Britain, Scandinavia, and other countries in September 1931 stimulated early recoveries there while deepening the slide in the United States and Germany. The devaluation of the dollar in 1933–34 caused a late trough in France and some other countries that had not devalued earlier.

PRICES, OUTPUT, AND AGGREGATE SUPPLY

Equilibrium and Disequilibrium Approaches

Monetarists tend to rely on equilibrium aggregate supply (EAS) theories to explain the division of nominal income between prices and quantities. These theories—embodied, for example, in the work of Friedman (1968), Lucas (1973), and Sargent (1976)—view changes in the actual relative to the "natural" rate of output as the response to deviations of actual from expected prices, which cause a divergence of economic agents' expected and actual real wages.

Nonmonetarists, on the other hand, tend to discuss the same issue of price and quantity determination in terms of a disequilibrium-adjustment framework. In response to a demand shock, prices do not typically adjust rapidly enough to clear markets, so agents find themselves constrained by a level of sales or employment different from what they would voluntarily choose to demand or supply at prevailing wages and prices (see Barro and Grossman 1976, chap. 2). Under these circumstances, the demand for labor becomes a function not only of real wage, the capital stock, and technology—as in EAS theory—but also of actual or expected output or sales. Nonmonetarists do not claim that wages and prices are completely rigid but rather assert that in the short run wage and price adjustment to a situation of excess supply or demand is partial rather than complete.

Empirical Explanations of Unemployment and the Output Ratio

Empirical tests of the EAS approach have been carried out by Lucas and Rapping (1969) and more recently by Darby (1976*b*). Though Lucas and Rapping examined the period from 1930 through 1965 and concluded that their model was "consistent with the U.S. experience," in the ensuing debate with Rees (1972), they admitted that their approach could not account for the failure of the unemployment rate to decline more rapidly after 1933.

In an attempt to resuscitate the EAS explanation of the 1930s, Darby has presented new unemployment data that treat government workers on CCC and WPA projects as employed rather than unemployed. Darby's corrected data are claimed to exhibit a strong movement toward the natural rate of unemployment in the post-1933 period.[41] Darby expresses the actual unemployment rate at time t, U_t, as a function of a constant, representing the natural rate of unemployment, and the unanticipated component of the price level:

$$U_t = \alpha_0 + \alpha_1 \log (P_t/P_t^*), \tag{5}$$

where α_0 is the natural rate of unemployment, P_t is the current level of the implicit price deflator, and P_t^* is the expected level of the deflator. Darby specifies the expected component of prices, P_t^*, as being formed adaptively:

$$P_t^* = \lambda P_t + (1-\lambda) P_{t-1}^*, \tag{6}$$

with λ being the adjustment coefficient.[42]

Table 7 replicates Darby's results using Lebergott's original unemployment rate series U^L, Darby's "corrected" version of that series, U^D, and a measure of output relative to trend output, Q/Q^*. Each equation is estimated using a maximum-likelihood technique. The regression that minimizes the sum of squared residuals over various values of the expectations parameter is listed in table 7, along with the implied mean expectations-adjustment lag, calculated as $(1-\lambda)/\lambda$.

The annual regressions in section A indicate that regardless of the output measure used, Darby's measure of unanticipated prices is correctly signed and has a significant impact on output; for each of the output variants employed, the t-ratio on "price surprises" is very large. As in Darby's original paper, the estimate of the natural rate of unemployment obtained with the U^D unemployment series is relatively high, around 8 percent. The natural-rate estimate obtained using the Lebergott data is 5.9 percent.[43] Section B contains the results obtained when the output-ratio version is reestimated using quarterly data. The quarterly and annual regressions tell the same story, with a strongly significant impact of "price surprises" on output and an infinite lag in the adjustment of expectations.[44]

Darby's adjustments to Lebergott's unemployment series sharply alter the profile of unemployment in the Great Depression, especially in the late 1930s. A comparison of lines A1 and A2 in table 7 indicates, however, that the regression evidence in favor of the EAS hypothesis is no stronger using the Darby data then when the original Lebergott data is used. In fact, the natural-rate estimate and the t-ratio on the price-surprise variable are both more favorable to the EAS theories in line A1. Thus, we cannot conclude that the

Table 7. Effect of "Price Surprises" on Output, 1930–41

	Dependent Variable	Estimate of U^* (1)	Unanticipated Price Coefficient (2)	Mean Expectations Adjustment Lag (in years) (3)	SEE (4)	D-W (5)	Mean Expectations Lag Constrained Equal to 1 Year: SEE (6)	Mean Expectations Lag Constrained Equal to 1 Year: D-W (7)
A.	Annual data							
	1. U^L	0.059 (13.7)	−0.716 (−13.3)	∞	0.0123	2.22	0.0453	0.52
	2. U^D	0.080 (14.2)	−0.631 (−12.3)	9.0	0.0133	1.98	0.0302	0.66
	3. Q/Q^*	0.970 (46.4)	+1.208 (+9.5)	∞	0.0292	1.83	0.0790	0.56
B.	Quarterly data							
	1. Q/Q^*	0.990 (70.2)	+1.183 (15.2)	∞	0.0374	0.72	0.0876	0.13

Note: t-ratios in parentheses.

"corrected" data reveal that kind of consistency with the EAS framework where none previously existed.

One disconcerting element in these test results is the extremely long implied expectations formation lag. Column (3) indicates that only the Darby unemployment rate regression has a finite mean lag, and that lag is nine years.[45] The infinite adjustment lag estimated for the remaining regressions in table 7 means that expected price level remains a constant equal to the actual price level in 1924.

While it is true that prices in the 1920s exhibited little variance, so that expectations of nearly constant prices like those implied by a 9-year mean lag seem credible, from 1929 through 1940 prices fell sharply and then recovered appreciably. In every year during this 12-year stretch, agents overestimated the price change, whether the adjustment lag is 9 years or infinite. These overestimates are both large and persistent. For instance, in 1931 with the economy sliding further and further into depression, the estimated expected 1932 inflation rate implied by a 9-year lag is over 12 percent. Actually, 1932 brought 11 percent *deflation*. The question must be whether these estimates can plausibly describe the behavior of rational economic agents. We think not. And we suspect that many monetarists, particularly those who stress the excess of real over nominal interest rates in 1931–33, would agree (see Meltzer 1976).

Columns (6) and (7) of table 7 present standard errors of the estimates and Durbin-Watson statistics for the models described above under the restriction that the mean expectations adjustment lag is limited to one year.[46] The imposition of this restriction in each case substantially decreases the explanatory power of the EAS hypothesis. The standard errors more than double, and the Durbin-Watson statistics indicate strong serial correlation in the residuals, suggesting that Granger and Newbold's (1974) warning regarding the possibly spurious nature of the entire relation cannot be disregarded.

Thus Darby's new unemployment data make a minimal contribution to the case for the EAS framework. Unconstrained expectations estimates imply incredibly long, sometimes infinite, lags. Lag estimates restricted to a one-year mean revive problems of serious autocorrelation. The EAS theory cannot yet account for the behavior of prices and output in the Great Depression in a manner consistent with the rational formation of expectations.

Explaining Price Change

If the EAS explanation cannot satisfactorily account for the behavior of prices and output during the Great Depression, how adequate is the expectations-augmented Phillips curve (EPC) favored in recent postwar econometric work?

Table 8 presents the results of an attempt to determine whether the level or change in either unemployment concept in table 7 can explain the rate of change of prices in annual data.

Contrary to the usual Phillips curve relation between the level of the unemployment rate and the rate of change of prices, table 8 indicates that there is no significant effect of the level of unemployment in equations that also include the change in the unemployment rate. These equations and numerous others not reported here demonstrate conclusively that the relation between prices and unemployment (or the output ratio) in the interwar period links levels of each variable or the rates of change of each. There is no evidence of any empirical effect of the *level* of unemployment on the *rate of change* of prices, as called for by the Phillips curve hypothesis. The EAS results in table 7 are completely consistent with this finding, of course, because an infinite adjustment lag causes equation (5) to be converted into a relation between the *level* of unemployment and the *level* of a price index with 1924 as base:

$$U_t = a_0 + a_1 \log (P_t/P_{1924}). \tag{7}$$

The role of government intervention as a source of price and wage behavior in the 1930s has been stressed by several recent authors (R. J. Gordon 1976; Darby 1976*b*; Weinstein 1981). One method of identifying such effects is to introduce dummy variables into time-series regressions for years identified as "special" by external evidence. For instance, we know that the NRA (the National Recovery Administration) was established in June 1933 and abolished in May 1935. Any effects of the NRA on price behavior must therefore contaminate the annual-average data for each year between 1933 and 1936. To allow for the possible effect of the NRA, the equations in table 8 have been reestimated with the addition of annual dummy variables for 1933–36.

The results are interesting, particularly those presented in line 3. The coefficients on the individual dummies for the years when the NRA operated (1933 and 1934) are positive and are almost exactly counteracted by negative coefficients for the years of the dismantling of the NRA after its enabling legislation, the National Industrial Recovery Act, was declared unconstitutional (1935 and 1936).[47] It appears that the addition of these annual dummy variables fails to change the conclusion that the Darby unemployment variant (line 4 of table 7) provides a relatively poorer explanation of price change in the interwar period. The addition of the NRA dummies reduces the standard error of the estimate, as compared with line 2, but the sum of the dummy-variable coefficients in that version is an implausible, albeit insignificant, −8.4 percent.

Meltzer argues that "anticipations of inflation depend upon the prevailing monetary standard" (1977, p. 189), implying that empirical schemes that ap-

Table 8. Effect of the Lebergott and Darby Unemployment Concepts on the Rate of Change of Prices in the United States, Annual Data, 1922–41

	Coefficients of										
	Sum of Two Lagged Rates of Price Change (1)	U_t^L (2)	ΔU_t^L (3)	U_t^D (4)	ΔU_t^D (5)	*Dummy Variables* 1933 (6)	1934 (7)	1935 (8)	1936 (9)	*SEE* (10)	*D-W* (11)
1.	0.135	0.042	−0.951	———	———	———	———	———	———	0.0237	2.51
	(1.04)	(0.59)	(−6.67)								
2.	0.203	———	———	−0.037	−0.921	———	———	———	———	0.0265	2.50
	(1.36)			(−0.34)	(−5.31)						
3.	0.432	0.042	−0.957	———	———	0.029	0.049	−0.050	−0.035	0.0200	2.94
	(3.05)	(0.49)	(−6.75)			(+1.03)	(1.76)	(−1.89)	(−1.40)		
4.	0.495	———	———	0.069	−1.060	−0.001	0.029	−0.059	−0.053	0.0241	2.81
	(2.73)			(+0.41)	(−4.35)	(−0.02)	(0.78)	(−1.64)	(−1.66)		

Note: Additional variants yield the conclusion that U^L_{t-1}, U^D_{t-1}, $(U^L\text{-}U^D)_t$, and $(U^L\text{-}U^D)_{t-1}$ have no significant effect when added to the listed equations.

proximate expectation formation with fixed coefficient functions of past inflation rates may provide poor estimates of rationally formed expectations when the operation of the monetary system undergoes a basic alteration, as when "the international gold standard ended after . . . 1931" (p. 190). In order to test whether the relationship between the changes of prices and of output is sensitive to the empirical method used to estimate expectations, we have drawn upon Meltzer's method and estimated expectations as a function of lagged money growth and lagged average money growth.[48] Regardless of whether expectations of inflation are proxied by lagged inflation rates or are independently estimated functions of lagged money growth, and regardless of whether or not separate annual dummies for 1933–36 are included, and regardless of whether the whole sample or various subsamples are examined, the same result emerges: the rate of change of prices is significantly influenced, not by the level of output, but only by its current rate of change.

The European Experience

While dummy variables provide a crude method to gauge the impact of government intervention, another approach is to compare the division of nominal income between price and output change in the United States and some other countries or group of countries where government intervention was less important. This section compares the behavior of the United States with that of an aggregate of six European countries.[49] Without further research, it is impossible to determine whether any or all of these countries were completely free of new government measures that interfered with the setting of prices; the presumption here is that the degree of New Deal intervention in the 1933–38 period represents an extreme case that might be identified by a comparison with countries with less intervention.[50]

Figure 6 displays real output and the GNP deflator for the United States and Europe during the period 1925–38. It is clear that the division of nominal income change between price and output change was quite different in Europe. Expressed on a 1929 base, the U.S. output index was lower than its price index in every year between 1930 and 1935, whereas the reverse was true in Europe for every year of the 1930–38 period. On an annual basis, only 45 percent of the decline in U.S. nominal income during the 1929–33 contraction was expressed as price change, whereas during the same interval the equivalent figure for Europe was 73 percent. Because the greater extent of price flexibility in Europe was evident well before the advent of New Deal intervention in the price system, it appears that some other factor must have been primarily responsible for sluggish price adjustment in the United States. One possibility

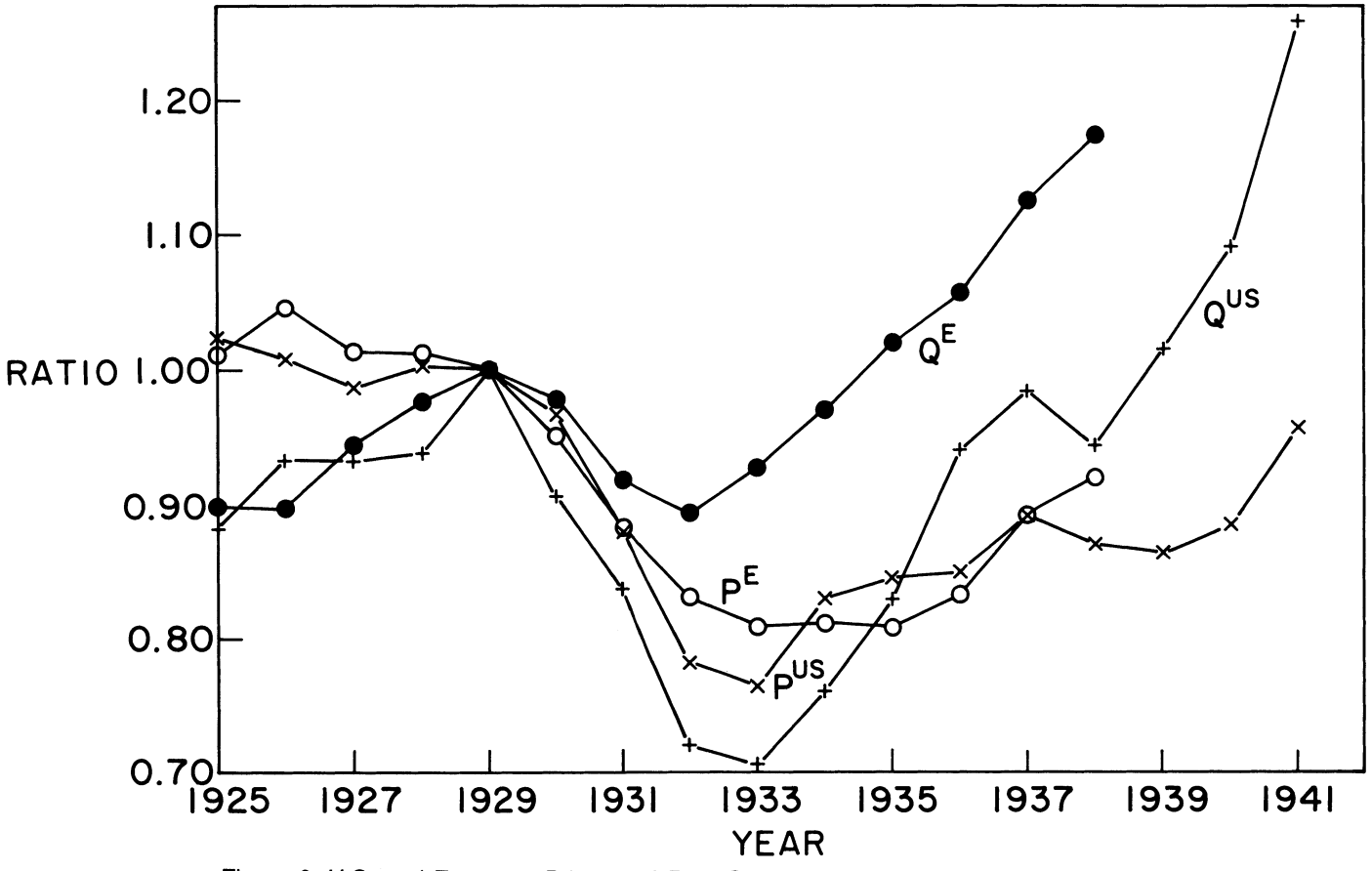

Figure 6. U.S. and European Prices and Real Output, 1925–41 (1929 = 1.0)

is the influence of the Hawley-Smoot tariff discussed above, although Meltzer's analysis (1976) relies heavily on retaliation by foreign countries to explain how the tariff contributed to the severity of the U.S. contraction.[51]

The difference between the U.S. and European aggregate-supply response is summarized in table 9, which presents the same specification as table 8 but replaces the alternative unemployment variables by the ratio of output to a trend (Q/Q^*), in order to compensate for the lack of comparable unemployment data for this period. The equations for Europe duplicate the U.S. result that the rate of price change is a function of the rate of change of the Q/Q^* level, not its level. Further, in the European equations the coefficient on the rate of change of Q/Q^* is significantly higher than in the United States, indicating that any given change in nominal income was reflected more in the form of price change and less as quantity change in Europe than in the United States.[52]

Since an identity links the rates of change of nominal income, the price level, and real output, the equations in table 9 can be reestimated in a form that makes the rate of change of prices a function of the current rate of change of nominal income and the lagged rate of change of prices. This allows a direct comparison of the impact of the differences between the European and the U.S. aggregate-supply functions, holding constant the behavior of nominal income. In figure 7 are plotted the annual level of the U.S. implicit GNP deflator (P_t) and the fitted values of prices in two dynamic simulations. The first ($\hat{P}_t$) is based on coefficients from a regression of U.S. price change on U.S. nominal-income change and lagged price change. The second ($\hat{\hat{P}}_t$) is based on coefficients from a regression of European price change on European nominal-income change and lagged price change fitted to 1928–38. Each simulation is calculated by multiplying these two alternative sets of coefficients by the *actual* rate of change of U.S. nominal income and the *fitted* values of lagged U.S. price change.

Several interesting features of figure 7 stand out prominently. First, the impact of government intervention on the price level is evident in the difference between P_t and $\hat{P}_t$. The rise in actual P_t relative to the simulated series in 1934 reflects the influence of NRA, and the subsequent slowness of increase in P_t presumably reflects the demise of NRA in 1935. Even more notable is the increase in P_t relative to $\hat{P}_t$ in 1937, caused at least partly by the influence of unionization.[53] Further, the simulated $\hat{\hat{P}}_t$ series based on European coefficients indicate that, given actual U.S. nominal-income behavior, the U.S. price level would have declined by 33 rather than only 24 percent during the period 1929–33 if prices had been as flexible as in Europe. The rapid increase in the U.S. price level during the period 1933–37, often cited as evidence of cost-push, instead appears to have been due to the very rapid growth of nomi-

Table 9. Effect of Output and Output Change on the Rate of Change of Prices, United States and Europe, Annual Data

		Coefficients of				
	Sample Period	*Sum of Two Lagged Rates of Price Change*	*Q/Q**	*Rate of Change of Q/Q**	*SEE*	*D-W*
A. United States						
1.	1922–41	0.156	0.031	_____	0.0460	1.42
		(0.61)	(0.33)			
2.	1922–41	0.016	_____	0.423	0.0318	2.44
		(0.09)		(4.20)		
B. Europe						
1.	1928–38	0.380	0.195	_____	0.0339	1.32
		(0.56)	(0.48)			
2.	1928–38	0.467	_____	0.794	0.0188	2.10
		(1.46)		(4.08)		

Note: t-ratios in parentheses.

nal income during this interval. In fact, the simulated series $\hat{\hat{P}}_t$ based on European coefficients and actual U.S. nominal-income growth registers a 1933–37 increase of 19.3 percent, greater than the 16.6 percent increase in the actual U.S. deflator during the same period. Thus, if the degree of price flexibility in the United States had been greater, U.S. prices would have rebounded even more in 1933–37 than they actually did.

CONCLUSION

Sources of Income Change

This paper has examined two different aspects of macroeconomic behavior in the United States during the 12-year period between 1929 and 1941—both the proximate determinants of the severity and duration of the slump in nominal income, and the factors influencing the division of those changes in nominal income between changes in the price level and in real output. The first topic involves the sources of shifts in aggregate demand, and the second concerns the slope and source of shifts in the aggregate-supply function. The link that unifies attention to both issues in a single paper is their relation to present-day monetarism. The preference of monetarists for monetary rules rather

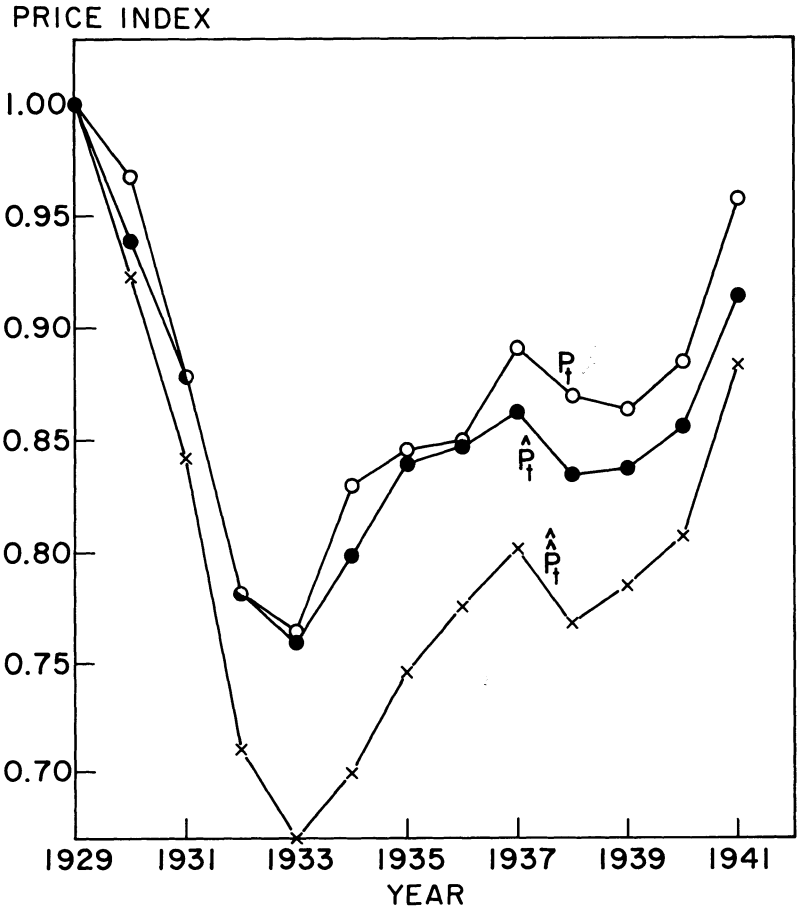

Figure 7. Comparison of Actual and Simulated Prices, 1929–41 (1929 = 1.0)

than countercyclical activism is based on their assumptions that private spending is basically stable in the absence of government interference, that government intervention does more harm than good, and that the price mechanism provides a powerful self-correcting force that insulates the economy from long-lasting swings in real output in the absence of government activism.

The first question, the sources of nominal-income movements, has been the subject of much recent controversy and debate, but we are persuaded that most of the heat has been unproductive. The common weakness of recent work has been its polemical and unscientific attempt to demonstrate that a single factor, the behavior of the money supply and monetary policy, either was solely responsible for the Great Contraction of 1929–33 (Schwartz and Darby) or played no role at all in the first two years of the contraction (Temin). The inherent weakness of single-factor explanations, or of denials of the influence of particular factors, is that they can be so easily contradicted. Schwartz and Darby must deny that *any* factor besides the 1928–29 deceleration in monetary growth was responsible for the rapid collapse of spending in the first quarters of the contraction. Temin must deny that a *single* deposit holder at a bank that failed in 1930 or 1931 was forced to cut back his spending on current goods and services by as much as a single dollar! Because such extreme positions fly in the face of common sense, we must register our surprise that they are still so firmly maintained.

This paper concludes that both nonmonetary and monetary factors played an important role in determining changes in nominal income during the period 1929–41. In holding that there must have been multiple causes rather than a single cause, we are only echoing a conclusion reached long ago by Haberler:

> Explanations which run in terms of one single cause have been more and more discredited and should be regarded with suspicion. The majority of modern writers on the subject are careful to point out that a whole set of factors, and perhaps not always the same combination of factors, contribute towards producing an alternation of prosperity and depression. [1958, pp. 5–6]

Four views ranging from extreme monetarism (*a*) to extreme nonmonetarism (*d*) were initially distinguished. Evidence has been presented that leads us to reject both views (*a*) and (*d*), leaving the intermediate soft-line monetarist and nonmonetarist views (*b*) and (*c*) as plausible explanations that differ only in emphasis.

View (*b*) essentially states that, while nonmonetary factors may partially have initiated the 1929–33 contraction, it was the failure of the Federal Reserve to offset the deflationary impact of bank failures that converted a serious recession into a severe depression. View (*c*) emphasizes the nature of the nonmonetary factors that played an important role in 1929–33 without denying that a countercyclical stimulus applied by the Federal Reserve could have less-

ened the severity and duration of the contraction. The difference between views (*b*) and (*c*) is inconsequential, representing mainly the greater interest of nonmonetarists in the 1929-31 phase of the contraction and of monetarists in the 1931–33 phase, and the two views are in fact almost perfectly complementary, each filling in the gaps left in the other's analysis.[54]

Weaknesses in a Purely Monetary Explanation

The present paper contains new evidence against a monocausal monetary explanation of the first two years of the contraction. Simulations based on the average relation between lagged values of the money supply and current values of nominal income in the 1920–28 interval suggest that the deceleration in monetary growth beginning in early 1929 cannot explain why the initial contraction of income was so severe. The initial slowdown of monetary growth in 1928–29 was no greater than in previous minor recessions in the 1920s, leaving unanswered the question why nominal income should have dropped by almost 30 percent during the first two years of the contraction.

The statistical relationship between lagged money and income is sufficiently weak, even in the 1920s, to raise serious questions about the ability of changes in the money supply and monetary policy to explain changes in nominal income during the interwar period. While an F-ratio on the joint contribution of the lagged money variables is significant in the 1920–28 period when the variables are expressed as growth rates, the F-ratio is insignificant for the level form of the variables. The t-ratios on the sum of the lagged money coefficients are insignificant in every period and for every variant of the equations. And questions may be raised as to whether a significant lead of money before income would have any meaning even if it could be found; some nonmonetarists might claim that money-supply swings reflect changes in the need of businessmen to finance inventory changes and that swings in these working-capital needs could precede business-cycle turning points.

As the sample period of the income-on-lagged-money regressions is extended into the 1930s, the coefficients on lagged monetary change become even weaker. After 1929 the relationship between money and income appears to be entirely contemporaneous, adding plausibility to the reverse-feedback hypothesis that the reflex influence of business on money was a primary determinant of money-supply swings during the period 1929–41. In the light of postwar time-series evidence indicating that swings in monetary growth induced by policy shifts require several quarters to influence income growth, it appears dubious that the purely contemporaneous relation of the 1930s could mainly reflect a money-to-income chain of causation.

Not only is a statistical relation between lagged money and income nonexistent after 1929, but in addition a purely monetary explanation cannot account for the duration of the slump of nominal income in the 1930s. The money supply grew very rapidly between 1938 and 1940 and in 1940 exceeded its 1929 average by almost 20 percent, yet income grew at a sluggish pace during the same period and in 1940 was still below its 1929 level.

A purely monetary explanation leaves unanswered why nominal income fell so rapidly during the 1929–31 period, why income grew so slowly during the 1938–40 period and so rapidly between 1940 and 1941, and why the relation between money and income in the 1930s should have been contemporaneous without the long lags that monetarists have emphasized in their critiques of policy activism. Yet nonmonetary explanations are available for each of these features of the period.

The first nonmonetary source of the 1929–31 contraction in income was the decline in residential housing construction, due both to a decline in population growth following the 1921 and 1924 legislation limiting immigration and to overbuilding during the mid-1920s. The decline in housing began in 1927 and became very steep in 1928 and 1929, and yet its impact on the aggregate economy was delayed by a temporary boom in consumption (and to some extent in nonresidential investment) stimulated by the speculative stock market bubble. The collapse in stock values brought about a rapid decline in consumption spending that added to and interacted with the impact of the housing slump. After the summer of 1930 the Hawley-Smoot tariff added to the contractionary pressure.

The timing of income change in the late 1930s also requires a mainly nonmonetary explanation. While money-supply growth was rapid and relatively steady between early 1938 and late 1941, nominal income grew slowly through mid-1940 and rapidly thereafter. Nonmonetarists point to the sluggishness of investment demand in 1938–40, and the enormous increase in defense spending in 1940–41, as an obvious explanation of this timing pattern. And, in the light of the weak relation between money and income in 1938–41, they would suggest that at least part of the simultaneous 1937–38 decline in money and income reflects, not the influence of an exogenous monetary policy shift, but rather the reverse-feedback effect of income on money following a very marked contractionary swing toward a full-employment fiscal surplus during the period 1936–37.

Weaknesses in a Purely Nonmonetary Explanation

There is no contradiction between the statements that (1) in the absence of a strong countercyclical monetary policy, the money-income relation in the

1930s was dominated by a contemporaneous feedback effect of income on money, and (2) an alternative activist monetary policy would have yielded a different set of data exhibiting a significant impact of lagged money on income. Nonmonetarists may rightly claim that, *given* the absence of monetary activism, nonmonetary factors were mainly responsible for the collapse in both money and income in 1929–33, but they thereby provide no proof that such activism could not have been effective.

Since the U.S. data are incapable of revealing the effects of behavior that did not occur, monetarists make a valuable contribution by pointing to the differences between European and U.S. behavior. The similarity in the behavior of velocity in Europe and the United States during the decade of the 1930s supports the monetarist conjecture that, had the United States followed Europe in preventing a collapse of the money supply, U.S. nominal income would have exhibited the milder contraction and earlier recovery actually observed in Europe.

Monetarists might also claim that the weakness of the effects of money-supply growth in 1938–40 could have been a consequence of earlier monetary inaction. As Hawtrey (1933) pointed out, once a depression has occurred and business expectations have become dominated by pessimism, a monetary expansion may not have the same stimulating effect that would have occurred earlier, and a combined monetary and fiscal expansion may instead be necessary to bring about a full recovery. It was such an expansion in 1940–41, of course, that finally brought the depression to an end in the United States.

Finally, the primary role of nonmonetary forces in explaining the initial phase of the 1929–33 contraction, and the inability of the small deceleration in monetary growth to explain why the contraction was so severe, may be admitted without precluding a role for money in determining the timing of the 1929 turning point. Without easy money in 1927–28 and tight money in early 1929, the stock market and consumption boom and collapse might have been dampened, and the course of nominal income might have more directly followed the path of the ongoing slump in housing investment.

The Aggregate-Supply Response

Neither the equilibrium aggregate supply (EAS) approach nor the expectational Phillips curve (EPC) appear at all adequate as explanations of the division of U.S. nominal-income changes between price and output changes in the 1930s. Deviations of unemployment or output from their natural levels, according to EAS, occur only when economic agents are surprised by the emergence of a price level different from that which they previously expected. While the EAS approach provides a plausible explanation of 1929–33, it cannot explain why

output remained so low and unemployment so high from 1933 to 1940. Price movements were sufficiently modest after 1934 to make surprises small by any reasonable version of how agents formed expectations; the computer is forced to conclude that an EAS econometric specification can explain unemployment and output in the late 1930s only if it is implausibly assumed that agents each year expected the price level to return to its 1924 value in the face of continuing evidence that no such return was occurring!

The EPC approach fails completely as well, because there is no evidence at all of a relation between price *change* and the *level* of unemployment or output during the 1930s, either for the United States or for an aggregate of six European countries. The statistical relation appears to have been between price change and output change, or between the level of prices and the level of output. These results lead to an interesting set of research questions to be explored in subsequent work. The finding that price change responds to output change but not the level of output is consistent with Meltzer's (1977) development of a price-specie-flow model of an economy operating under the gold standard. Changes in demand cause simultaneous changes in both output and prices, but the emergence of unemployment and an output gap is anticipated and has no independent effect on the rate of price change, as required in the EPC approach. Why the EPC appears to describe the postwar years but not the interwar years is attributed by Meltzer to the shift from the gold to the dollar standard, although R. J. Gordon's recent work (1977) on the postwar years suggests that even recently the dominant explanation of the rate of price change is the rate of change of the output gap rather than its level.[55]

Because the high *level* of unemployment had no independent effect on prices in the 1930s, the monetarist belief in the recuperative self-correcting powers of the private economy receives no support from the data. Some monetarist writings have stressed the role of government intervention as a source of cost-push pressure in the 1930s, but our results lead us to discount any crucial role for government in explaining the puzzles of U.S. aggregate supply behavior during that decade. With the exception of a temporary upward blip in prices in 1934, which vanished in 1935, and of a high rate of price increase in 1937, the year of greatest unionization, we find that a simple relation between price change, output change, and lagged price change fits the interwar data for both the United States and Europe quite well. Because Europe has much less price-raising intervention than the United States but exhibits the same type of supply response, doubt is cast on intervention as the main cause of U.S. behavior. The main difference between the United States and Europe—the steeper slope of the European supply function—was evident in 1929–33, well before the advent of the New Deal. The sources of sluggishness in U.S. price behavior prior to 1933 must stand high on an agenda of future

research topics, with an initial avenue of investigation being an attempt to quantify the role of the Hawley-Smoot tariff of 1930.

A Final Scorecard

In concluding that nonmonetary factors are essential in a complete explanation of the magnitude and timing of income movements in the 1930s, we deny the validity of Plank 1 of the monetarist platform with its emphasis on the inherent stability of private spending. But in agreeing with the basic Friedman-Schwartz proposition that a different policy response would have reduced the severity and duration of the Great Contraction, and in pointing to the harmful role of the Hawley-Smoot tariff, we lend our support to the message of Plank 3 that past government policy actions (and in 1929–33 the *absence* of appropriate policy actions) have done more harm than good. Finally, while denying any potency to the self-correcting mechanisms of price flexibility during the 1930s, as stressed in Planks 2 and 4, we must add that the underlying sources of aggregate supply behavior in the United States during the interwar period, and the reasons for changes in this behavior between the nineteenth century and the 1930s, and between the 1930s and the present day, must stand high on any agenda of unsolved research puzzles in macroeconomics.

APPENDIX: DATA SOURCES, 1919–41

IPC: Index of nominal value of industrial production. Calculated as the index of industrial production (Federal Reserve *Bulletin,* various issues) times the CPI (BLS).

M: Money.
United States: Friedman and Schwartz 1963*a*, Appendix A.
Europe (France, Germany, Italy, the Netherlands, Sweden, United Kingdom): Mitchell 1975, pp. 676–83.

P: Implicit price deflator.
United States: Annual—1919–21: Kuznets 1941; 1922–28: Hickman and Coen 1976; 1929–41: U.S. Department of Commerce 1976. Quarterly—generated using the Chow-Lin (1971) technique that distributes annual series into quarterly series using related, quarterly series; the related series used were the CPI and the WPI.
Europe (except France): Mitchell 1975, pp. 785–90. Calculated as the quotient of current-dollar divided by constant-dollar gross national product.

France: The German implicit price deflator was regressed on the German WPI, both in logs. The resulting coefficients were then multiplied by the French WPI to obtain an estimate of the French implicit price deflator. French and German WPI data series are from Mitchell 1975.

Q: Real output.

United States: Annual—1919–21: Kuznets 1941; 1922–28: Hickman and Coen 1976; 1929–41: U.S. Department of Commerce 1976. Quarterly—three related series (industrial production, real department store sales, and a linear trend) were employed to generate quarterly real output using the Chow-Lin (1971) technique. See *P*.

Europe: Mitchell 1975, pp. 785–90. The aggregate for Europe is the sum of real GNP for the six European countries, converted into dollars using 1929 exchange rates.

Q*: Natural rate of output. 1913–29: calculated as the exponential trend between the real GNP levels of 1913 and 1929; 1930–41: calculated as the extrapolation of the 1913–29 annual trend rate of growth of 2.54 percent using the actual rate of output in 1929 as the natural rate of output in 1929.

S: Index of department store sales. Federal Reserve *Bulletin,* various issues.

U^D: Unemployment rate. Darby 1976*b*.

U^L: Unemployment rate. Lebergott 1964.

Y: Nominal GNP.

United States: P multiplied by *Q*.

Europe: Mitchell 1975, pp. 785–90. The aggregate for Europe is the sum of nominal GNP for the six European countries converted into dollars using 1929 exchange rates.

NOTES

1. Outside of the context of the Temin debate, several monetarist authors have provided important recent interpretations of the price-output division of nominal income. See especially Meltzer (1977) and Darby (1976*b*).

2. The primary emphasis in this paper on monetarist interpretations reflects the topic selected by the organizers of the conference on the Great Depression and does not imply any belief on our part that nonmonetarist interpretations should be immune from detailed scrutiny.

3. The phrase "battle of the radio stations" comes from the initials (AM-FM) of the main protagonists in a 100-page debate published in 1965 in the *American Economic Review*. See Ando and Modigliani (1965) and Friedman and Meiselman (1965).

4. The development of the monetarist platform benefited from the suggestions of Milton Friedman, Allan Meltzer, Franco Modigliani, and Arthur Okun. It is supported by a more extensive discussion in R. J. Gordon (1978, pp. 335–43).

5. This explains the apparent oddity that the word *money* does not appear in the platform. Were it not for the popularity of the word *monetarist* among both economists and journalists, the platform might be better described by the term *antiactivist*.

6. A clear example is Friedman's (1968) statement that "the simultaneous fall *ex post* in real wages to employers and rise *ex ante* in real wages to employees is what enabled employment to increase." Empirical studies by Lucas and Rapping (1969), Darby (1976*b*), and Barro (1977) place unemployment or real output on the left-hand side of the equation and deviations of nominal variables from their expected values on the right-hand side. Schwartz writes in the same vein, "Why quantities changed as they did *in response to* price changes should be the goal of analysis" (1981, p. 21, emphasis added).

7. In response to a demand shock, prices do not typically adjust rapidly enough to clear markets, and so agents find themselves constrained by a level of sales or employment different from what they would voluntarily choose to supply at going prices and wages (Barro and Grossman 1976, chap. 2). Prices and wages are not completely sticky; instead, their adjustment to excess demand or supply in any given time period is partial rather than instantaneous and complete.

8. R. J. Gordon (1977) presents both wage and price equations with the coefficient of expectations constrained to be 1.0, thus placing the difference between actual and expected values on the left-hand side.

9. "A far more satisfactory explanation of 1929–33 than Temin's is, therefore, that a series of negative shocks, monetary in origin, reduced real output. . . . There are no unexplained changes in spending that serve as deus ex machina. . . . The behavior of the economy was determined by public policies. Different policies would have resulted in different behavior" (Schwartz 1981, p. 26).

10. The clearest admission of the possible role of nonmonetary forces comes in Friedman and Schwartz (1963*a*): "True, as events unfolded, the decline in the stock of money and the near-collapse of the banking system can be regarded as a consequence of nonmonetary forces in the United States, and monetary and nonmonetary forces in the rest of the world. . . . Prevention or moderation of the decline in the stock of money, let alone the substitution of monetary expansion, would have reduced the contraction's severity and almost as certainly its duration. The contraction might still have been relatively severe" (pp. 300–01). On the endogenicity of the money supply, see Friedman and Schwartz (1963*b*): "The reflex influence of business on money, the existence of which is not in doubt in light of the factual evidence summarized above, would then become part of the partly self-generating mechanism whereby monetary disturbances are transmitted" (pp. 49–50).

11. "The first thing to be said about [the Great Depression] is that it was a *double* slump. It began with the Wall Street crash in 1929, a repetition, at least at first sight, of that of 1907, leading to a depression just as that had done. But the recovery from the depression, which on previous experience might have been expected to follow within a year or two, did not take place. Instead there was a double slump, superimposed upon the first. Now there is no doubt at all that this second slump was monetary in character" (p. 210). Hicks dates the second stage from the fall of 1931, thus differing from the Friedman-Schwartz emphasis on the role of bank failures in the fall of 1930.

12. In 1974 R. A. Gordon was close enough to the Friedman-Schwartz position to agree that "vigorous action by the Fed could have substantially reduced the severity of the depression" (p. 72).

13. The peak of the cycle was 1929:3. Sources of data are identified in the Appendix at the end of the paper.

14. P is the quarterly GNP deflator. See Appendix.

15. Econometric studies of consumption functions generally support a real-balance effect that makes IS depend on M/P, but this added factor does not alter our conclusions; it simply makes the DD curve flatter without changing the variables that cause it to shift its position.

16. We believe that Temin was unwise to use IS–LM curves in a problem involving variable prices without also examining the SS–DD diagram shown in the bottom frame of figure 2. For a full development of this diagrammatic apparatus, and a discussion of the variables that make SS shift its position, see R. J. Gordon (1978, pp. 143–98).

17. This represents a graphic translation of an argument made by Schwartz (1981). The distinction between nominal and real interest rates is incorporated into IS–LM analysis in R. J. Gordon (1978, pp. 289–91).

18. The inverse of velocity, M_2/PQ, is displayed in figure 1.

19. This aspect of the Great Depression is emphasized by Fisher (1933).

20. Two channels by which the decline in nominal GNP could have caused the bank failures are (1) by reducing the nominal sales of individuals and firms to which banks had lent money, turning initially sound loans into loans that could not be repaid, and (2) by reducing the prices of bonds, as securities markets reflected the increased "price of risk," thus contributing to the insolvency of banks holding risky bonds (Temin 1976, pp. 103–21).

21. Schwartz (1981) has independently used the Granger method to evaluate the Temin interpretation of the 1929–33 contraction. While our results are consistent with hers in rejecting Temin's extreme position (d), we go beyond her results by running simulations that tend to reject her own extreme position (a).

22. In table 1, $K = 4$ and $L = 8$ when money is the independent variable, and $L = 4$ when income is the independent variable. This difference in the value of L occurs because quarterly income data are not available before 1919, and we were urged by a discussant to start our sample period in 1920:2 in order to capture the relation between money and income in the 1920–21 recession.

23. In table 2, $K = L = 8$. Inclusion of extra lagged values beyond eight yielded insignificant coefficients and did not alter the results displayed in table 2.

24. Between 1929:3 and 1933:2, M_1 fell by 31.8 percent and M_2 fell by 35.1 percent. The sums of coefficients on lagged money in the 1920–28 sample period are as follows in each version:

	Level	*Rate of Change*
M_1	0.970	0.213
M_2	0.936	0.520

25. Nonresidential fixed investment in 1940 was exactly the same as in 1937 in real terms and grew only 3 percent in nominal terms, whereas nominal M_2 grew 18.5 percent between 1937:2 and 1940:2.

26. Between 1940 and 1941 nominal defense spending grew by an amount equal to 11.5 percent of 1940 nominal GNP. The annual growth rate of M_2 in the six quarters after 1940:2 (12.2 percent) was little different from that in the six quarters before 1940:2 (11.0 percent).

27. For sources and data, see Gordon (1978, p. 496).

28. In the case of the regressions in which the variables are in the form of growth rates, the sums of coefficients on lagged money are almost always negative and are never significantly different from zero. Also, the F-ratio on the significance of lagged-M_2 growth rates declines appreciably as the sample moves through the 1930s.

29. The European data refer to the total of France, Germany, Italy, the Netherlands, Sweden, and the United Kingdom aggregated with weights according to 1929 GNP in dollars. Sources are described in the Appendix.
30. As examples, the ratio of real fixed gross investment to Q^* was 0.144 in 1955–57, 0.130 in 1958–61, 0.150 in 1965–69, and 0.124 in 1975–77.
31. Hickman and Coen (1976, p. 194) estimate a multiplier for changes in real autonomous spending (for five years after the change) of 5.09 under the conditions of 1926–40 and only 2.10 under the conditions of 1951–65. R. J. Gordon (1978, p. 494) calculates that the automatic fiscal stabilizers absorbed only 5.5 percent of the decline in GNP in 1932, but 36.9 percent in 1975.
32. Meltzer himself shows that industrial production had already fallen by 25 percent at the time of the first bank failures in October 1930 (1976, p. 464).
33. "Standardized households" are calculated by applying fixed 1940 household-headship rates to each age group. Since headship rates among children are negligible, the endogenous decline in the birth rate caused by the drop in income during the depression could not have altered the number of standardized households in the 1929–41 period.
34. See Hickman's presentation (1973, table 3, p. 307) of results for each year both for simulation II (standardized households grow at 1924–25 rate) and simulation III (standardized households follow actual path). Temin's summary of this same paper states that "holding income constant in this model eliminates most of the fall in construction in the 1930's by eliminating the observed fall in the rate of family formation in that decade" (1976, pp. 46–47). But Hickman's simulation III, which holds income constant while allowing standardized households to follow their actual path, does *not* eliminate the observed fall in the rate of household formation (in simulation III the rate of household formation falls from 579 million in 1925 to a trough of 377 million in 1937, for a decline of 34.9 percent). Nor is most of the decline in construction eliminated, since housing starts fall in simulation III from 977 million in 1925 to 372 million in 1940, for a decline of *61.9 percent.*
35. The oral tradition was passed on to us by George R. Neumann, to whom we are indebted.
36. The ratio was 8.6 percent in 1924–27 (Hickman and Coen, 1976, table A.2, p. 222). None of the postwar *individual* peak years of residential construction spending (1950, 1955, 1959, 1964, and 1972) came close to the ratio of *any* of the four successive peak years of the 1920s (the ratios for these postwar years are 6.2, 5.4, 5.3, 5.0, 5.3).
37. The expanding role of government is evident in the following comparison of ratios to actual real GNP in 1929 and 1940:

	1929	*1940*	*Change*
Consumption expenditures	68.5	67.1	−1.4
Gross private domestic investment	17.8	13.0	−4.8
Net exports	0.7	0.9	+0.2
Government purchases	13.0	19.1	+6.1

38. Between 1922 and 1923 U.S. real output grew 13.1 percent, compared to 3.5 percent in the United Kingdom, 8.2 in France, and 6.3 in Canada. At the same time, the relative price of exports fell by 9.2 percent. The Emergency Tariff Act of 1921 and the Fordney tariff of 1922, in the absence of perfectly elastic supply schedules, prevented relative export prices from falling further.
39. If income elasticities for U.S. exports and imports were equal, the smaller decline in European income should have led to a smaller decline in U.S exports, abstracting from relative price effects. The zero change in the trade balance argues that relative price effects cut the trade balance. But this result depends on equal income elasticities and would be invalidated by an income elasticity higher for U.S. exports than for imports.

40. In this era the economy of western Europe plus Canada was about equal in size to that of the United States. The GNP in dollars of the subset of six countries plotted in figure 5 was 75 percent of U.S. GNP in 1929 (see Appendix).
41. For a critique of Darby's redefinition of unemployment, see R. J. Gordon (1976, pp. 195–96) and Kesselman and Savin (1978).
42. Actual and expected prices were assumed equal in 1924.
43. This may still be an overestimate, since, even in the presence of far more lucrative welfare and unemployment benefits in the 1970s, current estimates of the natural rate hover around 6 percent (R. J. Gordon 1977).
44. At the suggestion of Robert Lucas, we conducted separate experiments in which lagged dependent variables were added to each of the regressions reported in table 7. In each case the lagged variables were insignificant.
45. This exceeds Darby's estimate of 5.7 years, presumably because of our use of a slightly different price series prior to 1929. The methods and other data series used here are identical to Darby's.
46. This lag restriction is arbitrary but is much closer to postwar estimates than a nine-year or infinite lag.
47. The sum of the individual-years coefficients is -0.7 percent, implying that whatever impetus the NRA gave to price and wage increases was completely reversed after the NRA was abolished. Since the NRA was both established and declared unconstitutional in midyear, the significance of the NRA's initial (1933)- and terminal (1935)-year coefficients are probably understated. An F-ratio on the joint significance of the 1933–36 coefficients, which is probably similarly understated, passes a significance test at the 10 percent level.
48. Over the 1922–41 period, agents choosing between predictions of inflation based on lagged inflation rates or on measures of lagged money growth would have been better off choosing the former. The R^2's and standard errors of the two versions are 0.22 and 0.0455, and 0.24 and 0.0448, respectively, in equations explaining the annual rate of change of the GNP deflator.
49. The six European nations are the same as those identified in n. 29 and for which money and nominal-income data are plotted in figure 5.
50. The most important cases of intervention in Europe were the German price and wage controls and the French Blum experiment. Bry (1960) suggests that German price controls caused the official cost-of-living index to be understated from 1937 on. Kalecki (1938) concludes that the Blum experiment raised wages and wholesale prices by 60 percent in France in 1937, without having any appreciable effect on real output. The upward push on prices in France in 1937 thus, to some extent, offsets the German controls, which became tighter in 1936–38.
51. If retaliation had been complete and instantaneous, then the tariff could make no contribution to the explanation of the greater degree of price flexibility in Europe.
52. The fraction of nominal income going into price change in the short run (given lagged prices) is equal to $\alpha/1 + \alpha$, where α is the coefficient on the rate of change of Q/Q^* in table 9. This fraction is 30 percent for the United States on line A2 and 44 percent for Europe in line B2.
53. The ratio of union members to civilian employment more than doubled between 1936 and 1938 and showed little change before 1937 or between 1938 and 1942.
54. Mayer (1978*b*) and others comment on the notable lack of attention to the nature of the 1929–30 phase of the contraction by Friedman and Schwartz.
55. Our research here supports Meltzer's in linking the rate of change of prices and the rate of change of output but conflicts with his in finding no conclusive evidence that price expectations were based on the recent behavior of monetary growth. Instead, we find that price change is better predicted by past price change than by past monetary change and that the money-to-prices link was particularly weak in 1937–40 (in 1940 the GNP deflator was below

its 1937 value, despite the 20 percent growth in M_2 and 30 percent growth in M_1 that occurred during the interval). In part, our differences with Meltzer may reflect the fact that Meltzer actually fits no equations that include only the interwar period. In his regressions for 1901–40, any looseness of the money-to-prices relation in 1937–40 must be dominated by the high variance of both money and prices during the World War I period, 1916–1920.

REFERENCES

Andersen, L. C., and Jordan, J. L. 1968. "Monetary and Fiscal Actions: A Test of Their Relative Importance in Economic Stabilization." Federal Reserve Bank of St. Louis *Review* 50: 11–24.

Ando, A., and Modigliani, F. 1965. "The Relative Stability of Monetary Velocity and the Investment Multiplier." *American Economic Review* 55: 693–728.

_____. 1976. "Impacts of Fiscal Actions on Aggregate Income and the Monetarist Controversy: Theory and Evidence." In *Monetarism,* ed. J. L. Stein. Amsterdam: North-Holland.

Artus, J. R., and Sosa, S. C. 1978. "Relative Price Effects on Export Performance: The Case of Nonelectrical Machinery." *IMF Staff Papers* 25: 25–47.

Barro, R. J. 1977. "Unanticipated Money Growth and Unemployment in the United States." *American Economic Review* 67: 101–15.

Barro, R. J., and Grossman, H. I. 1976. *Money, Employment, and Inflation.* Cambridge: Cambridge University Press.

Bolch, B. W., and Pilgrim, J. D. 1973. "A Reappraisal of Some Factors Associated with Fluctuations in the United States in the Interwar Period." *Southern Economic Journal* 39: 327–44.

Bry, G. 1960. *Wages in Germany, 1871–1945.* Princeton, N.J.: Princeton University Press.

Chow, G. C., and Lin, A. 1971. "Best, Linear, Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series." *Review of Economics and Statistics* 53: 372–75.

Darby, M. R. 1976*a*. *Macroeconomics.* New York: McGraw-Hill.

_____. 1976*b*. "Three-and-a-Half-Million U.S. Employees Have Been Mislaid; Or, An Explanation of Unemployment, 1934–41." *Journal of Political Economy* 84: 1–16.

Fisher, I. 1933. "The Debt-Deflation Theory of Great Depressions." *Econometrica* 1: 337–57.

Friedman, M. 1968. "The Role of Monetary Policy." *American Economic Review* 58: 1–17.

Friedman, M., and Meiselman, D. 1965. "Reply to Ando and Modigliani and to Deprano and Mayer." *American Economic Review* 55: 753–85.

Friedman, M., and Schwartz, A. 1963*a*. *A Monetary History of the United States.* Princeton, N.J.: Princeton University Press, for the National Bureau of Economic Research.

———. 1963*b*. "Money and Business Cycles." *Review of Economics and Statistics* 45: 32–78.

Goodwin, R. M. 1955. "A Model of Cyclical Growth." In *The Business Cycles in the Post War World,* ed. E. Lundberg. London: Macmillan.

Gordon, R. A. 1951. "Cyclical Experience in the Interwar Period: The Investment Boom of the Twenties." In Universities-National Bureau Committee for Economic Research, *Conference on Business Cycles.* New York: National Bureau of Economic Research.

———. 1974. *Economic Instability and Growth: The American Record.* New York: Harper & Row.

Gordon, R. J. 1976. "Recent Developments in the Theory of Inflation and Unemployment." *Journal of Monetary Economics* 2: 185–219.

———. 1977. "Can the Inflation of the 1970's Be Explained?" *Brookings Papers on Economic Activity* 8: 253–79.

———. 1978. *Macroeconomics.* Boston: Little, Brown.

Granger, C. W. J. 1969. "Investigating Casual Relations by Econometric Models and Cross-Spectral Methods." *Econometrica* 37: 424–38.

Granger, C. W. J., and Newbold, P. 1974. "Spurious Regressions in Econometrics." *Journal of Econometrics* 2: 111–20.

Haberler, G. 1958. *Prosperity and Depression.* Cambridge, Mass.: Harvard University Press.

———. 1976. *The World Economy, Money, and the Great Depression, 1919–39.* Washington, D.C.: American Enterprise Institute.

Hawtrey, R. G. 1933. *Trade Depression and the Way Out.* London: Longmans.

Hickman, B. G. 1973. "What Became of the Building Cycle." In *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz,* ed. P. David and M. Reder. New York: Academic Press.

Hickman, B. G., and Coen, R. M. 1976. *An Annual Growth Model of the U. S. Economy.* New York: American Elsevier.

Hicks, J. R. 1974. "Real and Monetary Factors in Economic Fluctuations." *Scottish Journal of Political Economy,* November: 205–14.

Jonung, L. 1981. "The Depression in Sweden and the United States: A Comparison of Causes and Policies." In this volume.

Kalecki, M. 1938. "The Lesson of the Blum Experiment." *Economic Journal* 48: 26–41.

Kesselman, J. R., and Savin, N. E. 1978. "Three-and-a-Half Million Workers Never Were Lost." *Economic Inquiry* 16: 205–25.

Kuznets, S. 1941. *National Income and Its Composition.* New York: National Bureau of Economic Research.

Lebergott, S. 1964. *Manpower in Economic Growth.* New York: McGraw-Hill.

Lucas, R. E., Jr. 1973. "Some International Evidence on Output-Inflation Trade-Offs." *American Economic Review* 63: 326–34.

———. 1976. "Econometric Policy Evaluation: A Critique." In *The Phillips Curve and Labor Markets,* ed. K. Brunner and A. Meltzer. *Journal of Monetary Economics* 2 (supp.): 19–46.

Lucas, R. E., Jr., and Rapping, L. A. 1969. "Real Wages, Employment, and Inflation." *Journal of Political Economy* 77: 721–54.

Maddison, A. 1964. *Economic Growth in the West.* New York: Twentieth-Century Fund.

Mayer, T. 1978*a*. "Consumption in the Great Depression." *Journal of Political Economy* 86: 139–46.

———. 1978*b*. "Money and the Great Depression: A Critique of Professor Temin's Thesis." *Explorations in Economic History* 15: 127–45.

Meltzer, A. H. 1976. "Monetary and Other Explanations of the Start of the Great Depression." *Journal of Monetary Economics* 2: 455–72.

———. 1977. "Anticipated Inflation and Unanticipated Price Change." *Journal of Money, Credit and Banking* 9: 182–205.

Mishkin, F. S. 1978. "The Household Balance-Sheet and the Great Depression." *Journal of Economic History* 38: 918–37.

Mitchell, B. R. 1975. *European Historical Statistics, 1750–1970.* New York: Columbia University Press.

Rees, A. 1972. "Real Wages and Inflation: Rejoinder." *Journal of Political Economy* 80: 192.

Sargent, T. J. 1976. "A Classical Macroeconometric Model for the United States." *Journal of Political Economy* 84: 207–38.

Schwartz, A. J. 1981. "Understanding 1929–1933." In this volume.

Sims, C. A. 1972. "Money, Income, and Causality." *American Economic Review* 62: 540–52.

———. 1977. "Macroeconomics and Reality." Department of Economics, University of Minnesota, Discussion Paper No. 77–91.

Stein, J. L. 1976. *Monetarism.* Amsterdam: North-Holland.

Temin, P. 1976. *Did Monetary Forces Cause the Great Depression?* New York: W. W. Norton.

Tobin, J. 1975. "Keynesian Models of Recession and Depression." *American Economic Review* 65: 195–202.

U. S. Department of Commerce. 1976. *The National Income and Product Accounts.* Washington, D.C.: Government Printing Office.

Weinstein, M. 1981. "Some Macroeconomic Impacts of the National Industrial Recovery Act, 1933–1935." In this volume.