The Legacy Of The Monetarist Controversy

INTRODUCTION

It is not quite true, as one (hostile) commentary has asserted, that Monetarism was developed by "Milton Friedman at the Federal Reserve Board (sic) of St. Louis," but it is nevertheless the case that the intellectual environment created at this Bank by Homer Jones ensured that the doctrine took root and flourished here when it was very much a minority taste elsewhere. And indeed, at least two early and seminal contributions to the Monetarist controversy, Andersen and Jor -
dan (1968) and of course Brunner (1968), which gave the controversy its label, first appeared in the Bank's Review. The Monetarist controversy, therefore, is surely a suitable topic for this lecture. Now Monetarism has been much defined and debated over the years, to the point at which one may find authority for applying the term to almost any economic and/or political doctrine one likes, or more probably dislikes. However, it is not so much my purpose here to define that doctrine in detail yet again, as it is to discuss the consequences for the development of monetary economics, both in theory and practice, of the debates to which it gave rise during the 1960s and 1970s.

In this lecture, I shall first of all describe the issues that were at stake at the outset of those debates. I shall show that although the Monetarist policy agenda was very different from that of what we might call "Keynesian" ortho -
doxy, the positive differences in economic analysis which underlay the policy debate were at first empirical in nature, raising no fundamental questions of economic theory. I shall also show, however, that, whether it was logically necessary or not, theoretical considerations of profound importance did get introduced into the Monetarist controversy as it progressed. Although these at first seemed to strengthen the Monetarist position, I shall go on to argue that these very considerations in the longer run undermined it, leaving Monetarism without distinct theoretical foundations, and incapable of coping with the empirical difficulties which it began to encounter from the mid-1970s onward. Furthermore, I shall suggest that most, though not all, of those empirical problems stemmed from an attempt by "Keynesian" orthodoxy to adapt Monetarist ideas to its own use. Finally, I shall argue that though the Monetarist controversy was to all intents and purposes over by the early 1980s, the problems which it bequeathed to monetary economics continued to affect theoretical, empirical- and policy-oriented aspects of the sub-discipline into the 1980s, with results which

I cannot help but view with considerable discomfort.

THE MONETARIST CONTROVERSY IN THE 1960s

The Monetarist controversy concerned the role played by monetary variables in general, and the quantity of money in particular, in the macroeconomy. Different exponents of Monetarism stressed different propositions, but it would be fair to say that, from the point of view of the non-academic observer whose main concern was the conduct of economic policy, Monetarism involved first a theory of inflation, second a theory of the cycle, and third, as a corollary of these, a recommendation for the conduct of monetary policy. Specifically, inflation was said to be explicable in terms of the rate of growth of the money supply, and the cycle, or more precisely its turning points, in terms of changes in that rate of growth. From these propositions, it immediately seemed to follow that both inflation and cycles could be avoided by choosing an appropriate rate of growth for the money supply, and binding the monetary authorities to deliver it year in and year out by imposing upon them some quasi-constitutional rule of conduct. The central item on the Monetarist policy agenda was thus to eliminate inflation and stabilize the real economy by taking discretionary power away from the central bank. ²

Whatever position may be taken about its validity, this Monetarist agenda is nowadays treated as having been (and perhaps as still being) worthy of serious discussion. It was not so 25 or 30 years ago. Then, though the Phillips curve, about which I shall have more to say below, was coming onto the scene, the predominant view treated inflation as largely a matter of “cost-push” forces. The cycle was regarded as having its roots in investment fluctuations, and some mixture of wage-price guidelines and fiscal measures was thought best able to cope with

---

²In an earlier essay, Laidler (1982), Ch. 1., I analyzed the essential characteristics of Monetarism from an academic perspective, concentrating there on the role of a stable demand for money function and the expectations-augmented Phillips curve in defining the doctrine. As the reader will see, these relationships play a large part in the following discussion, and I regard this essay as supplementing rather than in any way contradicting this earlier piece.

²Thus Laidler (1969) used this framework to motivate its discussion of the significance of the demand for money function, while Friedman (1971) also used it as an expository device. Note, however, that Brunner and Meltzer (e.g., 1976), because of their insistence on the importance of credit market effects in the generation of the money supply, were led to extend this framework to a point at which it became sufficiently different in its characteristics to make it misleading to treat it as simply one more variant on IS-LM.
seemed to confirm directly the irrelevance of variations in autonomous expenditure, while at the same time attributing a considerable influence on money income's behavior to the growth rate of the money supply.

The appealing simplicity of these results did not long survive further empirical investigation. A number of studies soon found a role for interest rates to play in influencing velocity, hence re-opening the theoretical possibility of variables other than money affecting the time path of money income, and the Friedman-Meiselman results were not robust in the face of small changes in the way in which the Keynesian concepts of autonomous and induced expenditure were measured. Nevertheless, from a practical point of view, the modifications to the basic Monetarist position required by these results were rather minor. The quantity of money did seem to be an economic variable of potentially strategic importance; and real shocks originating in the private sector, not to mention impulses coming from fiscal policy, did seem to play a potentially less important role in determining money income's time path than the conventional wisdom of about 1960 would have had it. By 1967 results such as these were sufficiently well established and widely accepted that it was possible for the present writer to begin work on a supplementary textbook (Laidler, 1969) which summarized the empirical evidence on the demand for money function, and explicitly interpreted it in terms of the above-mentioned IS-LM model along just such lines as these.

Now the reader will be well aware that our earlier confidence in the empirical stability of the demand for money function has not been entirely justified by subsequent experience. I will take up this matter below. For the moment it is more important to concentrate upon another anomaly in this aspect of the Monetarist controversy, namely that empirical evidence on the stability of the demand for money function, though relevant to Monetarist propositions about the causative role of money vis à vis both inflation and the cycle, and hence also to proposals to tie down monetary policy by way of a growth rate rule, stops far short of establishing them. Thus, if monetary elements in the generation of inflation on the one hand, and the cycle on the other, are to be discussed coherently, one requires more than a link between the time path of money and money income. One also needs a theory of how fluctuations in money income are divided up between its price level and real income components. Moreover, a stable demand for money function is quite compatible with the existence of cost-push inflation and a cycle whose origins lie in the private sector of the economy, provided only that institutional arrangements are such as to render the supply of money as an essentially passive variable.

Empirical evidence about the stability of the demand for money function, that is to say, got Monetarist analysis taken seriously enough for it to become controversial, but it could not in and of itself guarantee its victory in a debate which rather centered on the two issues raised above. The first issue became known as the problem of Friedman's "missing equation," and debate about it overlapped heavily with discussions of the "Phillips curve," otherwise known as the "inflation-unemployment trade off." As to the second, it was addressed by such workers as Philip Cagan (1965), and Brunner and Meltzer (e.g., 1964, 1976) who opposed their findings to the then "new view" of money propounded by James Tobin (e.g., 1969) and his associates. It will be helpful to discuss those two aspects of what we might term the second stage of the Monetarist controversy in turn.

The problem of decomposing variations in money income into their real and price level components was a longstanding one in macroeconomics, dating back at least to Hicks' original (1937) formulation of the IS-LM framework as a model of the determination of money income. Though subsequent developments of this system reinterpreted it as dealing with real income, that still left prices unexplained. Here, the usual solution was to have them determined by the behavior of the money wage, and as I have already noted, to treat the latter variable as being driven by exogenous "cost-push" factors. However, this is not the whole story, for "demand-pull" explanations of inflation also had their adherents, and the Phillips (1958) curve relating the rate of money wage inflation inversely to

---

4 Among early papers finding a significant interest elasticity of demand for money were Meiseler (1963) and Laidler (1966). Both Ando and Modigliani (1965) and De Prato and Mayer (1965) showed that fairly small changes in the definition of "autonomous" expenditure led to rather large effects on the assessment of its influence on aggregate demand. Strangely enough, at this stage of the Monetarist controversy, questions about vagueness in the definition of money were not raised.
the level of unemployment soon came to be seen as a device for synthesizing these two points of view. Wage inflation, according to the Phillips curve was proximately caused by an excess demand for labor, for which variable the unemployment rate was a proxy. Such excess demand could either result from "pull" forces shifting the labor demand curve, or "push" influences shifting the supply curve. In either case, however, price inflation would be determined by the difference between wage inflation and labor productivity growth. Hence there seemed to exist a structural inverse trade-off between inflation and unemployment, or equivalently a positive relationship between the level of real output and inflation.\(^5\)

From the point of view of microeconomics, the above analysis was fatally flawed, being based on a theory which had the supply and demand for labor determine the money, rather than the real, wage, and it was not long before Phelps (1967) and Friedman (1968) independently pointed this out. The latter, moreover, brought his critique into the center of the Monetarist controversy, by noting, first that the orthodox Phillips curve predicted that an inflationary monetary policy could generate permanent gains in real income, and second, that when its underpinnings were corrected to take account of the money wage-real wage distinction, the inflation-unemployment trade-off was reduced to a temporary phenomenon at most. Hence, a key ingredient of the theory which yielded the characteristic Monetarist propositions about money growth affecting only prices in the long run but quantities too, in the short run, was created after the empirical observations upon which those propositions were based had been published, and turned out to be a theory not of money, but of labor market behavior. It seemed to be, moreover, in its original form, a modification of a device borrowed from one branch of the orthodoxy to which Monetarism was opposed. As a practical matter, rather than replace it with some alternative equation, the Phelps-Friedman critique of the Phillips curve simply added the expected inflation rate to the relationship's right-hand side, with a coefficient of unity.

As with Monetarist propositions about the demand for money function, then, questions raised by the Monetarist "correction" to the Phillips curve seemed to be inherently empirical. Either the labor force was immune to money illusion in the long run, so that there was no permanent inflation-unemployment trade-off, or it was not, in which case such a permanent trade-off existed. This question generated much empirical work in the late 1960s and early 1970s, and it is a fair generalization that as more and more evidence was added by the passage of time, the more the work in question came to support the Monetarist position. Evidence from the 1950s and early 1960s was compatible with the existence of an inflation-unemployment trade-off, but the experience of higher inflation rates from the mid-1960s onward made this hypothesis harder and harder to support.\(^5\) Once again, though, there was no reason why this result could not be incorporated into the framework of a suitably extended IS-LM model. However, two factors militated against so simple and harmonious an outcome to this stage of the Monetarist controversy.

To begin with, the original Phelps-Friedman critique of the Phillips curve had been advanced on theoretical grounds, and before empirical evidence seemed to require economists to change their notions about labor market behavior. When confronted with a choice between what empirical evidence seemed to show, and what elementary economic theory required to be the case, Phelps and Friedman had chosen the latter. \textit{Ex-post}, they were vindicated by empirical evidence, and this vindication served notice on macroeconomists that they would be wise to pay more attention to the microeconomic foundations of their empirical generalizations than had typically been the case up until then. Second, Friedman's version of the critique had been accompanied by a brief account of labor market behavior, part of which was quite incompatible with the then conventional interpretation of real income and employment fluctuations as manifestations of variations in excess demand in the economy.

\(^5\)Phillips' original analysis was much elaborated by Lipsey (1960), but responsibility for explicitly treating the Phillips curve as a structural relationship constraining policy choice probably rests with Samuelson and Solow (1960).

\(^6\)The relevant empirical literature was voluminous, but is surveyed by Laidler and Parkin (1975).
Conventionally enough by the standards of the 1960s Friedman described the early stages of the economy's response to a monetary expansion in the following terms.\footnote{Friedman's (1969) Optimum Quantity of Money reprints many of his seminal contributions, and the page references here and elsewhere to quotations from his work are to this source unless otherwise noted.}

...much or most of the rise in income will take the form of an increase in output and employment rather than in prices. People have been expecting prices to be stable, and prices and wages have been set for some time in the future on that basis. It takes time for people to adjust to a new state of demand. Producers will tend to react to the initial expansion in aggregate demand by increasing output, employees by working longer hours, and the unemployed by taking jobs now offered at former nominal wages. This much is pretty standard doctrine. (p. 103, my italics)

However, Friedman immediately went on to elaborate this account of the mechanisms that brought about short-term fluctuations in income and employment:

Because the selling prices of products typically respond to an unanticipated increase in nominal demand faster than prices of factors of production, real wages received have gone down—though real wages anticipated by employees went up, since employees implicitly evaluated the wages offered at the earlier price level. Indeed, the simultaneous fall \textit{ex-post} in real wages to employers and rise \textit{ex-ante} in real wages to employees is what enabled employment to increase. (p. 103-04)

Unemployment fluctuations in the conventional view of the Phillips curve had been treated as manifestations of variations in the pressure exerted by excess demand in goods and labor markets characterized by less-than-perfect price flexibility, and hence likely to be operating out of equilibrium in the wake of any shock to aggregate demand. The first passage quoted above is quite compatible with this view. In the second passage quoted, however, employment fluctuations are explicitly pictured as arising from voluntary decisions taken in response to price changes and, in the case of the suppliers of labor, on the basis of faulty expectations which would in due course be corrected.

Thus Friedman's critique of the Phillips curve potentially involved much more than the addition of an extra variable to the right-hand side of an equation. It also pointed toward a fundamental reinterpretation of the labor market behavior underlying it. Viewed with hindsight, it provides as an unmistakable sketch of the microfoundations of the short-run aggregate supply curve which was to become the central analytic device of New-classical economics. As we shall see below, this device would in due course, and quite paradoxically, not strengthen but thoroughly undermine the very Monetarist explanation of the business cycle to which Friedman's analysis of inflation-unemployment interaction was particularly addressed.

The analysis of inflation-unemployment interaction was by no means the only area in which, during the 1950s and 1960s, macroeconomists were seeking to strengthen the microeconomic foundations of their analysis. A whole set of questions concerning the determination of the money supply, and the mechanisms whereby monetary changes might interact with aggregate demand in the economy, were also addressed in such terms, both by adherents of the conventional Keynesian macroeconomic wisdom of the time, and by Monetarists. Here as elsewhere, the contentious issues were more empirical than theoretical. As Harry Johnson noted as early as 1962, there was no debate in principle between, say, Brunner and Meltzer on the one hand and James Tobin on the other about the basic nature of the linkages between the monetary and real sectors of the economy. Both saw these as involving disturbances to the structure of the portfolios of the banking system and the non-bank public generating changes in the relative rates of return on various assets, financial and real, which would in turn provoke attempts on the part of agents to restore equilibrium by way of sales and purchases of various assets. Both sides also agreed that financial institutions and the non-bank public alike should be analyzed as maximizing agents.

What defined the Monetarist position here was not its general approach, but rather a set of specific hypotheses about the quantitative nature of the responses in question. First, as far as the non-bank public was concerned, Monetarists took a broad view of the array of assets whose rates of return were relevant to what we might term the transmission mechanism of monetary policy. Specifically, they argued that monetary policy would have effects not just on rates of return borne by financial assets, but also on the implicit rates of return yielded by producer and consumer durables. Hence they saw its effects as being both more pervasive, and quanti-
tatively more significant too, than did proponents of the Tobinesque "new view." Furthermore, while not denying that disturbances originating in the private sector of the economy would impinge upon the behavior of financial institutions, so that the quantity of money was undoubtedly in this sense an endogenously determined variable, they nevertheless strongly resisted the idea that the endogeneity in question also involved that variable being passively demand-determined, even when the monetary authorities used short-term interest rates as their policy instrument.

Such an outcome was logically possible, to be sure. If interest rate changes disturbed only the margin between financial assets (let us call them bonds) and money, then a reduction, say, in interest rates would lead to the public simply offering bonds to the banks in exchange for money with no further consequences. However, consistent with their broad view of the range of assets relevant to the transmission mechanism, Brunner and Meltzer argued that the principal margin likely to be disturbed by a change in interest rates was that between bonds and physical capital, including consumer durables. The public would sell bonds to the banks, of course, but as part of a process of replacing those bonds with physical capital. Furthermore, and crucially, this would be only a first-round effect, for the sale of bonds to the banks would be in exchange for newly created money which would have to be accommodated in the portfolios of the non-bank public. Since the banks had presumably changed interest rates for a reason in the first instance, the possibility of their simply acquiescing in the destruction of newly created cash by the public discharging debts to them could be discounted. Hence, further substitution effects, changes in aggregate demand, and ultimately in the price level, would be set in motion.

Monetarists' theoretical position vis-à-vis the behavior of the quantity of money in circulation, then, may be described succinctly as follows. Rather than being a passive demand-determined variable, money also had a separate supply function which was derivable from analysis of the interaction of banks and the public in the market for bank credit. In turn, the quantity of credit which banks were willing to grant, as well as the terms on which they would make it available, were both subject to a strong (though not unique) influence flowing from the quantity of reserves which the central bank made available to them. The quantity of money was an endogenous variable, certainly, arising from a complex set of interacting portfolio choices, but it was nevertheless controllable by the monetary authorities. This theoretical position, moreover, was supported by a good deal of empirical evidence, some yielded by formal econometric studies, and some by less-formal historical work to the effect that the behavior of bank reserves, or more precisely of the quantity of high-powered money, was the principal determinant of the quantity of money in circulation and that variations in the ratio of high-powered money to the money supply proper could be modeled as the outcome of systematic maximizing portfolio choices.*

**THE SUBVERSION OF MONETARISM BY NEW-CLASSICAL ECONOMICS**

In the previous section of this essay, I have described the main elements of Monetarist doctrine, and have tried to show that they were all reasonably well-developed in the academic literature by the end of the 1960s. However, it took the inflationary experience of the early 1970s, particularly in the United States and Britain, to draw popular attention to that doctrine by providing, during its early stages, something as close to a controlled experiment as one ever gets in economics. In both countries, monetary expansion preceded an inflation which the authorities attributed to cost-push forces and attempted to control with wage-price control programs, and in both countries those programs failed. Though the OPEC-led energy price increases of 1973 contaminated the later stages of the experiment by introducing an extraneous cost-side impetus to the upward progress of prices, the early 1970s was nevertheless the last time that wage-price controls were deployed as an alternative to monetary restraint in an anti-inflation policy. Moreover, the fact that the high

---

*Perhaps the best known of early econometric studies of the money supply function is Brunner and Meltzer (1964). The standard historical study is that of Cagan (1965). Brunner's (1968) St. Louis Review piece contains perhaps his single best exposition of the Monetarist position on the generation of the money supply, and of the characteristics distinguishing it from the Tobinesque "new view."
inflation of the 1970s was combined with an obvious deterioration in real economic performance, rather than an improvement, ensured that Monetarist ideas about the absence of a long-run inflation-unemployment trade-off quickly became conventional wisdom.

At the very time at which Monetarist ideas were gaining popular acceptance, however, their academic foundations began to shift dangerously as New-classical economics was developed. I have suggested elsewhere that the work of Robert E. Lucas, Thomas Sargent, Neil Wallace and Robert Barro is more usefully treated as separate and distinct from Monetarism, and for what I still regard as good reasons. However, this was and remains, something of a minority viewpoint; and the fact remains that the two central characteristics of New-classical economics, namely the interpretation of the Phillips curve as a reflection of an aggregate supply relationship, and the rational expectations hypothesis, were quickly adopted by leading Monetarists. This was unfortunate, because as the Monetarist controversy moved into the 1970s and 1980s, it increasingly became, as a matter of fact, a controversy about New-classical economics. The majority of economists failed to distinguish between Monetarism and New-classical economics, and accepted James Tobin's characterization of New-classical economics as Monetarism Mark II.

When New-classical economics was academically discredited, so too was Monetarism in general, leading to what, as I shall argue in the final section of this paper, is a dangerous gap in the structure of contemporary monetary economics.

I have already noted that Friedman's 1968 analysis of the inflation-unemployment trade-off relied on two incompatible theories of labor market behavior. In the passages quoted earlier, we had quantities moving instead of wages and prices, which had already been set and hence were unable immediately to change; and we also had quantities moving in response to asymmetrically perceived money-price changes. We had, in short, both an informal account of the effects of wage and price stickiness which was, as Friedman said, "pretty standard doctrine" in the macroeconomics of the 1960s; but we also had an unmistakable sketch of an aggregate supply curve interpretation of the short-run Phillips curve in which money-wage and price flexibility was of the essence.

There is little point in speculating about the extent to which Friedman was aware of the tensions inherent in his analysis in 1968; but it is interesting to note that, at that time, he characterized Phillips' work as "...containing a basic defect—the failure to distinguish between nominal wages and real wages..." (p. 102, Friedman's italics) whereas in 1975 he referred not only to this point, but also attributed to Phillips an error in having "...taken the level of employment [instead of the rate of change of prices] as the independent variable..." in the relationship. By 1975 Friedman clearly was aware that there was a choice to be made concerning the microeconomic underpinnings of the Phillips curve and had explicitly rejected that which hinged on interpreting the unemployment rate as a proxy variable for some excess demand for labor concept. Brunner and Meltzer made a similar choice concerning the modeling of the linkages between output and price-level variations in designing the basic framework which they and their associates used to analyze the inflationary process in a number of countries in the mid-1970s. Like Friedman too, they chose to model expectations as being formed not adaptively but, following Lucas and Sargent and Wallace, rationally, as the inflation forecast of the "true model" of the economy under analysis.

These developments seemed at the time to be in no sense revolutionary. I have already remarked that the Phelps-Friedman critique of the Phillips curve initially involved using simple microeconomic analysis to mount a theoretical attack on what at the time seemed like an hypothesis well-supported by empirical evidence; and in the event, microeconomic principles proved a sounder guide than what turned out to be some misleadingly special observations. To show that price-output interaction could be derived from a supply and demand apparatus without resort to purely empirical generalizations about price stickiness, and to show that the analysis in quas-

---

9 This argument is developed in some detail in Laidler (1982), Ch. 1.

10 This characterization of Tobin's is developed and defended by him in Tobin (1981), in a paper prepared for the same conference as the above mentioned Laidler (1982), Ch. 1.

11 The relevant work here is contained in Brunner and Meltzer, eds. (1978). Note that the details of their formulation of the aggregate supply curve, with output's rate of change rather than level playing a major role, set it apart from the standard Lucas (e.g., 1973) formulation. These matters, discussed by McCullum (1978), are not central to the matter under discussion here.
tion was compatible with agents making use of “all available” information in a utility maximizing fashion, simply seemed to be making even more secure the microeconomic foundations of a particular piece of macroeconomics, and hence to be rendering it less prone to excessive dependence on theoretically unsupported empirical observations of a type that had proved so misleading in the recent past.

Also, and crucially, New-classical macroeconomics still seemed to yield Monetarist implications, namely that inflation in the long run was a monetary phenomenon, and that, in the short run, so was the cycle. To be sure, it broadened the menu of monetary rules that would enable the cycle to be avoided to any that the general public could understand and therefore use as a basis for expectations formation, but a constant growth rate rule was still a particularly simple, and therefore viable, item on the menu in question, and it was of course a particularly appropriate choice if price stability in the long run was added to the elimination of the cycle as a proper goal for monetary policy.

Even in the mid-1970s, it should have been apparent that the attempt to underpin the Monetarist position with New-classical foundations was dangerous. As early as 1958, Friedman had suggested that monetary policy affected the economy with a “long and variable” time lag. The evidence with which he supported this suggestion identified a change in monetary policy as a change in the rate of growth of the quantity of money in the economy: and it showed that downturns in that rate of growth occurred on average 16 months before the corresponding upper turning point of the cycle, while upturns in money growth led cyclical troughs by about twelve months. Furthermore, as Friedman (1987) was later to note more explicitly, the effect of changed money growth on nominal income

...typically shows up first in output and hardly at all in prices. If the rate of monetary growth increases or decreases, the rate of growth of nominal income and also of physical output tends to increase or decrease about six to nine months later, but the rate of price rise is affected very little...the effect on prices comes some 12 to 18 months later... (p. 17)

In 1963, Friedman and Schwartz had presented a more elaborate analysis of money’s role in generating the cycle, in the course of which they explained the length of the time lags involved in terms that amounted to an informal sketch of a dynamic version of the Monetarist version of the transmission mechanism which Brunner and Meltzer were also expounding at the time.

The central element in the transmission mechanism...is the concept of cyclical fluctuations as the outcome of balance sheet adjustments, as the effects on flows of adjustments between desired and actual stocks. It is this interconnection of stocks and flows that stretches the effects of shocks out in time, produces a diffusion over different economic categories, and gives rise to cyclical reaction mechanisms. The stocks serve as buffers or shock absorbers of initial changes in rates of flow, by expanding or contracting from their “normal” or “natural” or “desired” state, and then slowly alter other flows as holders try to regain that state. (p.234)

The empirical evidence referred to here, and particularly that part of it dealing with the timing of output responses relative to those in prices, and an explanation of that evidence in terms of a transmission mechanism involving portfolio disequilibria working themselves slowly out over real historical time are quite incompatible with New-classical analysis. Since both the evidence in question and the above explanation of it were available and well established before the development of New-classical ideas, the incompatibility in question ought to have prevented those ideas from being adopted as a basis for Monetarist propositions, but it did not. Whether this was because the problem was not fully appreciated at the time, or because that shift in methodological priorities away from empirical evidence and toward “sound” theoretical foundations upon which I have already commented caused those who were aware of it to opt for the latter is hard to say. I suspect that a strong element of the latter consideration must have been at work, though, since the inconsistency in question is hardly subtle.

To begin with, the price flexibility postulate of New-classical economics creates problems for the

---

12The reader’s attention is drawn to the recent (1987) vintage of this statement. I have not been able to find so clearcut and concise an exposition of the point in Friedman’s earlier work, though I believe that the basic message contained in the passage quoted here can be distilled from the evidence presented in Friedman (1969), Chs. 10-12.

13Though I do not claim to have understood this matter fully from the outset, I did discuss it in some detail as early as 1978 in an essay reprinted as Chapter 4 of Laidler (1982).
Monetarist account of the transmission mechanism. Slow adjustment of portfolios in the wake of a monetary disturbance is of the very essence here, and it is usual to explain the slowness in question in terms of transactions costs. But if the price level moves freely and instantaneously to keep markets cleared it also, in the process, eliminates portfolio disequilibria quite costlessly for agents. According to Friedman and Schwartz, excessive money holdings develop because, when money growth increases, the rate of inflation does not respond immediately. But according to New-classical economics it does, and so increased money growth cannot cause a temporary rise in buffer-stocks of money as a preliminary to increased expenditure flows. The conflict here between traditional Monetarism and New-classical analysis concerns rival theoretical constructions. Much more serious is the conflict between theory and evidence which arises when the New-classical aggregate supply curve is confronted with the empirical evidence concerning the interaction of money, output and prices over the course of the cycle, evidence upon which traditional Monetarism laid considerable stress.

As is well known, the New-classical aggregate supply curve explanation of the decomposition of nominal income changes into their real and price-level components hinges upon the distinction between demand side shocks whose price level effects can be anticipated by agents, and those that cannot. The word “anticipated” normally means “expected and acted upon,” but since the New-classical model is one in which flexible prices always move costlessly and instantly to equate supply and demand in all markets, the second phrase is redundant in its context. If a price level change is anticipated by agents, there will be no quantity changes associated with it; but if it is not, then the specific money-price changes in particular markets that are associated with it will be misinterpreted as reflecting relative price changes and voluntary responses in quantities of goods and services supplied will occur. Cyclical fluctuations in real variables may therefore be interpreted as the consequence of unanticipated price level changes. The trouble here is that it is hard to see how output and employment fluctuations can be responses to price level fluctuations if they precede those price level fluctuations; but the empirical evidence tells us that they do just that.

This inconsistency of the timing of data with the basic structure of New-classical theory, which should have led monetarists to reject that theory from the outset, did eventually undermine it. So long as empirical work was confined to testing the proposition that output and employment fluctuations could be modeled as responses to “unanticipated” money, all seemed well. However, Robert J. Barro (1978) noted that the theory in question made specific predictions about the relationship between monetary shocks and price level changes as well. When he came to test the latter predictions, he found that, in order to reconcile them with his data, the price level’s response to unanticipated monetary shocks had to be characterized by a rather slow-moving distributed lag process, in an economy in which, however, prices could respond with no lag to anticipated money. Furthermore, his results also seemed to require that the aggregate demand for money display a greater sensitivity to transitory than to permanent changes in income, the very opposite result to that implied by a wide variety of other studies. New-classical analysis could be forced to fit the data generated by the U. S. economy, that is to say, only by way of some extremely implausible subsidiary assumptions.

Nor did the results of subsequent empirical work enhance New-classical economics’ claim to be taken seriously. Mishkin (1982) repeated Barro’s tests of the irrelevance of anticipated monetary shocks for output using more sophisticated econometric techniques, and found that this resulted in a reversal of the initial results—apparently anticipated monetary changes did have real effects. Boschen and Grossman (1982) noted that money supply estimates are published weekly, and that only the errors in these estimates properly can be regarded as constituting the unanticipated component of the money supply. They further noted that the latter were implausibly small to form the basis of a monetary explanation of the cycle, and that output changes were in fact correlated with monetary changes about which information had previously been published.

In addition to these problems raised by academic work, of course, there was the experience of the early 1980s recession, which played the same role in publicly discrediting New-classical economics, as did the inflation of the 1970s in undermining “Keynesian” theory. New-classical economics discounted the importance of real world wage and price rigidities, and placed considerable faith, therefore, in the public’s willingness to react immediately to a well-publicized
anti-inflation policy based on monetary contraction in such a way as to reduce inflation without serious real income and employment consequences. The policies in question certainly did reduce inflation, but the recession which accompanied that reduction was, in some respects, the worst since the 1930s. Moreover, the United Kingdom and Canada carried out similar experiments at about the same time with similar results. By the mid-1980s, the New-classical development of the Monetarist account of the role of money generating the business cycle was thus widely recognized to have failed in its encounter with empirical evidence. This could have led to a revival of interest in more traditional Monetarist analysis; but it did not, because the very postulate that had established the respectability of that analysis in the first place, namely a stable aggregate demand for money function, had also run into difficulties.

THE KEYNESIAN ADOPTION OF THE STABLE DEMAND FOR MONEY FUNCTION

The early studies of the aggregate demand for money function, which established Monetarism's respectability, were carried out using long runs of U.S. data, some stretching back into the 19th century. Moreover, the data themselves were highly time aggregated. Thus Friedman's (1959) seminal study covered the years 1869-1956, and used cycle averages of variables in estimating its basic equation. One business cycle, that is to say, lasting on average about four years, provided one observation. Later studies, such as those of Meltzer (1963) or Laidler (1966) dealt with essentially the same time period, but used annual observations. It was an obvious enough extension of such work to test the hypotheses at stake in it against data drawn from other countries and also against more time disaggregated data too, and such extension proceeded apace in the 1960s mainly at the hands of people far more interested in exploiting new data and computing techniques than furthering any particular policy agenda.

At first the hypotheses in question—that the demand for money varied with some real income or wealth measure, some measure of the opportunity cost of holding money, and in proportion to the general price level—displayed remarkable robustness; so much so that, by the early 1970s, the demand for money function was a prime candidate to become the centerpiece of stabilization policy, much as had the Keynesian consumption function or the Phillips curve at earlier times. For the demand for money function's full potential for such a use to be exploited by policy makers, detailed knowledge of the function's contemporary form was obviously needed, and in the early 1970s a remarkably simple version of the equation seemed to be able to deal with quarterly U. S. data with a high degree of precision. This relationship, nowadays known as the "Goldfeld equation" after its most careful exponent (Goldfeld, 1973), had the long-run average value of money holdings determined by real income and interest rates, but involved the hypothesis that when some disturbance took money holdings away from this long-run average, they would move back toward it slowly over time, with the speed of adjustment in question being proportional to the size of the gap to be closed.14 The Goldfeld equation fitted U.S. data well, appeared to be stable over time, and crucial for policy purposes, dealt with data at a relatively low degree of time aggregation. It did indeed appear to be so policy relevant that, in his Presidential address to the American Economic Association, Franco Modigliani (1977) argued that it could, and should be used as the basis of an activist monetary stabilization policy in the Keynesian mold.

Modigliani's address in fact appeared after the first signs of trouble with the relationship had appeared. Before dealing with that, however, it is worth reiterating that the progress of the demand for money function so briefly dealt with above was by no means synonymous with the progress of Monetarism, but rather it involved a component of Monetarism being taken over by the Keynesian opposition. Just as the association of Monetarist ideas about the cycle with those of New-classical economics was to prove destructive, so too was this association, and once more this could have been, but was not, discerned at the time in the light of evidence then available. That the association in question was indeed being formed was, of course, obvious enough once Modigliani made the existence of a

14 Though the relationship in question is now irrevocably known as the "Goldfeld Equation," it was in fact used before him in studies of the demand for money by, among others, Teigen (1964) and Chow (1966). Goldfeld, be it explicitly noted, did not claim originality for the relationship in question.
stable demand for money function an important part of the basis of his case for monetary fine-tuning, a policy stance to which Monetarism was root and branch opposed; and I am not here claiming otherwise. However, I am also claiming that the equation fitted by Goldfeld, and the interpretation he put upon it, were quite antithetical to earlier Monetarist ideas about the demand for money function in particular, and the role of money in the economy in general.

To begin with, the demand for money function for which Friedman had initially claimed stability was the "long run" relationship. In 1959 he had tested it against cycle average data, among other reasons, in order to abstract from the complex interactions among money, output, prices and interest rates that characterized the cycle, and which would tend to obscure the underlying stability of the relationship in question. Nor was such a procedure a quirk of one particular paper: the empirical work of Friedman and Schwartz's Monetary Trends..., though not published until 1982, was largely completed in the late 1960s and was based upon cycle phase average data. The apparent stability of demand for money functions such as Goldfeld's, which used quarterly data and hence were dominated by within-cycle interactions among variables, should have been a source of puzzlement to Monetarists, therefore, not of satisfaction.

Monetarists should also have seen that Goldfeld's interpretation of his equation ran quite counter to their ideas about the transmission mechanism.15 He used the money supply as the dependent variable of his relationship, and treated it as responding passively, albeit with a distributed lag, to variations in the arguments of his demand for money function. It is of the very essence of the Monetarist view that there exists a supply function of money that is independent of the demand function, and that the interaction over time of the money supply, income, interest rates and the price level involves causation running predominantly, though not uniquely, from money to the other variables. Goldfeld's work on the demand for money, and many other studies in the same vein, were thus based implicitly on the "new view" of the money supply process discussed above, of which Brunner (1968) had been so critical. And indeed, in the 1970s, as central banks became interested in controlling the time path of monetary aggregates, their procedures involved measuring or forecasting the values of all the right-hand-side variables of a Goldfeld-style equation except the interest rate, and then setting the latter in order to achieve a value of the quantity of money demanded equal to the money supply target. Clearly such a procedure left no room for taking account of the subtle interactions among markets for money, credit and equity that lay at the heart of the Monetarist view of the matter.

Be that as it may, the stability of the Goldfeld short-run demand for money function that ought to have puzzled Monetarists did not last long. By 1976 he was inviting his readers to solve "The Case of the Missing Money," while by the early 1980s, a number of commentators were contemplating an unexplained decline in money's velocity of circulation.16 Nor were problems with the demand for money function a uniquely American phenomenon; Canada and the United Kingdom too had problems with badly behaved demand for money functions in the same years; so did Australia and New Zealand a little later; and this list is far from exhaustive. There is little point here in attempting a survey of the voluminous literature generated by these events. Suffice it to say that a wide variety of ad hoc explanations, often relying upon particular institutional changes were proposed, and often (not always) proved fragile. The upshot was that in the eyes of the majority of commentators the postulate of a stable aggregate demand for money function was discredited, and along with it the very foundation of Monetarism.

15The first sign of discomfort about this matter that I can find in my own writings appears on pp. 143-44 of the second (1977) edition of my Demand for Money, where I refer there being a fallacy of composition involved in proceeding from the individual to the market experiment when analyzing adjustment processes in the demand for money function. Chapter 2 of Laidler (1982) developed my doubts about all this in much more detail.

16And this decline in velocity was associated with rapid money growth failing to produce renewed inflation in the recovery from the 1982 recession. Monetarists should not have predicted renewed inflation then, because actual and expected inflation fell dramatically in the preceding downswing, and hence should have led to an increase in the demand for real balances which could be met only by a falling price level or a growing nominal money supply. What ought to have been done, and what was done, however, are different matters, and careless Monetarist predictions at this time did help to discredit the doctrine.
Of course this upshot was preposterous. The stability that Monetarism had from the outset attributed to the demand for money function was long run in nature, and the relationships that collapsed were short-run formulations, espoused by Keynesian economists intent on establishing a basis for a policy of monetary fine-tuning. Moreover, those relationships were based on a view of the money supply process which Monetarists had vigorously opposed from the earliest stages of the controversy. An alternative inference to be made from the collapse of the empirical stability of short-run demand for money functions from the mid-1970s onward was that this was the result of their failure to model the dynamic relations among the quantity of money, interest rates, real income and prices as the outcome of the interaction of an independent supply of money function with the arguments of the demand function; that the problem stemmed not from instability of the latter relationship at all, but from the inability of simple single equation distributed lag techniques to come to grips with the dynamic complexities involved in the transmission mechanism of monetary policy.

Those of us who advanced the latter explanation, however, did not find much of a sympathetic audience, even among Monetarists. No doubt this was partly because we did not make our case as clearly as we might have done. But it was mainly because the case in question had as a key component the notion that markets failed to clear instantaneously in the manner demanded by New-classical economics. Rather, it was argued that the transmission mechanism worked along the portfolio disequilibrium lines sketched by Friedman and Schwartz (1963) in the passage quoted on page 56. In the wake of the success of the Phelps-Friedman critique of the Phillips curve, Monetarism had come to attach great importance to adhering to “sound” microeconomic foundations, so much so that it had, as we have seen, become intertwined with New-classical economics. Rather, it was argued that the transmission mechanism worked along the portfolio disequilibrium lines sketched by Friedman and Schwartz (1963) in the passage quoted on page 56. In the wake of the success of the Phelps-Friedman critique of the Phillips curve, Monetarism had come to attach great importance to adhering to “sound” microeconomic foundations, so much so that it had, as we have seen, become intertwined with New-classical economics.

New-classical economics was underpinned by a strong methodological preference on the part of its exponents for grounding macroeconomic theorizing on explicit microeconomic foundations. Such foundations, however, are capable of yielding a wider variety of macro models than the monetary explanations of the business cycle that lay at the heart of the work of Lucas and Sargent and Wallace. Thus the empirical failure of those models did not lead to the aban-
donment of the methodological agenda which had produced them, but merely to an attempt to replace them with an alternative explanation of the cycle with equally well, or even better, defined micro premises. I refer here to that body of research known as “real business cycle theory” which is based on a stochastic version of the New-classical growth model of Meade, Swan and Solow, and attributes cyclical fluctuations to exogenous shocks to the aggregate production function, in much the same way as, over a century ago, Jevons attributed the cycle to fluctuations in agricultural productivity associated with sunspot activity.

The exponents of real-business cycle theory, though hostile to econometric testing, nevertheless do not ignore empirical evidence, and have begun to address the question why, if it is not a causative factor in cyclical fluctuations, there are nevertheless systematic relations among the quantity of money and other variables. The very manner in which the question is posed virtually dictates the answer offered, namely that the relations in question are the result of reverse causation running from the cycle to money, rather than vice versa. Thus, in what is surely one of the greater ironies in the recent history of economics, a research agenda which is widely regarded as a direct descendant of the Monetarism of the 1960s has ended up adopting a view of the role of money in the economy directly opposed to that of its intellectual antecedent, and virtually identical to that of the most extreme form of “post-Keynesian” economics. This is no accident, for a view of the world which has markets functioning perfectly and without friction leaves no more room for money to play an important role than does a view in which markets do not function at all.

Though one important group among the academic heirs of Monetarism has thus systematically adopted hypotheses that downgrade the importance of monetary phenomena, and has also abandoned traditional econometric methods as a basis for empirical research, this does not mean that econometric work on monetary economics ceased in the early 1980s. On the contrary, a diverse body of contributors, including unrepentantly old-fashioned Monetarists, econometricians in search of an area in which to try out new techniques, not to mention economists associated with central banks in various countries which do, after all, still have to carry out monetary policy regardless of the state of academic opinion concerning its importance, have generated an extensive empirical literature on the demand for money function during the 1980s. The literature in question has been lively and, as I shall now argue, productive in two lines of inquiry.

First, questions arising from the fact of institutional change in the monetary sector have been examined using both historical and contemporary data. In both cases, this phenomenon has been found to be sometimes important. Thus Bordo and Jonung (1987), using data going back to the 19th century for five countries, have shown that the slow decline in velocity which occurred largely before the first world war seems to have been associated with the increasing degree of monetization of the economies involved, and its later rise with the increasing efficiency of monetary exchange. Closer to our own time, the sharp increase in the velocity of M1 balances in Canada that occurred at the turn of the 1980s has been shown, beyond reasonable doubt, to have been associated with the simultaneous spread of daily interest checking accounts. As to the United States, the work of Barnett (e.g., 1990) using Divisia aggregates, provides evidence that various instances of financial deregulation have produced shifts in the demand for more conventionally measured aggregates; while the very fact that it proved necessary not so long ago to redefine those aggregates is itself testimony to the importance of institutional developments.

The other line of inquiry to which I refer above has involved the application of far more sophisticated techniques than were previously available, both to the explicit modeling of the so-called short-run demand for money function, and closely related, to the task of extracting information about the underlying long-run relationship from data with a high degree of time disaggregation. Here the results have been quite startling. Short-run dynamics dominate quarterly data and have to be modeled with a great deal more care than exponents of, say, the Goldfeld equation brought to bear or, given the state of econometric technique, could have been expected to bring to bear, on the task. Never-

---

18I refer here to King and Plosser (1984) which uses a form of the old “money-income-causes-money” hypothesis to reconcile monetary phenomena with a productivity-shock theory of fluctuations in real variables.
theless, once this is done, stable long-run relationships, very much like those which were first estimated by Friedman (1959), Meltzer (1963) or Laidler (1966), are after all to be found buried in those data. This same result emerges with powerful simplicity from a recent "low-tech" study by Robert E. Lucas (1988), who shows that data for the last 25 years or so of United States history are still scattered (albeit widely, and with complex serial correlation) around a velocity function directly derived from Meltzer’s (1963) estimates of the long-run demand for money. Nor are results of this sort confined to the United States. Similar conclusions arise when British, Canadian or Japanese data are analyzed.

As yet, this empirical evidence has not attracted the academic attention it deserves, largely, I suspect, because the treatment of the short-run dynamics in the studies which generate it has been based more on econometric technique than economic theory. Exponents of the so-called “buffer-stock” approach to modeling the demand for money, who have of course followed up the analysis of the transmission mechanism of Brunner and Meltzer as well as Friedman and Schwartz, have tried to bridge this gap. They have tended to ground their explicit analysis on rather simple special cases, however, which have not proved empirically robust; and this in turn has led to an identification of the general approach with those special cases and a tendency to dismiss prematurely the broad insight which it yields: namely that the dynamics of what we have learned to call the short-run demand for money function are not the property of a structural relationship at all, but rather reflect that complex interaction of the quantity of money with other variables to which the Monetarists of the 1950s and 1960s used to refer as a transmission mechanism subject to long and variable lags.19

The Monetarist controversy was about theory and empirical evidence, to be sure, but it was also about monetary policy; and here it has left its strongest mark. To begin with, the idea that inflation is fundamentally a monetary phenomenon, so outlandish in the 1950s, has by now become something close to conventional wisdom. We nowadays hear very little about cost-push forces and the need to control them with wage and price guidelines, controls and so on. Closely related, central banks routinely pay attention to the behavior of monetary aggregates in designing their anti-inflation policies. Though even such hard-core Monetarists as Brunner and Meltzer (1987) no longer argue that the dominant impulse driving the business cycle is always the quantity of money, neither they, nor the practitioners of monetary policy, have shown the slightest sign of taking the productivity shock hypothesis of the real business cycle theorists seriously.20 Rather, the exclusively monetary interpretation of the cycle has been replaced by an eclectic approach in which monetary factors have an always potentially important, and sometimes an actually important, role to play. That is why the pursuit of price stability by monetary means is tempered by caution concerning the short-term costs that could be incurred if the pursuit in question was to become too vigorous. And eclecticism about the causes of cyclical instability has not been accompanied by a revival of interest in the use of fiscal policy for stabilization purposes—though this probably has as much to do with the fiscal deficits that are the legacy of supply-side economics as with any lessons learned during the Monetarist controversy itself.

A superficial reading of the above record might suggest that, whatever its academic standing, Monetarism is alive and well in policy circles. That it has left a lasting mark in those circles is beyond doubt, but its success on the policy front has been far from complete. Indeed, on one interpretation of the evidence, Monetarism’s partial success here may have been worse than failure. The Monetarist agenda, after all, involved establishing the importance of the quantity of money as a policy variable as a preliminary step to removing from the monetary authorities their discretionary control over it; and the record

---

19For an informal but more complete exposition of the arguments involved here, see Laidler (1987). One of the better known pioneering special case empirical formulations of the approach is that of Carr and Darby (1981), who refer to it as a “shock-absorber” approach.

20Let it be clear, however, that I do not regard “real business cycle theory” as a total waste of time, and that I do think it merits policymakers’ attention. Though I find it hard to believe that shocks to technology will turn out to be the sole or even main source of real world cyclical fluctuations, it is nevertheless the case that, from a theoretical point of view, such shocks do originate in such phenomena are likely to be welfare improving. The valuable message of real business cycle theory, then, is that we should not take it for granted that stabilizing the cycle is always and everywhere desirable. It might not be in particular instances. That reminder is surely worth having.
shows that only this preliminary step was completed. Monetary policy plays a far more important role in macroeconomic affairs now than it did 30 years ago, but those in control of it have, as a result, much more power than previously, not less, as the Monetarist agenda intended. This cannot really be helped, because part of the academic legacy of the Monetarist controversy has been a greater appreciation of the role of institutional change in influencing velocity's long-run behavior; but once that role is admitted, the case for tying down monetary policy with rules becomes far more difficult, perhaps impossible, to make.

All this, though, poses a challenge. If monetary policy cannot be tied down by rules, then the next-best solution is to subject it to the discipline of constant public scrutiny and criticism. Academic economics has a large role to play here, in providing the intellectual basis for such activity, but at the moment it is not in good condition to do so. We do have, as I have noted above, a large and growing body of empirical work which points to the long-run stability (not constancy, note) of money's velocity, institutional change notwithstanding. That same body of work, however, tells us that the short-run dynamics of that function are at least as complex as anyone thought they were 30 years ago: and we are no further forward now than we were then in understanding just why this is the case.

At the same time, the gap between theoretical work on the role of money in the economy, and our empirical knowledge in the area, which was opened up when the Monetarist controversy became a debate about New-classical economics and began to pay more attention to microfoundations than to data, still remains. Intellectual vacuums of this sort rarely remain unfilled for long, but until this one attracts more attention than it has to date, the Monetarist controversy's legacy must be judged to be a distinctly uncomfortable one. The controversy weakened the links between academic research and contemporary policy which in the past made such research a natural source of constraining criticism of the conduct of policy, while simultaneously enhancing the discretionary powers of policy makers. That is hardly what its participants intended.

REFERENCES


_________. The Optimum Quantity of Money and Other Essays (Aldine Publishing Co., 1969).


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


Monetarist Perspectives (Harvard University Press, 1982).


