The Foundations of Monetary Theory
Frank H. Hahn
UNIVERSITY OF CAMBRIDGE

I INTRODUCTION

To the pure theorist, at the present juncture the most interesting and challenging aspect of money is that it can find no place in an Arrow–Debreu economy. This circumstance should also be of considerable significance to macroeconomists, but it rarely is. Much of current macroeconomics is written as if the ‘real’ economy could be looked at as an equilibrium of an Arrow–Debreu economy. It is true that these economists have given the economy a sequential characterisation, but they believe that the postulate of rational expectations renders this inessential: the underlying economy could just as well be that described by Debreu. In this, I believe, they are not only careless (since the matter is never given proof) but also mistaken. It is one of my central concerns in what follows to argue this. Indeed, I shall wish to maintain the view that from our present standpoint Keynes’s theory, and in particular his monetary theory, is best understood as a denial of the realism and relevance of Arrow–Debreu equilibrium. Of course, Keynes did not and could not have put it in this way and in any case, while he had a poet’s insight, he lacked the seriousness and care of a theoretician.

In recent years the most notable attempt at a foundation of monetary theory was that of Patinkin (1956). There cannot be many economists who have not been profoundly, and rightly, influenced by this work. Nonetheless, it left a number of important questions unanswered and in one particular aspect it did not go nearly far enough. It was, of course, also largely written before the great advances in general equilibrium theory. Although Patinkin did develop the beginnings of a
theory of non-Walrasian equilibrium he did not complete it, and although he had much of interest to say on the underlying determinants of the demand for money, he did not establish in any rigorous way the viability of a monetary economy. The world which he considered was one of perfect competition. This, I shall argue, precludes a number of very important matters in monetary theory from being considered. He, by and large, assumed away all distribu-
tional aspects so that the economy could be looked at as if it consisted of only one household. He did not study the full asset equilibrium of an economy nor its evolution over time. Expectations were single valued and arbitrary. In short, although Patinkin clearly pointed the way, he fortunately left much for his successors to do. In the sequel some of this unfinished business will be considered.

Through much of the history of monetary theory there runs one consistent theme: the view that the nominal quantity of fiat money should not matter; that is, the real equilibria of the economy should be invariant to the magnitude. It is one of the paradoxes of present circumstances that those who hold this doctrine most fervently are also those who are most vociferous in their demands that the nominal money stock be closely controlled. But perhaps there is no paradox here at all. Perhaps these two views can be reconciled within a distinction between the short and long run, perhaps within a distinction between equilibrium and disequilibrium. We shall have to consider this, but in doing so we must not expect too much help from the current literature. It is true that Lucas (1972) has given a plausible answer: variation in the money stock can affect the real equilibrium by making it impossible to disentangle ‘real’ from purely ‘nominal’ disturbances. But to reach this conclusion he needed to suppose that the money stock itself was not directly observable. This is a very implausible assumption. In any case, Lucas set the tone and there are now quite a number of accounts of how nominal magnitudes can confuse signals and thus have real effects. There are also some fairly obvious accounts of how lack of short period price flexibility can lead nominal magnitudes to exert short period real effects. Very little of all of this has been formulated at any convincing level of generality and all of it takes Walrasian equilibrium as the appropriate point of departure. This equilibrium is always taken to be unique. Certainly the Neutrality doctrine as well as its exception merit further study.

But although these matters are all central to monetary theory and certainly central to any analysis of monetary policy, there are other
problems connected with money which while, perhaps at first sight, somewhat academic nonetheless require answers. In particular, I have here in mind the old conundrum of how fiat money can survive as an institution. By this I mean the question of what it is that ensures that money has at all times a positive exchange value.

At a common-sense level almost everyone has an answer to this, and old-fashioned textbooks used to embroider on some of the banalities at great length. But common sense is, of course, no substitute for thought and certainly not for theory. In particular, most of the models of an economy which we have, and I am thinking here of many besides those of Arrow and Debreu, have no formal account of the exchange process. The recognition that this was the case was the point of departure for Clower (1965) and Leijonhufvud (1968) who had disequilibrium situations in mind, although Clower especially attempted to model explicitly what he took as axiomatic, namely ‘that only money buys goods’. This axiom, however, was not underpinned by an account of the process of transaction nor was the axiom itself subjected to extensive scrutiny. While it was an advance to take such explicit cognisance of exchange constraints, the offered solution is more like a sledge hammer than a scalpel. One still wants to ask: why does only money buy goods? I suppose that questions like this are real ‘foundation questions’ and I shall start with them. They are of interest for the following reason: in trying to answer them we are led to build features of the economy into our models which we might otherwise have neglected, and these features in turn will have consequences beyond the immediate desideratum of accounting for the viability of fiat money.

I have now taken note of many of the issues which I propose to discuss. Before I do so, it may be useful to attempt a synoptic formulation of what I conceive to be my task.

By monetary theory I shall understand any theory of the economy which includes a means of exchange in its specification. I define a means of exchange as a good which can be exchanged against any other good on terms at least as favourable as a good not designated a means of exchange can be exchanged. In general, although stronger than needed, we can stipulate that the terms using the means of exchange are strictly better than they are when some other good is used. We must enquire into the circumstances in which an economy with a good thus singled out can indeed sustain such singling out. The question will be sharper if the means of exchange is intrinsically worthless. In what follows these matters will be made more precise.
But it is clear that a good which is a means of exchange must necessarily be a means of intertemporal reallocation and thus a 'store of value'.

If this view of monetary theory is adopted then we shall have to consider only those economies in which there are transactions at every date. This rules out the Arrow–Debreu economy. Consequently, some of the propositions based on the latter are at risk. This will be particularly true of the Fundamental Theorem of Welfare Economics. Moreover, the specification of the economy will now essentially have to include agents' expectations of the future values of economic variables. In the Arrow–Debreu economy one is only concerned with beliefs concerning states of nature. In addition there will now emerge an interesting distinction between short- and long-run equilibrium. To pursue monetary theory thus entails large reformulations of our most powerful paradigm.

In addition, there are the traditional questions: the 'classical dichotomy', the 'neutrality' and 'super neutrality' of money. In recent years, in spite of the literature, these matters have been treated rather cavalierly, and rather badly, by the new macroeconomists. They certainly belong here, since so long as macroeconomics is far away from the 'foundations', it is also incoherent.

II THE MEANS OF EXCHANGE

I do not regard it as a very promising enterprise to attempt to give a theoretical account of the origin of monetary exchange. No doubt there is a place for theoretical history but I doubt that economists are very good at it. Indeed, there is some evidence that they are not. Thus one often encounters a demonstration that a medium of exchange may allow an allocation which Pareto-dominates that attainable by barter. It is then an easy step to the proposition that it is this circumstance which explains why money is used. But this Panglossian teleology lacks all merit until a process is described which brings about the superior outcome. Hicks has attempted a historical sketch (1969), but he leaves much unaccounted for.

A preferable route is to take the institution of money as given and to ask why and how it survives. Given all the common sense, not to say banal, answers which come to mind, this turns out to be harder to do at the level of formal theory than one might have thought. A rather simple condition for the survival of money (of no intrinsic worth) is that
the probability of it not being exchangeable against goods is for some agents bounded away from zero. This would then suffice to ensure a positive exchange value for money at all times. It is important to understand two things about this explanation: first, while, as it were, it starts in the mind, it is easy to reconcile with expectations rationally held; in every evolution of the economy the exchange value of money will in fact be positive. Secondly, the explanation implies that when a high enough price level is reached it is expected almost surely to fall. Hence the expectational assumption ensures a positive rate of return on money at a high enough price level.

Grandmont (1974) has shown that an expectational assumption of this kind is needed if one wants to demonstrate that every short period equilibrium of an economy is a monetary equilibrium. Yet I do not believe that their analysis goes far enough. The trouble is that an economy in which there is a medium of exchange is quite different from one without one. For instance, there will be large differences in the transactions and production choices — after all, Adam Smith taught us to make an important connection between the division of labour and the ease of exchange. This is also one of the drawbacks of Patinkin’s approach since formally his description of the economy is unaffected whether money has a positive exchange value or not.

The difficulty can perhaps be best illustrated as follows. One of the oldest questions of monetary theory is why an agent should hold money which may have a zero-expected return when there are other assets whose expected return is positive. The introduction of uncertainty of future prices may fail to provide an answer since the dispersion of the returns on money may be as large or larger than those of other assets. In order, so it is argued, to have money held, one may require the expectation of a falling price level. But even that may fail to resolve the difficulty when there are other assets, like government bonds, denominated in money (Hahn, 1981). But we now notice that it makes a great deal of difference to the question whether the returns on ‘real’ assets are realisable in a monetary economy or not. When I invest all my money in corn which I plant, then the return depends a good deal whether the harvested corn can be sold for money and the money exchanged for goods, or whether I have to hawk corn from the candlestick-maker to the blacksmith. In other words, if corn were really to drive out money then it may no longer be profitable to plant corn. The limit of a monetary economy as the price level in terms of money goes to infinity, is not the economy we started with. To enunciate a paradox: corn can only
drive out money if agents believe that money will not be driven out.

This leads me to the formulation which I favour. I want to think of
an economy in which transactions are mediated by money as being a
Nash equilibrium of an 'exchange game'. It is thus advantageous for
any given agent to mediate his transactions by money, provided all
other agents do likewise. This, of course, is not a novel idea, but its
formalisation is not straightforward. Moreover, there will be other
Nash equilibria which are not monetary so that it is not possible to
dispense with history altogether in this 'explanation' of a monetary
economy.

An agent who lives in an economy where all other agents mediate
their exchange by money will receive money in exchange for the
goods which he sells and must offer money for the goods which he
buys. This is the picture behind the Clower dictum. But it is too
strong. For while other agents may always be prepared to offer and
accept money, it does not yet follow that they would not be ready to
offer and accept something else. What we need is that the given agent
should calculate that these alternative modes of transactions will be
less advantageous than those that are mediated by money. That is,
we need to give some flesh to what Hicks called the 'non-moneyness'
of other goods.

It is here that the traditional account is in need of some correction.
This account invokes the low probability of a 'double coincidence' of
wants, so that an agent wishing to exchange one good directly for
another faces higher transaction uncertainty than he would if he had
mediated his exchange by money. But this view is based on the
presupposition that in every exchange each agent acts as if it were his
last and that is a very strong presupposition. There will be terms on
which a vegetarian agent will accept beef for shoes since there will be
a probability that he in turn will be able to exchange beef for health
foods with someone else. In other words, there is no intrinsic reason
why exchange should not be speculative in this sense if the terms are
right. What one needs is that the vegetarian will accept beef for shoes
on terms less favourable to the beef supplier than the latter could
have had by exchanging beef for money and money for shoes. At
least the terms for direct exchange should not be more favourable
than they are for the indirect route.

The easiest answer is that the vegetarian in accepting beef for his
shoes will have to undertake two transactions to get his health food
instead of one. What the beef supplier saves in transaction (not
selling his beef for money), he has put on the shoemaker. If transac-
tions are costly then the shoemaker will charge a premium for accepting beef rather than money and the beef supplier loses as much as he has gained by saving on the number of his own transactions. Behind this argument in turn is the supposition that the vegetarian will not be able to get his health food as favourably by direct exchange (now against beef) as he would against money.

It will now be seen that transaction costs are not enough. It may simply lead to agents being indifferent between a barter and a monetary deal. Thus the beef supplier would simply pay the vegetarian to undertake the indirect exchange which he has avoided. Some additional element has to be brought into the story.

The missing element, in a way, is to be found in Adam Smith's analysis of the division of labour and the extent of the market. In a monetary economy agents can specialise not only in production, but also in the selling of what they produce. In my example I want to argue that the cost of selling beef (for money) is less for the supplier of beef than it is for the vegetarian shoemaker. For the idea of two agents meeting at random to engage in bilateral exchanges is not one of any descriptive merit and it misses an important feature of a monetary economy. This, of course, is the presence of anonymous markets.

There are quite clearly increasing returns to specialisation in transactions. Thus there is a set-up cost in acquiring the right initial information of where and how best to transact and there are informational economies in being known as someone willing to exchange beef for money. The beef-seller may rely on a middleman who can reap even greater advantages from increasing returns, but he will know more about such middlemen than does the vegetarian. The asymmetric transaction costs will then explain why indirect barter is more costly to the beef supplier than is monetary exchange.

We do not say that there is a market for $x$ if there is simply a probability of bumping into someone who is willing to exchange $x$ for something else on some terms. When we postulate a market for $x$ the very least we can mean is that agents know where to go – to whom to go – to buy and sell $x$. We do not necessarily mean that agents know where to transact most advantageously. Thus, as in the current literature, agents may search for the cheapest store, but they know that it is amongst stores that they have to search. Someone wishing to exchange his house goes to estate agents or advertises – he does not, like some crazed-particle, wait to bump into a buyer.

Monetary exchange is seen to be the cheapest way in which
markets can be organised. For the information – willing to exchange beef for money – is much less costly to convey than a vector of goods for which there is willingness to exchange beef. In general, markets will have middlemen to mediate between buyers and sellers – a simple consequence of the increasing returns to transactions. Once there are middlemen, they must be known to at least one side of that particular market. I know that I can get meat for money at the butcher’s, but I need not know – not being a meat producer – where to sell meat for money. Moreover, if I routinely sell meat, the appropriate middleman has information about Hahn-meat. If I turn up once at the meat wholesaler with a side of beef, I must have considerable transaction uncertainty and surely must be pretty uncertain of the terms which I can get. In short, a monetary economy by greatly economising in information through transaction specialisation leads to an asymmetry in transaction cost between agents. This in turn makes it advantageous for agents to engage in transactions mediated by money rather than barter when everyone else is doing the same.

The literature to date on this ‘foundation’ question of monetary theory seems to me to be rather inadequate. The reason is that it is almost entirely cast in a framework of pairwise exchange without any attention to the market arrangements which monetary exchange makes possible. In particular, what is missing from these accounts is recognition of both the increasing returns from production and from transactions. A barter economy is not the limit of a monetary one – there is a large discontinuity in the kind of economy as one passes from the former to the latter.

Let us look briefly at an example. Consider an economy in which all transactions are mediated by money and by middlemen. The latter exist because they economise in information. A middleman in good \( i \) buys at the money price \( q_i \) and sells at the money price \( p_i \). His money costs are \( a_i + b_i t_i \), where \( t_i \) is the number of transactions which he mediates. The costs per transaction are declining in \( t_i \). If middlemen’s profits are competed away then \( p_i/q_i \) will become closer to unity as the number of transactions mediated increases, provided the amount bought or sold per transaction is approximately constant.

An agent wishing to exchange one unit of good one for good two via the middlemen obtains \( q_i/p_i \) units of good two. Another agent wishing to exchange good two for good one gives up \( p_i/q_i \) units of good two for one unit of good one. If the first agent wants to cut out
the middleman by resorting to barter, he must offer the other agent
(if he can find him) terms at least as good as these. Hence his
maximum gain in good two is

\[
\left[ \frac{p_1 - q_1}{q_2} \right] = \frac{p_1}{p_2} \left[ \frac{p_2}{q_2} = \frac{q_1}{p_1} \right] > 0
\]

However, he must incur a search cost – say c and the possibility that
the search will be fruitless. Let \( \pi \) be the probability of successful
search and let c be measured in ‘utils’ (say of time). Then the
expected gain of cutting out the middlemen on the marginal exchange
is

\[
\pi u_i \frac{p_1}{p_2} \left[ \frac{q_1}{p_1} - \frac{p_2}{q_2} \right] - c - (1 - \pi) \left[ - u_i + u_i \frac{q_1}{p_2} \right]
\]

(1)

where \( u_i \) is the partial differential coefficient of the utility function
with respect to its ith argument. If the first agent initially consumed
both the goods, then the last term is zero. So there are three
considerations which will stop agent one from attempting barter: (a)
the probability of finding the ‘right’ partner, \( \pi \), may be small. That is
the usual double-coincidence of wants argument. (b) The gain to be
had in a monetary economy from cutting out the middleman may be
small because of the increasing returns of the middleman and the
large number of transactions since everyone else is making use of the
middleman in a monetary economy. That is \( p_i/q_i \) may be close to one.
(c) The costs in time and effort of search may be relatively high since
they are largely of the ‘set up’ variety.

If this sketch is on the right lines then one will see that in certain
commodities (e.g. with a thin market) exchange mediated by money
and middlemen may be at risk. Barter and partial barter deals are not
unknown in a monetary economy. However, in most cases the sign of
(1) will be negative and the monetary exchange mechanism will be
stable.

It is a consequence of this approach that a theoretical history of the
origin of paper money is rather harder than one might have thought.
A barter economy may easily be a Nash equilibrium as well. That is,
if all other agents barter, whether through middlemen or not, then
there are strong reasons why no individual agent should accept
promissory notes. The first person to accept a goldsmith’s IOU took
considerable risks. That is, a barter mechanism may also possess
considerable stability and certainly this is so for a species mechanism.
There seems to be the need for some co-operative move – perhaps
through a government – to move from one of these Nash arrange-
ments to another.

III HOLDING MONEY

All accounts which we have of why agents hold money appeal to the
axiom that exchange must be mediated through money although they
probably only need to postulate that exchange not so mediated is
very costly. Even so, this needs to be supplemented by the postulate
that selling and buying either are not, or may not, be simultaneous
and that there is a utility loss from not being able to buy at a
particular moment or from not being able to meet obligations in
terms of money incurred in the past. Not being able to obtain money
from sale at a particular moment is similar in its effects to having to
pay (a brokerage fee) to insure a sale. This fee, for instance, is a
central feature in the Baumol (1954)–Tobin (1958) account of the
transaction demand for money.

The transaction demand for money seems to be of central impor-
tance since sole reliance on other motives seems to lead to troubles.
These other motives turn on the uncertainty of returns on assets in
general which, it is argued, will lead to the inclusion of money in the
portfolio of a risk averse agent. The cause of the trouble I have
discussed elsewhere (Hahn, 1981) and I will give only a bare outline
here.

Suppose that there is a government perpetuity denominated in
money. Its return is made up of interest and changes in its price. Let
$p(s)$ be the price of the perpetuity in rational expectation equilib-
rium, if $s$ is the ‘state of the world’. For simplicity’s sake suppose
there is only a finite number of possible states $s_1, \ldots, s_n$ and that
by suitable numbering one has $p(s_1) > p(s_2) > \ldots > p(s_n)$. Then, if
at the portfolio investment date, the state is $s_n$ the investor is certain
that he will make no capital loss by holding the perpetuity. Hence on
these motives he will hold no money.

It is of incidental interest to ask whether this argument also applies
in reverse: if at the decision date the state is $s_i$ then the investor
knows for sure that he will not make a capital gain, but that he may
make a capital loss. This might be thought to be the case where
‘liquidity preference becomes absolute’. But the two situations are
not symmetrical since the interest payments can, given beliefs over
states, compensate for this downward risk and, in an equilibrium, will do so.

I conclude that uncertainty of returns is not on its own sufficient to account for the demand for money. (Of course, if there were no government bonds or other assets denominated in money, the above difficulty need not arise since price level changes between two date-cost pairs can give money a positive expected real return. But there are government bonds, etc. in the world.) The transaction motive and the attendant institutions of exchange of a monetary economy must be the foundation on which all else rests.

The question, as always, is the appropriate formalisation of an intuition. Until that has been done one cannot even be certain that the intuition constitutes a grammatical insight – for instance, the credit card company is a device to economise on the cash balances of at least one side of the transaction by guaranteeing the purchaser's delivery. It would economise on both sides if credit cards were universally, or at least widely, used. In other words, there is an opportunity for profitable intermediation by means of a system which guarantees settlement. But it needs to be done on a wide scale to be attractive and viable and that is perhaps why it has been slow to develop.

The arguments of this and the previous section lead to a rather natural way of ordering assets by liquidity. Suppose, first, that there are no transaction costs so that any asset can be exchanged against anything else costlessly. Think of goods distinguished by date and state of nature as well as by the other usual characteristics. The agent is uncertain about the state of nature, but knows prices given the state of nature and date. Then for a given amount of asset $A$ we can define $E(A)$ the set of goods which can be acquired given that one has that amount of asset $A$ (and nothing else). Next consider an amount of another asset $B$ which can be acquired at the same cost as the amount of $A$ did. Calculate $E(B)$. Assume that $E(A) = E(B)$. Now introduce transaction costs and for definitions think as them as either brokerage fees or the use of leisure, or both. Calculate $E^*(A)$ and $E^*(B)$ taking these costs into account. One can then say that $A$ is more liquid than $B$ if $E^*(A) \supset E^*(B)$ but not vice versa when $A$ and $B$ satisfy the condition that they cost the same and that $E(A) = E(B)$. These conditions should be looked at as a 'Gedanken Experiment': suppose I spend equal amounts on $A$ and $B$ and neglecting transaction costs they give the same transaction opportunities, which is the more liquid?
Now it does not follow that the more liquid asset will command a liquidity premium. For an agent may not be interested in the whole of $E^*(A)$. But if $A$ is money, I think we can take it that $E^*(A)$ and $E^*(B)$ share only one efficient point – that of no transactions. Then it is safe to argue that money will indeed command such a premium and thus, as Keynes argued, set a lower bound on the returns of other assets if they are to be held. Of course, agents will in general hold mixed portfolios. We can then calculate our sets for two portfolios costing the same amount, but differing in the proportion in which money is held. As the proportion of money held declines the fact that the portfolio becomes less liquid, on my definition, will become increasingly relevant to choice, since transactions in other assets will become increasingly frequent on a utility maximising consumption path. Hence lower liquidity will indeed result in a liquidity premium (in due course) and eventually in a rising premium. All of this is easy to formalise in a simple way.

It is for these sorts of reasons that I always considered the claim of an interest-inelastic demand for money as shown by data to be a ‘fact’ unsupported by theory and so to be rejected. For, on my way of analysing the matter, it is clear that there are many assets for which $E^*(B)$ is ‘close’ to $E^*$ (money). While these other assets do not drive out money for reasons already given, they are good substitutes for money as far as liquidity is concerned. Radcliffe was surely nearer the mark then Friedman! These considerations are highly relevant to the proposal to control the quantity of money. Perhaps it can be done, but it may be that what is needed is to control the liquidity of the economy and that is something else again. I suppose all of this is reflected in the endless proliferation of money measures which monetarists tell us to concentrate on.

Before I leave these particular matters there are two further theoretical hurdles which I want to consider.

If one is interested in short period analysis – that is, if one has not tied down expectations to be not systematically disappointed with the passage of time – there is a difficulty noted by Grandmont (1982) and others. If, for instance, people hold more money than they now want, this can be rectified for the economy by a higher price level provided the expected price level can be made not to rise in the same proportion. For, if people want to get rid of money because the price level is expected to rise, then they will still want to do so when the current price level is higher, but expected prices are higher in the same proportion. This difficulty of attaining a short period monetary equilib-
rium is acute when the transaction rôle of money is ignored. One then has the possibility of a Wicksellian situation where demand is always out of joint with supply. Thus in this case a certain conservatism in expectations is postulated (‘uniform tightness’ – of expectations) in order to make a short period monetary equilibrium possible. My own view is that consideration of transactions obviates the need for the postulate.

The second difficulty arises from the need to suppose that the probability agents attach to money having no exchange value at some future date is bounded away from one. For this cannot be true of a finitely lived economy with rational expectations unless one introduces the fiction that agents who are short at the last gasp of the world have to pay the government a certain amount of money. (How that is to be enforced is obscure.) I do not propose to discuss this difficulty which is entirely of the theorist’s own making. I am quite happy to live with a model in which expectations held today are rational for $10^4$ years, but not $10^{20}$ years.

IV MONEY AND THE REAL ECONOMY

I have already argued at length that a monetary economy requires a quite different description than does a non-monetary one. I do not suppose that this is controversial, but the formal literature for the most part takes no account of this. In any case ‘money matters’ profoundly to the real economy and the question is whether the quantity of money also matters.

One strand of this question has to do with something called ‘the optimum quantity of money’ although it is the real stock that is here at issue. A good deal of what has been written on this is of little worth because no attention was paid to all those deviations from an Arrow-Debreu world which must be present if there is to be money at all. Thus it happened that the Welfare Theorems which apply to an Arrow-Debreu economy were cavalierly applied to one where they cannot be shown to hold. A notable exception to these strictures is the work of Truman Bewley (1980). He starts with what predecessors had blissfully ignored, namely the absence of the requisite number of contingent futures markets. One of the motives for holding money is then self-insurance. Although Bewley does not look at transaction costs, he at least has a well-defined objective for people to hold money. Unfortunately, it turns out that for infinitely long lived agents
the ‘optimum’ quantity of money is not finite. It is clear that no simple argument about a positive marginal benefit of money balances and a zero marginal cost in their production stands much chance. For instance, it turns out that the taxation arguments for financing interest payments of money are relevant to the welfare argument.

It is, of course, true that when no return accrues to money holdings more resources will be used in transacting than would be the case if money had a positive return; that is, that will be so ceteris paribus. But for the optimum quantity the marginal liquidity premium of money should be zero – agents are liquidity satiated. Even if that can be done with a finite real stock there are other problems. Differently placed agents will be liquidity satiated at different rates of return on money and one really requires individualised returns in the manner of Lindahl prices for public goods. Even without this complication the payments would have to be conditional on the states of nature. One would also have to investigate carefully the ‘second best’ aspects due to missing contingent markets.

There is only one point that I wish to make on this topic. The power of money is its purchasing power, i.e. the ability to buy. If money bears interest it is cheaper to insure the possibility of purchase and more insurance will be taken out. But that means, for welfare, that the insurance must be real, which in turn means that in states where for transaction cost reasons, when money had no return an agent abstained from purchasing as much, he will now purchase more. But that means that there must be more to purchase. This in turn may imply that inventory levels will on average have to be larger when agents are liquidity satiated than when they are not. But that is a real cost! What I am proposing here is pretty innocuous: if liquidity satiation goes with more insurance then society cannot provide it costlessly. But it is easy to show that transaction costs will reduce the variance of purchases and so the need for inventories. I believe that something on these lines has been argued before.

I now turn to more urgent matters, although it must be recalled that in the absence of non-indexed taxes and rational expectations the main cost of, for example, inflation is said to be the extra transactions induced by the economising of money balances. To remove this cost Mrs Thatcher has engineered over 3 million unemployed.

The last remarks are motivated by the doctrines of neutrality and superneutrality. These doctrines are so well known that it does not
seem necessary to rehearse their essentials – but it is worthwhile to
get them right.

The basic proposition is simply a consequence of rational actions
by an agent whose preferences are over commodities. Such an agent
must be indifferent over two identical opportunity sets in commodity
space. Moreover, his preferred set of actions depends only on this
opportunity set. Abstracting from taxes it then follows that if it is
possible for an agent to have a \( k \)-fold change in the value of his
financial assets in every date-event pair in which all prices are \( k \)-fold
what they were before, then indeed his opportunity set is invariant to
these nominal changes. Notice that his financial assets include debt
denominated in money.

This is a result of extraordinary simplicity and common sense.
What is truly surprising is that it should have led to the claim that the
real state of the economy is invariant to the quantity of money.
Certainly for this claim to make even minimal sense either the
quantity of money must have been correctly forecast as well as prices,
or the statement must refer to ‘the long run’. But in particular, and
most importantly, it must be a Walrasian equilibrium allocation
which is said to be invariant to the quantity of money. However, all
one can possibly hope to show is that the set of such equilibria is
invariant. These are basically Arrow–Debreu equilibria or economies
in which money is what I have called ‘inessential’ (Hahn, 1973) so
that the equilibria of a monetary sequence economy are just the
Arrow–Debreu equilibria. But lack of uniqueness is endemic to these
equilibria and therefore the proposition is a non sequitur even if we
suppose the economy to be always in equilibrium.

I have discussed this aspect of the neutrality issue at length else-
where (Hahn, 1981) and shall not do so again. Instead I shall first
consider a simple example and then state the correct neutrality
proposition for it and proceed to draw a lesson from it for the widely
loved ‘ineffectiveness’ proposition. The example will be of a ‘macro-
kind’, that is, it will use some notion of the representative firm. In
general this is simply bad economics, but what is sauce for the
gander . . .

There are many reasons why a firm should value the employees it
has differently from those who are not yet employed by it. For one
thing, more is known about the presently employed; for another,
they are trained and familiar with the firm. Lastly, the firm may have
a kind of contract with its existing employees. This too I leave to one
side since it is examined at length in another study (Hahn, 1983). In any case, it is perfectly reasonable to suppose that there is a hiring cost which is a sunk cost. Call it $H$. If a new hired worker were to stay with the firm forever then $rH$ would be the flow cost which must be added to the worker's wage ($r$ is the constant interest rate). In general, however, this cost will depend on the length of time the firm expects to employ the worker. Suppose that is given and write $h$ for the flow cost measured in good(s).

Next I turn to the production function which is written as $f(k, \ell)$ where $k$ is 'capital' and $\ell$ is labour. (Of course, everything I have to say could be said in a full input–output way and nothing turns on $k$.) Since we are supposed to be concerned with the long run, I make the usual assumption of constant returns to scale. Let $m(k, \ell)$ be the marginal product of labour. Then if $w$ is the real wage the firm will be in equilibrium with $\ell^*$ workers and $k^*$ if

$$w + h \geq m(k^*, \ell^*) \geq w$$

(2)

Because of constant returns to scale this gives us a continuum of pairs $(k, \ell)$ which satisfy (2) – the latter can only help to fix $k^*$.\text{

Next, to stick closely to the new macroeconomics, let it be a condition of equilibrium that the real wage equals the marginal disutility of work:

$$u_r = w$$

(3)

where $u_r$ is the marginal disutility of work, here assumed to be the same for everyone. Hence if there are unemployed workers in equilibrium they do not care. But the unemployed are demanding leisure not goods. Hence the demand for the latter must depend on employment or more generally on $we^* + rk^* = y^*$. Notice that because of constant returns to scale there is no supply function of output.

The demand for goods will also, for the usual reasons, depend on real cash balances $m$. Hence the 'clearing of the goods market' is a third equilibrium condition:

$$y = f(\ell, k) = x(w, y, m, r)$$

(4)

It is important to emphasise that $y$ is an argument of the demand function as long as there are (voluntarily) unemployed workers. To this we now add the fourth obvious equilibrium condition:
\[ f_i(\ell, k) = \ell \]  

(5)

and as the fifth, the equality between the demand for money and its stock, \( (\bar{M}) \):

\[ pm(w, y, m, r) = \bar{M} \]  

(6)

where \( p \) is the price level. The equations (2) to (6) together with the production function constitute a miniature general equilibrium system. With the usual assumptions, it can be meaningfully solved for \( (w^*, r^*, \ell^*, k^*, y^*, p^*) \).

Suppose that the economy has an unemployment equilibrium. (I re-emphasise that this unemployment is voluntary by (3).) If \( y^* \) is the output in this equilibrium consider \( y^{**} > y^* \) where the available labour supply suffices to produce \( y^{**} \) when the technique of production corresponds to \( [\ell^*/r] \). Now give (6) the Cambridge form

\[ m = \lambda y = \frac{\bar{M}}{p} \]  

(7)

This is not needed for my argument, but the devil must have his due. Then if \( y^{**} \) is to be part of a new equilibrium it must be that \( m^{**} > m^*_1 \). But (7) enables us to write (4) as

\[ y = yx(w, r, 1, \lambda) \]  

(8)

so that that equilibrium condition becomes

\[ x(w, r, 1, \lambda) = 1 \]  

(4*)

which is independent of \( y \). It is then clear that by a proper choice of \( p^{**} \) we have a new equilibrium \( (y^{**}, p^{**}, r^*, w^*, \ell^{**}, k^{**}) \). In fact the model has many equilibria.

I shall comment on the robustness of this conclusion below. But first let me emphasise the message. Suppose we start in an unemployment equilibrium and somehow the quantity of money in the hands of every agent is increased in the same proportion. Suppose further that we have complete trust that the economy, after the disturbance, will settle down again in an equilibrium. What can be predicted?

(1) Certainly if the price level increases in the same proportion as the money stock while real wages remain constant, the old equilibrium will be restored.
(2) On the other hand, $y$ can increase at constant prices to restore the equilibrium in (7).

To say what will be the case we need dynamics and that is why the usual neutrality formulation is a non sequitur.

It is clear that a good deal will depend on expectations. For instance, we might model the dynamics as a sequence of short period equilibria under rational expectations. But if these paths are not unique, one agent’s particular expectations are only rational because others expect the equilibrium which he himself expects. Thus, to come to the ineffectiveness proposition, if agents accurately predicted the increase in $\bar{M}$ they would not be much wiser unless they knew whether it would induce a purely nominal or a real adjustment. Keynesians have been at fault in neglecting (1), but monetarists never consider (2). To give an inaccurate but telling summary: the rational expectation of Mrs Thatcher’s policy was that it would reduce output; it could have been rational to expect all nominal prices to be reduced by the implied amount. It could have been, but it was not the equilibrium we were in.

Now I must emphasise that although the little model which I have presented is on impeccable neo-classical lines, it is also special and not at all satisfactory. This is true of almost all of macroeconomics, but that is no excuse. So I shall now comment briefly on some of the special features.

(i) Constant returns to scale is what is usually postulated for long-run analysis. The realistic alternative is not diminishing, but increasing, returns to scale. That then requires imperfect competition. But since now a higher $y$ can go with a higher marginal product of labour I can do without the special assumptions in (3). I shall also be able to Pareto-rank equilibria.

(ii) Inputs, in general, precede outputs and hence the demand for them will depend on expectations in the widest sense. In particular, firms producing under constant or increasing returns must anticipate demand. This will introduce a further source of multiple equilibria. Abstracting from uncertainty there will be many expectation fulfilling equilibria.

(iii) In the model which I have given, there is a continuum of equilibria. This, of course, is rather strong and not supposed to be generally so. But unique equilibria are almost as rare, except when the model consists of one firm producing under diminishing returns to labour.
(iv) My model is played by the textbook rules and they are not very
good rules. In addition, I want to say that the labour market is in
equilibrium when there is no incentive for anyone to change wages.
This emphatically does not mean that in an equilibrium the market
and the shadow wage must coincide. For instance, when there is wage
bargaining then a Nash-bargaining equilibrium will, under the training
assumptions which I have made, lead to a wage equal to \( \bar{w} + h \), if
the alternative to working is the unemployment wage \( \bar{w} \). But then
when workers are similar the wage exceeds the reservation wage, but
there is no way in which it can be lowered. Thus, one can have a
proper equilibrium with 'involuntary' unemployment. In that case
there will be many equilibria even with diminishing returns. There
are many other directions towards 'realism' which have similar impli-
cations.

Of course, all this deserves lengthy considerations, but this is not
the occasion.\(^1\) My claim, in short, is (a) that multiple equilibria make
neutrality claims non-predictive even if the economy is always in
equilibrium, (b) that there are equilibria which are non-Walrasian
but properly founded on the theory of rational actions for which
multiple equilibria are endemic, and that for the leading case of
constant returns to scale this is also true of Walrasian equilibria.

Let me conclude by reiterating an observation which I have already
made elsewhere (Hahn, 1981) but which, it seems to me, cannot be
reiterated too often. Suppose the economy is Walrasian, but not,
miraculously, always in Walrasian equilibrium. Suppose that at \( t \)
the real market wage exceeds the shadow wage of labour. Neo-
macroeconomists assure us that this will start a process in which real
wages fall until they clear the labour market. But it will be a process –
a messy and costly one if experience is any guide. Suppose instead the
government injects helicopter money. Everyone believes that money
prices of goods will rise. But money wages by assumption will not,
since real wages are going to fall anyway. What a quick and painless
way to equilibrium! More pertinently: what real effects! These will be
particularly good if the helicopter money was anticipated. Keynes
made this point: every time you need lower real wages it is cheaper
and surer to use monetary policy. I have no idea why the current
advisors of monetarist governments fail to support a point which is so
much in harmony with all their theorising, bar one: that the labour
market is always in Walrasian equilibrium. But that they can hardly
hold to be even a good 'as if' hypothesis.
V  FURTHER CRITICAL REMARKS

There is a close connection between the classical dichotomy and money neutrality, although it cost Patinkin much effort to stop the belief that these two were identical propositions. The classical dichotomy proposes to determine all real variables before money is brought into the picture. The neutrality proposition claims that all real variables including the real stock of cash are independent of the quantity of money; it, however, determines the values of all real variables at one go. The proper scientific statement of the neutrality proposition is as follows: the set of equilibrium allocations is invariant to the quantity of money where an allocation is in the space of goods and the real money stock.

Patinkin argued that the classical dichotomy was logically untenable. He was criticised (Archibald and Lipsey, 1958) for not considering stationary equilibria, in which no agent wished to accumulate or decumulate money. In such an equilibrium his argument via Walras’s law does not work. Moreover, it may be possible, it was argued, to describe all stationary equilibria without bringing money into the picture. This, however, seems only possible when all agents are liquidity satiated which, it will be recalled, is not equivalent to the demand for money being equal to its stock. The demonstration is easy: if it is possible to liquidity satiate agents then that must be a Pareto-improvement – a real improvement – over an allocation in which satiation does not occur. Since there are also unsatiated equilibria, there is no more to say.

The mistake of Archibald and Lipsey, and indeed of much writing in this area, is that transactions are not modelled and in particular the relation between money and resources used in transactions are not modelled. But if money is only a store of value and transaction technologies are not modelled, one can hardly expect a monetary theory. Patinkin, although largely right on the dichotomy, also neglected to provide a full picture of a transacting economy and that must excuse his critics.

This brings me back to the beginning. Monetary theory can only exist in the context of a sequence economy in which transactions take place at all dates. If transactions use no resources, then even in a sequence economy the demand for money can only be rationalised when money has a return equal to its opportunity cost. Something quite analogous is true when uncertainty is introduced. Hence in that
case one finishes with a monetary theory which is daily conclusively falsified.

Monetarists, by and large, do not claim the dichotomy but only neutrality. The quantity equation \( m = \lambda y \) encapsulates the equilibrium relation between real balances and output. Simple economics suggests that \( \lambda \) is itself dependent on other serious real variables in the economy. It would, for instance, be a surprising claim by a neo-classical theorist, that transaction frequencies are independent of their real cost. Friedman announced, however, that \( \lambda \) is a stable function and so recommended the equation as the best aid to policy. However, if the other functions are not stable then the logic of this argument is obscure for we shall not be able to determine the appropriate values of the arguments of \( \lambda \). Moreover, by usual accounting if the excess demand for money is a stable function, then so must be its complement in the Walras law. Since the demand functions for other assets are claimed to be notoriously unstable this is a hard requirement, since somehow this instability must be 'undone' by the excess demand function for goods. Lastly to return to the old point: there may be several equilibria with different values of \( \lambda \) when the latter is taken as a stable function.

But now we must note that the dichotomy does make an appearance in some monetarist arguments. Their preferred model is the stationary state equilibrium: \( f'(\ell) = w, f'' < 0, \mu_i = \mu \). From this, equilibrium deviations only occur through mistaken price level expectations. So, in long-run equilibrium \( y \) is determined independently of the monetary sector. This result is due to the neglect of durable inputs, that is to say, productive assets and the odd procedure of making the demand of leisure independent of real cash balances when the demand for nothing else is thus independent. That sort of 'separability' is highly implausible.

However, in general 'real cash balance effects' are now de rigueur and indeed much weight is put on them. In this, as I have already emphasised, the real money balance has become a 'real' variable to be determined by the 'real' economy. But how, exactly and precisely, is the latter specified? We know that tastes, endowment and production technology do not suffice for the specification of a monetary economy. At the very least we must endow agents with information and expectations. Even so, I have argued, money may slip through our fingers unless its rôle in transactions is made essential. How far the given transaction institutions and technologies are part of the
basic description of the economy and how far they are the consequence of genuinely basic features like information and geographical dispersion is open to argument.

Envoi

Monetary theory cannot be more convincing than the general theory of which it is a part. The latter, however, at the moment hardly carries conviction. The real fluctuations which we observe seem far too large for consistency with the world being in continuous Walrasian Rational Expectations Equilibrium. Moreover, the theory gives no account of how prices change and who changes them. There is a lot of work to be done and those of us not intent on manipulating the world should keep pretty quiet. As far as monetary theory proper is concerned it does not seem to have risen above the banality of the homogeneity proposition.

NOTE

1. At the time of writing this paper I was working on a study of general equilibrium with implicit contracts, the preliminary results of which confirmed the claim in the text. This is since published in Hahn (1984).

REFERENCES


