Money, Credit, and Business Fluctuations*

JOSEPH E. STIGLITZ
Stanford University

This paper provides a critique of standard theories of money, in particular those based on money as a medium of exchange. Money is important because of the relationship between money and credit. The process of judging credit worthiness, in which banks play a central role, involves the collection and processing of information. Like many other economic activities involving information, these processes are not well described by means of standard production functions. Changes in economic circumstances can have marked effects on the relevance of previously accumulated information and accordingly on the supply of credit. Changes in the availability of credit may have marked effects on the level of economic activity, while changes in real interest rates seem to play a relatively minor role in economic fluctuations.

This alternative view has a number of implications for policy, both at the macroeconomic level (for instance, on the role of monetary policy for stabilization purposes and the choice of targets) and at the microeconomic level.

1 Introduction

Money has long played a central role in popular conceptions of economics—and life more generally. 'Money makes the world go around!' 'Money is the root of all evil.' These are but two aphorisms which come to mind.

Professional economists give money an equally mixed review. The monetarists, whose enormous

* Sir John Crawford Lecture, to the 1988 Australian Economics Congress. Financial support from the National Science Foundation, the Olin Foundation, and the Hoover Institution are gratefully acknowledged. This paper is part of a larger research project, undertaken jointly with Bruce Greenwald and Andrew Weiss, attempting to construct a macroeconomic model based on solid microeconomic foundations, in which considerations of imperfect information play a central role. This work, sometimes referred to as New Keynesian economics, while it shares many features with traditional Keynesian models, differs in several crucial respects. For a review of the central features of these models, and a contrast between the New Keynesian and New Classical approaches, see Greenwald and Stiglitz (1987b). For a survey of the empirical evidence in support of these theories, see Greenwald and Stiglitz (1988c).

popularity of the early 80s seem subsequently to have waned, place money as a central determinant of economic activity. By contrast, when I was a graduate student at MIT, I was taught the classical dichotomy, in which the money supply did nothing but determine the price level. Monetary economics, as I was taught it by Paul Samuelson, was a seemingly curious course, devoted to explaining why we should not be studying the subject. Money had no real effects. I was persuaded by the theory. Since, as a young idealist, I thought economists should be concerned with matters of importance, with real variables, I paid little attention in the following years to monetary economics.¹

As a result, I missed several of the fads in empirical monetary economics: what mattered was unanticipated money; what mattered was

¹ Monetary economics displayed the same schizophrenia between microtheory and macrotheory that characterized then (and continues to characterize) the profession. In macroeconomics, we were taught that, at least some of the time, money did matter. It had real effects. For reasons which I shall explain later, I found the macroeconomic analysis less than compelling.
anticipated money; money did not matter, once one had correctly controlled for short-term interest rates; velocity (correctly calculated, with the right definition of money, a definition which was itself constantly changing) was or was not constant.

The Classical Dichotomy Strengthened

The theory, too, progressed, if in an uneven way. Some economists who, in other respects, insisted that models should not be ad hoc, that they should be based on principles of maximization, took the rough road around the difficulties posed by money by putting money into the utility function or the production function—a trick, which repeated often enough, took on a semblance of respectability! Others took the high road, creating a demand for money by assuming that it was required for transactions, modelling it as an old-fashioned cash in advance constraint—criticisms that it was an ad hoc assumption which was blatantly false being brushed aside with the remark that these were topics for future research.²

On the other side, I managed to strengthen considerably the theoretical foundations of the classical dichotomy, incorporating uncertainty, and showing that the conclusions did not depend either on the existence of a complete set of markets or even rational expectations. Using the approach I had previously employed to demonstrate the irrelevance of corporate financial policy (Stiglitz, 1969, 1974), I showed that public financial policy had no effect (Stiglitz, 1983, 1988).³ Establishing a form of Say’s law for government debt, I showed that if the government reduced taxes and increased its debt, the demand for government bonds increased by an amount exactly equal to the increase in supply. Not only did such a change in financial policy have no real effects, it had not effects on prices either. Furthermore, a change in the term structure of government debt had no effects. This result seemed particularly surprising, given the widespread acceptance of the Tobin portfolio approach to monetary economics in which changes in the supplies of bonds affect the level of investment as a consequence of the general equilibrium effects on the price of capital (Tobin’s ‘q’). The ‘mistake’⁴ was that Tobin (and others of that school) forgot to include, as one element in individuals’ portfolios, the variable tax liabilities. The probability distribution of these tax liabilities changes, of course, when the government undertakes a change in, say, the maturity structure of its debt. My analysis took those changes explicitly into account.

Of course, like any theorem, there were assumptions which went into the analysis. These seemed to be of two sorts: some, like the absence of distortory tax affects, while they would alter the qualitative result that taxes had no effect, seemed an implausible basis for an argument about why monetary policy should be important: surely its effectiveness did not hinge on the real effects arising from the difference in the change in the dead weight losses from an increase in taxes in one year compensated by a decrease in taxes in some later years! Another assumption in the analysis was the absence of intergenerational redistributive effects. While one might agree or disagree with Barro that the economy is best modelled as a set of dynastic families, with no inter-generational effects, surely short-run monetary policy does not hinge on these inter-generational effects.

The other set of assumptions, concerning perfect capital markets (though the analysis did not require that there be a complete set of risk and futures markets), was no different from that assumed in conventional economic models. If that assumption was struck down, with it would fall much of the standard theory. Of course, practical men have long

² Some recent research has attempted to address this criticism, but in ways which I find unconvincing. Models are constructed where, for instance, individuals only meet once, and hence it is impossible to establish creditworthiness. But even in these circumstances, there would be incentives to establish credit intermediaries. The fact is that credit facilitates a substantial fraction of all transactions, a fraction which itself is presumably endogenously determined Constructing a model in which that fraction is, of necessity, zero may provide limited insights into our economy

³ Several other versions of the irrelevance theorem, parallelling the various versions of the Modigliani-Miller theorem, were established, e.g. by Wallace. These required somewhat more stringent conditions, for instance, the existence of a complete set of securities markets.

Earlier, simpler versions of public irrelevance theorems, in which there was no uncertainty, were popularized by Barro (1974), though similar results were proven several years earlier in an unpublished paper by Robert Hall (1967), and Buchanan attributes the idea to Ricardo (hence the popular characterization of the result as the Barro-Ricardo proposition.)

⁴ I hesitate to call it a mistake, because I suspect, as a behavioural model, his probably provided a more accurate depiction of the economy than mine, which required, if not rational expectations, at least a certain form of ‘consistent expectations’, entailing a higher degree of rationality than, I suspect, is commonly found.
claimed that the economists' models of the capital market were unrealistic, and a host of institutional economists (and theoretical economists, when they found it to their convenience) have made use of the assumption of imperfect capital markets. But higher minded economists have looked derisively at those who made reference to imperfect capital markets, accusing them of, among other sins, ad hocery.

One of the most important developments in economic theory, however, of the past 15 years has been to explore the consequences of imperfect and costly information for the functioning of the capital market. These studies have shown that those models which assumed imperfect capital markets may have been much closer to the mark than those which, on the contrary, assumed perfect capital markets. These studies have shown that capital markets that are competitive, in the sense which that word is commonly used, may be characterized by credit and equity rationing. These models not only provide explanations of institutional details of the capital market, details which are either inconsistent with the perfect capital market models or about which they have nothing to say; but they also provide a basis of an explanation of macroeconomic (aggregative) behaviour which are also inconsistent with the conventional neoclassical model. It is not, however, my purpose here to provide a review of these results.

I have now come to believe that monetary institutions and policy do have important real effects, but for reasons quite different from those of the standard theory. My objective in this talk is to explain why it is that monetary policy is, sometimes, effective and why the conventional explanations of the mechanism by which it works, particularly those versions based on the transactions demand for money, are inadequate, though there may be some slight grains of truth in these conventional explanations.


The theoretical and empirical objections to the real balance effects are perhaps slightly better known and, I think, equally telling. The one approach which I find somewhat persuasive is the Tobin portfolio approach but, as I noted above, that approach faces certain difficulties. The theory we present below resolves some of those difficulties.

II A Critique of Transactions Based Traditional Monetary Theory

There are several reasons why one might be suspicious of the traditional explanations. The past 15 years have witnessed remarkable changes in transactions technologies. Computers enable the velocity of circulation to become virtually infinite, for instance, in the use of money market accounts. The relationship between conventionally measured money and income, while it has not been stable in recent years, has not moved in a way that a transactions-based theory would have predicted.

More fundamentally, most transactions are exchanges of assets, and are not related directly to the production of income. And there is no a priori reason to expect that the relationship between the two, the volume of exchanges of assets and the level of income, should be stable. Indeed, there is reason to believe that in periods of rapid change, such as when the economy is going into a recession, there will be large differences in opinions concerning future prospects of the economy, as well as large changes in relative net wealth positions, leading to corresponding larger than normal exchanges of assets. Of course, many, if not most, asset exchanges may not entail money (or may entail money with a much higher velocity than transactions involving labour services). But that is exactly my point: transactions do not require money, only credit.

Furthermore, with the recent growth of interest bearing accounts, the opportunity cost of holding money has, by and large, become the difference between the interest rates paid on money market accounts and those paid on other short-term financial instruments—a difference which is, for economic purposes, minuscule. It is this interest rate differential which should appear in the money demand equation (and in the LM curves of macroeconomic analysis). It stretches the credulity of even the most hardened Keynesian to believe that monetary policy operates through changes in the real value of this interest rate differential. Of course, in a general equilibrium system, any exogenous change will have affects throughout the system. Were the government to buy up peanuts and burn them, it would have real effects on output, investment, and employment. But though there is a long tradition in macroeconomics on focusing on third- and fourth-order effects, dating at least back to the real balance effect, there is a consensus that a peanut theory of macroeconomic policy will not do; so too should a theory based on changes in the differential between money market accounts.
and short-term government bills be looked upon with suspicion, though there may indeed be links between this differential and the appropriate long-term real interest rates, and between those interest rates and the level of investment.

But even that last link is suspect; or, to put it more accurately, it does not seem plausible that variations in long-term real interest rates have played an important role in fluctuations in investment. It is hard to know, of course, what the relevant real interest rates are. We have time series for nominal interest rates for loans of various maturities, but what is relevant for an investor (in the traditional theory), at the time he makes his decisions, is the real interest rate, and to calculate this, he must form a forecast of future prices. (This would not be true, of course, if all loans were appropriately indexed.) As a first pass at this problem, Dwight Jaffee and I (1988) looked at the realized real rates of interest. We constructed a time series for what the real rate of interest would have been on five-year indexed government bonds (so we can ignore variations in rates of default), assuming the market acted as if it had perfect foresight and was risk neutral. Such a hypothetical rate would have been the rate that would, in neoclassical models without equity and credit rationing, be relevant for investment decisions in five-year machines. Though there have been some secular changes in this real interest rate, rising in the 80s to record heights, there is virtually no relationship between this, the relevant real interest rate, and the level of investment. (Similar results hold for the ten-year real interest rate.) With a few exceptions—the Great Depression and the 1980s—one might as well treat the real interest rate as a constant. And constants do not provide a basis of a good theory of fluctuations.

This is not surprising, given businessmen’s accounts of how they make investment decisions. Though interest rates enter the calculations, they conventionally require (expected) real returns of 15, 20, or 25 per cent before they undertake a project; given the uncertainties associated with all aspects of long-term investment projects, a variation in the real interest rate from 3 to 4 per cent is absorbed in rounding error.

The skepticism that we have expressed here and elsewhere about the role of interest rates has, of course, been implicitly or explicitly shared by monetarists. They have noted the regularity between the money supply and the level of output, with interest rates seemingly playing a relatively minor role. Leaving no theorists the task of explaining this sometime regularity, they have been content to base policy prescriptions on the persistence of this relationship, to argue that if only we increase the supply of money, output will go up. The failure of this prescription to work in the 80s should not detract us from the essential insight of the monetarists—the seeming unimportance over long periods of time of the interest rate.

There are further objections to the transactions based theory of money: when Ireland faced a strike which closed down the clearing mechanisms for checks, while the transactions based theories might have suggested that the economy would have come to a screeching halt, alternative arrangements were easily worked out, and the effects were indeed limited. Italy has periodically gone through periods of shortages of small currency, with little impediment in trade, the major effects perhaps being the increased number of cavities resulting from the slightly larger number of pieces of candy consumed, as candy became conventionally used for small change.

The Survival Power of a Bad Theory

I have thought hard about the question of how could a theory with such seemingly little prospect for explaining major macroeconomic events gain such widespread popularity. (If bad managers had an ability to survive comparable to that of bad economic theories, the capitalist economic system would indeed be faced with problems!) There are three possible explanations.

(i) Money as an Accounting Device in Simple Games and Primitive Economies

One of the lessons we have learned from Freudian psychology is the importance of going back to events of our childhood if we wish to understand our neuroses, our obsessions, and the way we misperceive the world. Most of us when we were children played simple games, in which play money was used. In playing monopoly, if we wanted to buy a house or hotel or a piece of property, we had to have money—we could not buy on credit. There was a cash-in-advance constraint; money was the medium of exchange.

Money served a second function: it was the method by which we kept score. The person with the most money at the end of the game was declared the winner.

Later on, most of us pass on to more complicated games, such as bridge, where there are alternative ways of keeping score.

Similarly, primitive economies need a simple
accounting scheme. Since the expulsion from Eden, all economies have needed some method of keeping track of what each individual contributes to and takes out of the economy. Primitive economies use money ($M_0$) as their accounting device.

But even in simple economies, there is a realization that there may not be an intertemporal coincidence of wants. There are gains from intertemporal trades. Intertemporal trades require credit: one individual gets more than he has—up to that point in time—'earned' the right to, in return for a promise to pay back some of his future earnings. Hence, even primitive economies drop the cash-in-advance constraint.

In modern economies, the task of determining who is credit worthy, who is likely to repay a loan, is both more important and more difficult than in primitive economies. It is not only that (to use Bohm Bawerk's term) roundabout means of production are more productive, but there is a separation between entrepreneurship and capital: those with the best ideas for using resources are not necessarily those who have control of resources. It is not surprising then that institutions have developed to ascertain who is credit worthy (and indeed different institutions specialize in evaluating different kinds of risk). Banks are among the most important of these institutions.

Banks can thus be viewed as social accountants—keeping the records of what individuals contribute to and take out of society—and screening devices. (See Stiglitz and Weiss, 1988a.)

Though the 'model' of money which is derived from childhood games and primitive economies is not a good model of our credit economy, there is one important lesson to be derived from it. Even in our more developed economies, money is part of our system of accounts, of ascertaining who has rights to resources. If the government should print money to finance deficits, this will interfere with our accounting system. The claims on resources, at existing prices, will exceed the supply of resources, and inflation will result.² On the other hand, if the government borrows to finance government deficits, there is no necessary increase in the supply of current claims on resources; individuals simply trade with the government their rights to current resources for rights to future resources. Of course, the increased supply of government debt may provide the basis of a general expansion of credit (a monetary expansion) by the banking system, but it need not do so. The fact that the banking system may extend more or less credit than is required to attain full employment at stable prices is a more general problem, which we discuss below.

(ii) Sunk Costs of General Equilibrium Theory

Let us now turn to the second possible explanation of the persistence of the transactions-based view of money. In spite of the sermon that we preach to our students about letting bygones be bygones, and ignoring past sunk costs in making future investments, there is considerable evidence that businessmen do not behave this way—nor do we academics. Having invested an enormous amount of intellectual capital in learning the general equilibrium model, we want to apply it wherever we can. A simple change of notation from $C$ for consumption or $x$ for output to $M$ for money and we have a theory of money. Monetary economics becomes the study of the demand and supply curves for money, just like agricultural economics is the study of the demand and supply for the output of farms. Of course, if we pursued the analogy further, we might have thought that there would be more than just the estimation of demand and supply functions: there would have been an interest in understanding the underlying technology, of examining the transactions technology and the relationship between that technology and the demand for money. Students of money and finance who are not macroeconomists do study in detail the financial institutions, but they study the instruments with which credit gets extended, not by which transactions get recorded.

Credit as information

It is remarkably difficult to incorporate credit within the standard general equilibrium model. Credit can be created with almost no input of conventional factors, and credit can equally easily be destroyed. There is no easy way to represent the supply function for credit.

The reason for this is simple: credit is based on information. Ascertaining that an individual is credit worthy requires resources; and standing by that judgement, providing or guaranteeing credit, entails risk-bearing. But there is no simple

² Similarly, when gold was at the center of the accounting system, an increase in the supply of gold (as a result of the discovery of new gold deposits) constituted a disruption to the accounting system. The supply of 'claims' on resources was increased, without a commensurate increase in the supply of resources. Inflation again resulted.
relationship between these economic costs and the amount of credit extended.

The physical capital with which we produce in our factories and fields may be slightly affected by outside disturbances—rain may lead to rust—but only major cataclysms, such as wars, can have a significant effect in the short run. But informational capital can be far more easily lost. Changes in relative prices may require a drastic re-evaluation of individuals' and firms' credit worthiness.

*Interest Rates are not like Conventional Prices and the Capital Market is not like an Auction Market*

The standard general equilibrium model is not only not helpful for understanding credit markets; it may be positively misleading. It is misleading because we are apt to think of the price of credit—the interest rate—being a price like any one other price, and that it adjusts to clear the market.

The interest rate is not like a conventional price. It is a *promise* to pay an amount in the future. Promises are often broken. If they were not, there would be no issue in determining credit worthiness. As Andy Weiss and I (1981) have shown, raising the rate of interest may not increase the expected return to a loan; at higher interest rates one obtains a lower quality set of applicants (the adverse selection effect) and each of one's applicants undertakes greater risks (the adverse incentive effect). These effects are sufficiently strong that the net return may be lowered as the bank increases the interest rate charged: it does not pay to charge high interest rates. Market equilibrium may be characterized by credit rationing.

*An Analogy* The fact that the credit market is fundamentally different from a conventional market for goods should be familiar to us from another context. None of the private universities in the US—even those, such as Harvard, Stanford, Yale, Princeton, and Northwestern, where first rate economists have served as President, provost, and deans—has employed the price system to allocate the scarce number of places for students. Let me remind you of how we often talk about the auction for credit working; those who have the best projects are willing to pay the highest interest rates, and thus the auction market ensures that the best projects—and only the best projects—get funded. Of course, we recognize the possibility of human error. But then we say, if the individual makes a mistake in over-estimating the return, he bears the cost.

Similar language could apply to an auction for places in our universities. Those with the highest return to a Stanford degree would bid the highest, thus ensuring that the value added of our scarce university resources is maximized. Of course, there is a possibility of someone over-estimating the value of a Stanford degree, but if the individual makes such a mistake, he bears the costs. Lack of capital should be no problem: the university would simply take an IOU.

Put in this way, we can immediately see the fallacies in this reasoning. Students who bid too much will default on their IOU—just as those who bid too high an interest rate on their loan applications will default. Not only cannot we rely on individuals' judgements, there may be reasons to believe that those who are willing to bid particularly high are more likely to default. The bank or the university—not the borrower or student—bears at least some of the costs of these misjudgements. The scarce resources would not be used in a way which maximized value added. The auction system would result in the universities being flooded with those over confident and cocky students who are so unpleasant to teach, combined with those natural charlatans and cheaters, who feel no more moral compunction about defaulting on their student loans than they would on cheating on an exam.

So too in credit markets: in the face of uncertain prices, wages, and interest rates, the return to a project depends as much on expectations of those future prices as it does on the physical outputs. Those who are most willing to bid high for a loan are those who are most optimistic about these future prices and are least risk averse, and/or for whom the cost of default is lowest. But there is no reason to believe that allocating credit to these individuals maximizes either the private return to the bank, or the social return to society.

Just as universities spend resources screening applicants, so too banks spend resources screening applicants. The screening is far from perfect, yet some screening is far preferable to none. Prices (or in this case, interest rates)—market auctions—do not provide, by themselves, an adequate screening mechanism.

*Credit and Equity Rationing*

The informational problems, about which I have just spoken, may easily give rise to credit rationing. Recall, again, the conventional stories: when there is an excess demand for credit, an unsatisfied borrower offers the bank a higher interest rate. As interest rates thus rise, the demand for credit decreases, and the supply increases, until equilibrium is attained. But now, consider what
happens if, at the interest rate which maximizes the bank's expected return, there is an excess demand for credit. The bank would refuse a customer who offers to pay a higher interest rate, reasoning that he is (on 'average') a bad risk. The expected return to such a loan would be lower than for loans the bank is currently making.

Banks will, of course, change the interest rate which they charge as economic circumstances change. But there is no simple relationship between the interest rate charged—or even the interest rate paid depositors—and the state of the economy. As the economy goes into a boom, the returns to all projects may increase, and one might be tempted to argue that as a consequence the real rate of interest ought to rise, presenting a quandary, since in some instances—such as the Great Depression—the real interest rate moved counter-cyclically rather than pro-cyclically. But Weiss and I (1988) have shown that there may be important instances, where the returns to say risky projects increase relative to safe projects in some booms, in which banks' optimal interest rate (and accordingly, the market equilibrium interest rate charged borrowers) actually falls. And it is even possible that it falls by enough that the real return to depositors falls, even though the expected default probability has fallen. Though total expected rates of return have increased, that fraction of total returns which lenders, with conventional debt contracts, can appropriate decreases so much that their expected return is lowered.

Not only may these informational problems give rise to credit rationing, they may also give rise to what Bruce Greenwald and I have called equity rationing (Greenwald and Stiglitz, 1987a,b; 1988a,b). Because of adverse selection and moral hazard effects, the cost of issuing equity is sufficiently great that most firms act as if they were equity rationed. When they are denied bank credit, they do not raise capital by issuing new equity, but rather constrain their capital expenditures to retained earnings. Equity rationing is particularly important, because it means that firms cannot divest themselves of the risks which they face; the original owners cannot fully share the risks throughout the economy, and accordingly the firms does not act in a risk-neutral manner. (Of course, there are other reasons to believe that firms will not act in a risk-neutral manner. In large corporations, decisions are made by managers, whose compensation almost always is partially contingent on the performance of the firm. Such contingency pay is viewed to be necessary to provide the requisite managerial incentives.)

Credit and equity rationing, or more broadly, the informational problems associated with the capital market, provide insights into two of the puzzles we have noted above. If, as we have asserted, credit rationing is, at least at times, important, it would explain both why corporate financial policy is not irrelevant and why public financial policy is not irrelevant. It would also explain our findings concerning the seeming unimportance of real interest rates. When the economy is credit rationed, it is the quantity of loans, not the interest rates which are charged, which is critical. But even when the economy is not credit rationed, equity rationed firms may not be willing to borrow, given that they cannot divest themselves of the risk associated with production and investment (in the absence of perfect futures markets). Thus, in a recession, increasing the supply of funds available may have little effect on investment—providing a theoretical justification for the traditional characterization of monetary policy at such times as pushing against a string.

(iii) A Seeming Empirical Regularity

The third, and to my mind, most credible reason for the survival of the conventional monetary model is the seeming regularity of the relationship between money and income, a regularity which called out for a theoretical explanation. The transactions based model seemed to provide an explanation for this seeming regularity, though one which, we have argued, is both incomplete—it does not explain why money is required for transactions, it simply assumes it—and wrong, since money is in fact not required. It is not a serious advance in economic theory simply to assume—counter-

---

8 That is, the original owners-managers of firms are more informed about their firm's expected return than potential purchasers of shares. At any given price of equity, those with the lowest expected returns are most willing to sell their shares. See Stiglitz (1982), Greenwald, Stiglitz, and Weiss (1984), and Meyers and Majluf (1984).

9 Because the managers only receive a fraction of the returns to their entrepreneurial efforts, their incentives are attenuated. See Stiglitz (1974) or Ross (1973).

10 For evidence on this, see Asquith and Mullins (1983).

11 There is another possible explanation for this behaviour: discontinuities in the marginal cost of raising funds resulting from the differential treatment of different forms of finance by tax authorities. See Stiglitz (1973).
factually—that cash is required in advance, rather than to assume a money demand function takes on a particular form. The former seems no less ad hoc than the latter, and the latter—for an appropriately specified function—at least has the possibility of being correct.

The disadvantages of using reduced form relations are, by now, well known: one cannot use such models to analyze 'regime changes'. But can one have much reliance on the predictions of a theory based on the obviously unsatisfactory assumption of the cash-in-advance constraint?\textsuperscript{12}

Summary
Let me pause for a moment to review where we are in the argument. We have contended that (a) the transactions based theory of money is implausible: that even if there were a stable relationship between 'transactions' and 'income', such a theory would at best explain the differential between the return on a money market account and on a treasury bill; it would not explain the real rate of interest; (b) that even if it were shown that monetary policy affected the real rate of interest, there is little reason to believe that variations in real rates of interest have played an important role in investment variability; (c) informational imperfections can explain rigidities in real interest rates, equity rationing, and credit rationing; they provide a theoretical explanation for what, in the older institutional literature, was casually referred to as 'imperfect capital markets'; (d) in particular, it is inappropriate to view the credit market with the same tools and perspectives as employed to study markets for conventional goods and services; it should not be viewed as an auction market; (e) rather, it is a market in which allocations are made on bases other than 'price'; and (f) banks serve a critical role as social accountants and screening devices for the allocation of credit.

I have, perhaps, been more forceful in pointing out the defects in the competing theories than I have in providing a complete explanation of the mechanisms by which monetary policy and credit mechanisms work within our framework. Let me turn to this now.

III A Closer Look at the Allocation of Credit

Though we have provided a theoretical rationale for the widely observed phenomena of credit and equity rationing, many of the consequences (which we are about to describe) hold regardless of the explanation of credit and equity rationing.

Consequences of credit rationing: role of interest rates
We have already noted one consequence: that investment may depend less on the interest rate charged than on the availability of credit. Changes in availability of credit may be much more marked than changes in interest rates charged. Indeed, there may be no simple relationship between credit conditions (e.g. the riskiness of the environment), credit availability, and the interest rate charged. (Stiglitz and Weiss, 1987b.)

Consequences of credit rationing: need for central bank
We have also noted a second consequence: that because interest rates are not used to equilibrate the demand and supply for funds, the decentralized system of credit allocation of capitalist economies\textsuperscript{13} may result in discrepancy between the two at full employment levels of output. One can view central banks as a substitute for prices (interest rates) as an equilibrating mechanism.

Mechanisms by which central banks affect credit availability
In the theory we have put forward, the central bank affects the availability of credit both through effects on the willingness and ability of banks to make loans. When banks have no free reserves, then monetary policy operates through its effect on banks' ability to make loans; it changes the constraint. On the other hand, when banks have free reserves (as they did in the Great Depression), monetary policy operates through banks' willingness to make loans. One important aspect of this which to date has received insufficient attention is described below.

Consequences of equity rationing: risk-averse behaviour
A major consequence of equity rationing is that firms may act in a risk-averse manner. Equity rationing limits the extent to which the firms' risks can be shared. The fact that future markets are incomplete and there are lags in production means that every production, employment, and investment

\textsuperscript{12} I really do not want to single out a particular model. Similar objections can be raised against models in which, for instance, monetary policy exercises its effects through real balance effects.

\textsuperscript{13} We should emphasize that there are good reasons for this decentralization, associated with the wide dispersal of information. See below.
decision is a risk decision. Accordingly, changes in either the risks faced by firms, or in firms’ ability to absorb risks, will affect all of these decisions of the firm. Thus, an increase in the degree of uncertainty about future economic prospects may lead firms to reduce their level of production; and a worsening in a firm’s balance sheet (its net worth) will lead it to undertake less investment and production.

**Interactions between equity and credit rationing and their consequences**

There are some important interactions between credit and equity rationing. First, firms’ most serious fears are associated with bankruptcy. Bankruptcy is determined not just by cash flows. Firms go bankrupt when they make losses and cannot find lenders willing to lend to them. The fear of future credit rationing is one reason that firms act conservatively today.

Second, most firms are not only engaged in production, but also in lending. They are mini-banks. The reason that firms so frequently reject Polonius’ advice about ‘neither a lender nor a borrower be’—and become both lenders and borrowers—is simple: information in our economy is diffuse. Suppliers have information about their customers that others do not have, and customers have information about their suppliers that others do not have. This lending activity is, however risky. Thus, when a firm’s net equity position is worsened or when a firm’s sources of funds are restricted (it is credit rationed) or when a firm’s perception of the risk of lending is increasing, firms will reduce their lending activities.

Third, banks are like firms: their production activity is the screening of loan applicants. And just as firms are risk averse, so too are banks. A reduction in their net worth (equity) or an increase in their perception of the risk associated with making loans reduces their willingness to make loans.

**Multiplier effects**

In our previous discussion we have shown the market economies will be characterized by price (interest rate) rigidities and credit interlinkages (both directly among firms, and between firms and banks). This results in the economic system exhibiting at times instabilities; small disturbances are amplified through a multiplier process which is quite distinct from the traditional Keynesian multiplier process (or more accurately, contains that process as only one element).

(The distinction of how the economy is envisaged to respond to disturbances is another example of the marked disparity between macro- and microeconomics. Traditional micro-theory, while it recognizes the difficulties of proving stability, begins with a downward-sloping demand curve and an upward-sloping supply curve, in which a disturbance to either is dampened through adjustments of prices. The market economy does not amplify disturbances, but dampens them.)

Perhaps the best way to contrast our perspective with the traditional one is to trace out the consequences of some disturbance—an increase in the uncertainty associated with the demand for some good that a country exports—both in our theory and in the traditional theory. In the models of the true believers, a mean preserving increase in uncertainty, that is one which does not change the expected return, would have no effect on current investment, since all firms are fully diversified. (A reduction in the expected return would, of course, lead to a decrease in the demand for investment, leading to a decline in real interest rates; the subsequent reduction in savings would quickly restore equilibrium.)

In our theory, risk-averse firms (recall, our theory explains why firms should act in a risk-averse manner) would be less willing to produce and invest. This reduction in investment then has further repercussions. Prices for the investment goods the firm would otherwise have brought are lower (than they would otherwise have been, and lower than the firms had expected). The consequent reduction in profits erodes the equity base of these firms,

---

14 There are other possible explanations for why firms might behave in a risk-averse manner; the principal agent literature has stressed the importance of providing managerial incentives; any incentive structure must have managerial rewards depending significantly on firm performance; this may lead to risk-averse behaviour on the part of the managers.

15 There are other reasons for the interlinking of production and credit activities, having to do with moral hazard (the chance of a defective product). The credit provided by a supplier can be thought of, in part, as a guarantee. More generally, certain externalities associated with moral hazard frequently give rise to interlinking of markets. See Braverman and Stiglitz (1982), and Arnott and Stiglitz (1988).

16 In related work, we have shown how the labour market will be characterized by wage rigidities (the efficiency wage hypothesis) and the product market may be characterized by price rigidities. See, e.g. Stiglitz (1984, 1985a, 1986, 1987a,b) and Shapiro and Stiglitz (1984).
who, in turn, become less willing to invest. Furthermore, a fraction of these firms (if the original disturbance was large enough) will go bankrupt. The increased bankruptcy leads to perceptions of increased uncertainty associated with making loans (both on the part of banks and of firms making loans to suppliers and customers).

We now have third-round effects. These are of two forms. The reduced investment of the suppliers of firms has further effects, along the same lines as the original disturbance. In addition, the (higher) bankruptcy rate means that banks’ net equity position is deteriorated.

One can view banks like conventional firms. The business of banks is making loans; they ‘invest’ in cultivating clients and screening applications. When their net equity position is reduced, they become less willing to make loans.

We now turn to the central role of the banks. For simplicity, assume that borrowers can be classified into three categories, ‘good’, ‘OK’ and ‘bad’ risks. The bank knows that it places firms into these categories imperfectly, with some good risks being classified as bad, and some bad as good, etc. All good loans are, nonetheless, granted, no bad loans, and—depending on the availability of credit—only some OK loans. Again, for simplicity, we assume all borrowers apply for the same size loan. If truly good borrowers always applied to larger (or smaller) loans, that would be a signal, and the interest rate would be lower to reflect the higher quality. But that could not be an equilibrium.17

How does the disturbance we have just described affect the bank in its central function as a screening device. First, the greater uncertainty is likely to lead to greater errors in classification. This is likely to lower the expected return of loans to those classified as ‘good’ and perhaps those classified as OK. Accordingly, interest rates charged would have to rise for the bank to obtain the same expected return. Second, the greater uncertainty of returns lowers the expected return to the bank, at any interest rate, for any (correctly classified) borrower (Stiglitz and Weiss, 1981), again necessitating charging a higher interest rate to obtain a given expected return.

Firms do not, of course, need to invest all that they borrow inside the firm; they can hold additional liquid reserves (treasury bills), and the bank cannot observe precisely what the borrowers do with the funds.

Then even if the overall supply of credit were to have remained fixed, the level of investment would decrease, for several reasons.

Firms that are classified as good are likely to reduce their investment, both because of the increased risk and because of the higher interest rate being charged.18 Furthermore, misclassifications lead to reduce investment: the good firms that are classified as OK reduced their investment more than the bad firms that are classified as good increase theirs. Indeed, the good firms that are classified as OK may choose not to borrow at all, rather than pay the high interest rates associated with such loans.

Note that in this perspective, there are two different senses in which firms may view themselves as credit constrained. Some firms—those classified as bad and some of those classified as OK—simply do not receive credit. They cannot obtain credit at any terms.

Others—the good firms that have been classified as OK—cannot get credit at terms which they think is fair. They know that they are going to repay. Thus, they view the expected cost of a loan charging a very high interest rate to be very high—far higher than does the OK borrower, who knows that there is a higher probability of default. The irony is that the cost to good borrowers is higher than the cost (in terms of expected payments) to lousy borrowers. Good borrowers may choose not to borrow anything at these unfavourable terms. They think of themselves as effectively shut out of the market.

This is similar to the feeling a small business has when it has been turned down for a loan by its bank. There are secondary sources to which it could turn. There is a market specialized in screening loan applicants that have been turned down by banks. When these institutions grant a loan, they charge a high interest rate—exorbitant from the perspective of the misclassified customer, but perhaps accurately reflecting the true default probabilities. (Some of these individuals are, in fact, good risks.) The point, of course, is that the credit market is not a pure-price market. It classifies borrowers. The classification is the central allocative mechanism in this market.

Finally, the recognition of the increased likelihood of misclassification means that good firms face an increase in interest rates more than commensurate with the increased risk of default associated with their loans.

17 We are focusing, in other words, on what is called a pooling equilibrium. One can construct equilibria in which contracts with differing terms, (e.g loan size, interest rate) can serve as the basis of a self-selection (separating) equilibrium.

18 The good firms face an increase in interest rates more than commensurate with the increased risk of default associated with their loans.
borrowers think it more likely that they will be credit constrained (in either of the two senses in which we have used the term) in the future, and this concern induces them to maintain a more liquid position (invest less.)

These are all reasons why, even if the supply of credit today were unaltered by increased uncertainty, investment would be reduced.

Further, however, banks do reduce the supply of credit, for reasons that we have already referred to. The greater likelihood of misclassification and the higher probabilities of bankruptcy make loans riskier, even if the bank is able to increase the interest rate charged to offset. But the bank may find it disadvantageous to increase the interest rate charged, because doing so lowers the expected return. (These effects are in addition to the one to which we previously referred, the reduced equity base of banks leading to a lower willingness to make loans.)

The lower supply of credit means that more of those classified as 'OK' are denied credit, or have their loan size reduced. And this, in turn, lowers investment.

The reduction in bank credit has further multiplier effects on credit availability. While there are some instances where non-bank credit may serve as a substitute for bank credit, in other cases the two may better be looked at as complements. The reduction in bank credit—the refusal of a bank to make a loan—leads to a reluctance on the part of other lenders to extend credit.

The increased risk, lower net equity levels, and greater likelihood of future credit rationing all lead firms to reduce their lending activity.

We can then trace our fourth round and subsequent effects: the reduced credit availability leads to reduced investment, leading to . . .

IV Empirical Implications

The theory we have presented has a large number of empirical implications, which can be compared both with those of traditional Keynesian theory, including the fixed price and menu cost versions, and the real business cycles which have become fashionable in the US in recent years. In a recent Brookings paper, Bruce Greenwald and I (Greenwald and Stiglitz, 1988c) identified certain key properties of labour, capital, and product markets which would enable us to distinguish among the alternative theories. Many of the key empirical observations were either inconsistent or unexplained by the other theories. Using this 'crucial' tests approach to choosing among alternative theories, our conclusion was that the New Keynesian theories, in particular those versions focusing on the consequences of imperfect information and incomplete markets, won hands down.

Here I have only time to provide a couple of examples, to give a flavour of the argument.

(i) Inventories

The cyclical behaviour of inventories has long provided a puzzle for neoclassical and Keynesian economists. With concave production functions, relatively little variability in real interest rates, and shadow wages falling markedly in recessions, inventories should be used to smooth production. Yet they do not; if anything, they exhibit an even more marked cyclical variability than output, and there is a general consensus that they contribute to economic volatility, rather than serving as a buffer-stabilizing force.

Our analysis provides an explanation of this behaviour.

(ii) Sectoral Patterns

Traditional Keynesian theories, relying on the importance of fixed prices, would seem to predict that fluctuations should be smaller in those sectors exhibiting flexible prices. Among the most competitive sectors of the economy is the home construction industry, yet this is a sector which has experienced some of the most marked volatility. Equity and credit rationing are particularly important in these sectors, and thus the magnitude of these fluctuations is perfectly consistent with our views.

(iii) Explanations of price and wage rigidities

Traditional Keynesian theories, while stressing the importance of price and wage rigidities, have not come up with convincing explanations of these rigidities, or of why they should be more important in some sectors than in others. The implicit contract literature, for instance, while explaining why wages paid may not exhibit much variability, has been unsuccessful in explaining why the particular form of wage rigidity associated with implicit contracts should give rise to unemployment. (See Newbery and Stiglitz, 1987; or Stiglitz, 1986.) While many

19 Shadow wages reflect the under-utilization of labour in recessions, evidenced by cyclical movements in productivity. The discrepancy between marginal real (producer) wages in booms and recessions is particularly pronounced if one takes into account over time pay prevalent at later stages of economic recoveries.
cost models have stressed the importance of the costs of adjusting prices, they have failed to take into account the much larger costs associated with adjusting quantities, and the problem to be explained is why it is, when faced with certain disturbances, firms seem more willing to make quantity adjustments than price adjustments. An extension of our analysis focusing on the differential risk associated with different adjustments, has provided an explanation not only of why prices may be more rigid than quantities, but also why the extent of price rigidities may differ in different sectors. (See below and Greenwald and Stiglitz, 1988b, for a more extensive discussion.)

V Government Policy

The theory we have described provides both a rationale for government policy and an explanation of the mechanisms through which it operates.

Interest rates do not work as an equilibrating mechanism. The multiplier processes we have described in the previous section may lead to instabilities. If more credit is extended than would be in 'equilibrium' there is an excess demand for goods; prices rise; there are fewer bankruptcies; equity positions are better than anticipated, leading to still higher demands for investment. Though the higher than anticipated price level means that banks' net equity positions, in real terms, may actually be worse, in spite of the lower default rate, the seeming lower riskiness of lending more than offsets the 'wealth' effect\(^{20}\), leading to still more investment.

The opposite side—an insufficient supply of credit cascading into a credit crisis—is perhaps an even more familiar story.

We emphasized earlier how the credit allocation decisions are made in a decentralized manner. We have noted that in the absence of the coordinating function usually performed by interest rates, there appears to be a role for a central bank, to ensure that the number of certifications for credit worthiness correspond to the number required to maintain full employment.

The mechanisms by which policy works should be clear from the previous discussion. Again, we have noted that monetary authorities may affect either banks' willingness or ability to lend. The latter effect is, of course, the one upon which traditional monetary theory has focused; but while Keynesian analysis traces the effect through the indirect route—the lower supply of loans leading to higher interest rates leading to less investment—in periods of credit rationing, we would argue for a more direct channel.

If our theory is correct, the long-standing confusion about whether it is money or credit through which monetary policy functions becomes more understandable. For the traditional mechanism by which banks issue credit is intimately associated with the creation of money. A credit line is created. The firm draws upon the credit line when funds are needed; a demand deposit—that is money—is created precisely when the firm plans to spend more, e.g. by undertaking a new investment.

Our views also reconcile some of the controversies between monetarists, Keynesians, and neoclassical economists. We agree with the monetarists that monetary policy may not primarily work through an interest rate mechanism and that interest rates may not provide a good target for monetary policy (except to the extent that interest rates are correlated with inflation rates, through Fisher's law). But we agree with the Keynesians that a major channel of monetary policy is through its effects on investment and firm's willingness and ability to produce. Our view differs from monetarists, Keynesians, and neoclassical economists in its view of the details of how monetary policy affects the economy: it is not through a transactions mechanism (money burning a hole in people's pocket, driving them out to buy goods or driving down the interest rate), but through the credit mechanism.

Some Questions

A long-standing puzzle facing conventional monetary economics is that, while it assumes that the government may control the nominal money supply, it does not control the real money supply. If the classical dichotomy were correct, an increase in the money supply would instantaneously lead to an increase in prices, negating the effect of the monetary expansion.\(^{21}\)

The traditional Keynesian way of resolving this is simply to assume price rigidities. We have

\(^{20}\) Banks may be slow in calculating their real equity position; they may be fooled by the higher than expected nominal wealth.

\(^{21}\) Although again it is possible to argue that monetary policy—defined as a rule which specifies the magnitudes by which the nominal money supply will be increased in indifferent circumstances—will have real effects, because the probability distribution of returns to holding financial assets will thereby change, this effect is generally viewed to be a second-order effect. See Stiglitz (1983).
already noted that the recent work trying to justify this in terms of menu costs—the costs of changing prices—seems most unpersuasive. Again, it seems an instance of macro-economists focusing on third- or fourth-order effects. Surely, these adjustment costs are of an order of magnitude smaller than the costs of adjustments associated with quantities. If our objective is to explain why the economy adjusts more in terms of quantities than prices, we have to look elsewhere.

As we have observed, our theory provides an explanation of the relative rigidity of prices as well. For if firms are risk averse (as they will be, with either principal agent problems or if firms are equity constrained) and if there is instrument uncertainty, that is, uncertainty about the consequences of a change in any action, then the extent to which a variable will be adjusted will depend both on the degree of risk aversion and on the uncertainty of its effect.

The risk associated with reducing production is relatively little, particularly if (as Bob Hall has argued) marginal costs of production are constant, or decrease or increase only slightly. For what is entailed is basically an intertemporal substitution of production which, under the stipulated conditions, has relatively little costs. The sole risk is that of a stock-out of inventories, and given the availability of excess resources, remedial actions can quickly be taken to avoid that should it prove imminent. On the other hand, there is considerable risk associated with changing prices and wages, partly associated with how others (competitors, employees) will respond.

Unless firms are perfectly informed concerning the money supply, and unless they all believe, with complete confidence, in the classical dichotomy, firms will not increase prices proportionately in response to an increase in money supply. And, in this sense, their lack of confidence in the classical dichotomy is self-confirming.

Accordingly, a monetary-credit expansion will have real effects, even if all contracts were fully indexed. For so long as prices do not instantaneously increase in proportion to the credit/money creation, access to resources will be altered by a change in nominal credit availability. In particular monetary policy affects different sectors of the economy differently. If the government were to make credit more available, it is not as if each firm in the economy immediately had its deposit accounts increased proportionately. Some sectors are more credit rationed than others. The increased credit availability will accordingly affect relative prices, and this will lead to all the kinds of effects described in our previous discussion.

Of course, there is not full indexing. Our theory provides an explanation of why unanticipated inflation—or an unanticipated decrease in the rate of inflation—will have even stronger real effects, given that there is not complete indexing. Since, both at the time of the Great Depression as well as today, there is not complete indexing, one might argue that this proviso is not of much moment. But purists might claim that, until we have explained the absence of indexing, we have an incomplete theory. But our theory even provides a suggestion of why indexing may not occur.

Assume there were some, say, oligopolistic industry in which all firms believe that other firms do not adjust their prices fully in response to inflation. If each firm believes the others will not do so, it does not pay for it to do so: one can show there exists a rational expectations equilibrium with nominal rigidities. (It would be excessively risky for any firm to depart from the seeming convention, to try to establish a new convention.) Such firms will be unwilling to borrow or lend with fully indexed bonds, and would not change their prices proportionately to changes in the money supply.22 This does not explain why there isn't more indexing—although it should be noted that in economies in which inflation rates are high and variable, indexing does occur. Indexing clearly does not resolve all the relevant macroeconomic issues: such economies still seem subject to macroeconomic fluctuations, and monetary policy may still be effective. Again, our theory provides an explanation23, an explanation which is missing from traditional Keynesian theories which rely on nominal price rigidities.

Further Implications for Policy

Let me conclude my talk by drawing attention to three of the policy implications of our theory.

1 Social vs. Private Returns to Increasing Speed of Recording Transactions

22 Several other models, based on imperfect information, with nominal wage and price rigidities, have been formulated. See Sargent (1984, 1987a, 1989).

23 Lest we claim too much for our theory, it should be noted that we show that there exist Nash equilibrium exhibiting nominal price or wage rigidities (Stiglitz 1985b, 1987a,b), there are other equilibria which do not exhibit these nominal rigidities, but show real rigidities. We do not explain how it is that a particular equilibrium is "selected." This may be asking too much of a theory: the explanation may simply lie in history.
One major change in recent decades, to which we have referred earlier, is the faster speed with which transactions are recorded. We emphasized earlier the importance of having a good accounting system—this was one of the central roles of banking institutions. But it does not follow that there is much of a social return to recording a transaction one second faster, though given the conventions of our accounting system, there may be a private return from doing so. Elsewhere (Stiglitz and Weiss, 1988), we have shown that expenditures on such increases in speed in recording may be welfare decreasing.

2 Incidence of the Corporation Income Tax

Our theory has emphasized the importance of equity constraints and the availability of funds. Government policies affect not only the marginal return to, say, investment, but also the supply of funds within a firm. Accordingly, the incidence of the corporation tax may depend on average tax rates as well as marginal tax rates.24

3 Stabilization Policy

We have noted the importance of equity rationing in business fluctuations. An effective stabilization policy of the government should be directed at overcoming the limitations of this rationing. We can think of standard investment tax credits as grants to firms, grants which are cut off as soon as firms profits become too low. But those are precisely the times when the firm needs funds the most. Government policy exacerbates the limitations of financial markets, rather than offsetting them. The government's ability to tax implies that loans made through the tax system have somewhat different properties than loans made by banks; moreover, the objectives of government loans should be somewhat different from those of banks, which are only concerned with the returns they can obtain. The government is, after all, a silent partner in all business enterprises, a partner, however with somewhat different terms than those facing other partners. I emphasize these differences because I wish to avoid the objection: won't government funds simply displace private funds?

The importance of discrimination—of ensuring that funds do go to good investment projects—is sufficiently great to make a government program without screening likely to be socially unprofitable. The difficulties of ascertaining credit worthiness are sufficiently great to make government screening either excessively bureaucratic or particularly vulnerable to corruption. This is why one needs to look for self-selection schemes, where those who have good prospects (as evidenced, for instance, by their willingness to put up their own capital) seek funds. Though this is not the occasion to provide a detailed proposal, a system whereby firms could borrow limited amounts for new investment (say 10 per cent of the capital cost), to be repaid out of future earnings would ease the capital constraint and, if the loans were made at sufficiently favourable terms, increase the returns so as to encourage investment during slack times.

Concluding Remarks

I began this lecture somewhat in the spirit of 'the true confessions of a converted monetarist'. By now, it should be clear that my conversion has been far from complete. I remain skeptical of the relevance of the 'M' so loved both by Keynesian and monetarist economists. Money as a medium of exchange is, if not irrelevant, close to being so. It is not money that makes the world go round, but credit. It is not easy to describe the supply and demand for credit, its creation and destruction. One cannot simply write down a production function. The process of judging credit worthiness involves the collection and processing of information; changes in economic circumstances can easily destroy the relevance of previously accumulated information.

The close links between the money supply and credit creation have naturally caused a confusion between the two. If money creation were perfectly linked with credit creation, one might well object to the traditional stories told about how money affects the economy; but the policy implications might not be far from the mark (so long as one used money supply as the target variable, not an intermediate variable, such as the rate of interest.) But the link is far from perfect, and changes in institutions, technology, and policy can affect that link. Our monetary institutions become one, important set of institutions within a broader set of institutions to provide credit, monitor loans and, more broadly, to facilitate intertemporal transactions. I hope that I provided here an alternative framework within which we can begin to think systematically about monetary policy and credit institutions from this broader perspective.

REFERENCES


— (1987c), 'Financial Structure and The Incidence of the Corporation Tax', Stanford University, mimeo


— (1985a), 'Credit Markets and the Control of Capital', Journal of Money, Credit and Banking 17(1), 133-52.


— (1988a), 'Banks as Social Accountants and Screening Devices for the Allocation of Credit', in A. Courakis and C. Goodhart (eds), The Monetary Economics of John Hicks, Macmillan.

— (1988b), 'Credit Rationing and Its Implications for Macro-Economics', Stanford University, mimeo.