The theory of credit and macroeconomic stability

*Joseph Stiglitz*

The post-2008 world has been one dominated by monetary policy, as politics and ideology – and sometimes financial markets – constrain the use of fiscal policy. There have been massive increases in the balance sheets of key central banks – the Federal Reserve’s reaching 25 percent (2016) of GDP, Japan, 82 percent (2016), the Bank of England, 21 percent (2016), and the ECB, late to embark on quantitative easing, but as of 2016 already over 31 percent of GDP. But in spite of these increases, the best that can be said is that monetary policy prevented matters from becoming worse; growth in GDP in the advanced countries was an anemic 2 percent.

The growth in base money has become disjointed from the growth in the economies. Figure 7.1 shows the growth in central bank assets and the growth in real GDP for each of the four countries. Rather than GDP growing proportionately to the growth of the central bank balance sheet, the figure shows significant variability in the ratio of central bank assets to GDP, and especially large changes in the money supply being associated with small changes in nominal GDP in recent years in the US.

A simple regression shows a very low correlation between money supply and GDP in recent years, weaker than in the period immediately after World War II. This weak relationship appears robust to a variety of specifications, including variable lags and different measures of money (e.g. the Fed’s balance sheet or more standard measures of M2 – see Figure 7.2). These results naturally raise the questions:

* This is a chapter prepared for the volume in honour of Deepak Nayyar, from whom I have benefitted enormously in our many conversations and joint work over many years. The author is indebted to the Institute for New Economic Thinking for financial support, and to Martin Guzman and Andrew Kosenko, for comments and suggestions. This chapter represents a summary of my research in this area and hence the disproportionate number of references to my earlier work.
where is the extra liquidity provided by the Fed going? What’s happening? Standard theory suggests putting more money into people’s pockets should lead to more spending, leading either to higher prices or greater output. If this isn’t happening, it suggests a fundamental flaw with standard formulations of monetary theory.

The absence of a clear link between money (however measured) and output (nominal or real) has led naturally to a shift of attention of monetary authorities away from quantitative measures (base money, M2, etc.) to a focus on interest rates. But even here, without further massaging of the data, the relationship is weak. The Appendix discusses the weak relationship between output (nominal and real) and money supply and interest rates (nominal and real). Our empirical investigation suggests, moreover, that the relationship has not been stable over time. In particular, the relationship between money and output has become weaker in the last quarter of a century. As the later analysis makes clear, this should not come as a complete surprise: there have been large changes in institutional arrangements, and one might have expected such institutional changes to be reflected in the relationships discussed in the Appendix.

In the aftermath of the Great Recession, there is a growing consensus, even among central bank officials, concerning the limitations of
monetary policy. Central banks may have prevented another Great Depression, but they have not restored the economy to robust growth. Our analysis suggests that this experience sheds broader light on the limitations of monetary policy. The first part of the chapter provides an explanation for this extraordinary ineffectiveness of monetary policy, and in doing so provides a new framework for thinking about money and finance.

The second part of the chapter builds on the insights of the first part and shows how advances in technology allow for the creation of an electronic monetary system that enables better macroeconomic management and a greater share of the rents associated with “money”, that is, with the payments system, to be captured for the public treasury.

Towards a new theory of money and credit

Standard modern monetary theory is based on the hypothesis that the T-bill rate is the central variable in controlling the economy, and that the money supply, which the government controls, enables the government to regulate the T-bill rate.

Prevailing economic doctrines earlier argued that there was a simple link between the supply of money (say M₂), which the government could control, and the value of nominal GDP. This link is described by the equation:

$$MV = pQ$$

where, V is the velocity of circulation, p is the price level and Q is real output. (1) is essentially a definition of the velocity of circulation. Monetarism translated (1) from a definition into an empirical hypothesis, arguing that V was constant. This meant that nominal income and the money supply moved in proportion.

Monetarists like Milton Friedman claimed further that (at least over the long run) Q was fixed at full employment, so that an increase in M would lead to a proportionate increase in p. Shortly after these monetarism doctrines became fashionable, especially in central banks, the links between money supply (in virtually any measure), and the variables describing the economy (income, or even real interest rates) seemed to become tenuous. The velocity of circulation was evidently not a constant. Of course, there never had been a theory explaining why it should be.

Even before this, Keynesians had argued that V was a function of the interest rate. An increase in M is split in three ways – an increase in p, an increase in Q, and a decrease in velocity – with the exact division depending on the relevant elasticities (e.g. the interest elasticity of the demand for money, of investment, and of consumption.)

But, beginning in the 1980s, velocity was not only not constant, but it also did not appear to be even a stable function of the interest rate – not a surprise given, as we have noted, the large institutional changes going on in the financial sector (such as the creation of money market funds and the abolition of many regulations). The natural response was a switch from a focus on the quantity of money to the interest rate. But while this experience should have led to a deeper rethinking of the premises of monetary theory, it did not.

Prevailing theories also held that monetary policy provided the best (most effective, least distortionary) regulator of the economy, and that the way it did this was through adjusting the interest rate. A lowering of the interest rate led to more consumption and investment. In an open economy, it led to a lower exchange rate, which led to more exports. The extraordinary ineffectiveness of monetary policy to restore the economy to full economy in the aftermath of the Great Recession has led to a modification of the standard theory: monetary policy is the instrument of choice so long as the economy is above the zero lower bound (ZLB); and to the extent that the ZLB can be breached, it should be.

This chapter questions the primacy given to monetary policy, suggesting that the problem is not the ZLB, but a host of other limitations – effects of monetary policy that were given short shrift. Most fundamentally, we argue that standard theory has given too much attention to the interest rate and too little attention to the primary mechanism through which monetary policy affects the economy, the quantity and terms (including the non-price terms) at which credit is available. In normal times, money and credit represent two sides of a bank’s balance sheet, so they may be highly correlated. But more generally, and especially crises, credit may be only weakly related either to the supply of money, or even to the T-bill interest rate. This weak link and not the ZLB – helps explain the ineffectiveness of monetary policy at certain times, such as the period from 2008 through the present. We further argue that the expansion of credit itself is weakly linked to GDP, with increases in credit going to more uses other than an increase in the demand for produced goods – most notably, towards the acquisition of assets such as land.

The discussion in this section is spread across seven subsections. After setting out the basic argument for the focus on credit in the first, we turn to the determinants of the supply of credit – primarily through
the banking system, observing that changes in monetary policy may be limited in overcoming other changes in the determinants of credit
availability, in the second. The third focuses on the demand for credit,
noting that there are many other uses to which credit can be put other
than an increased demand for produced goods. The fourth then turns
to a more expansive explanation of the ineffectiveness of monetary
policy. The fifth subsection explains that the distortional effects of
monetary policy may be far greater than earlier analyses have assumed;
for instance, the conventional use of an aggregative model hides inter-
sectoral distortions. The sixth argues, by the same token, that there
may be serious adverse distributional effects which cannot be ignored,
and which contribute to the ineffectiveness of monetary policy. The sev-
enth reexamines these issues from the perspective of an open economy,
explaining why monetary policy may be more or less effective, and
more or less distortional, with a different set of distributive effects.

The analysis of the relative ineffectiveness of monetary policy pro-
vides the background for the second part of the chapter, where we show
how a move to an electronic banking system, combined with a
direct focus on credit availability, and the use of new monetary instru-
m ents described there, can increase the effectiveness of macroeconomic
management, even in an open economy.

The importance of credit—not money

In earlier work, Greenwald and Stiglitz (1991, 2003) argued that what
matters for the level of macroeconomic activity was neither the supply
of money (the quantity variable upon which monetarism was focused),
nor the T-bill rate (the rate of interest that the government had to pay
on its short-term bonds, and the focus of recent monetary policy),
but rather the availability of credit and the terms at which credit is
available. They thus criticized standard monetary theory in terms of
its theory of the determination of the lending rate, the relevance of the
T-bill rate, and the assumption that credit markets always clear.

In the standard model, the interest rate is determined by the intersec-
tion of the demand and supply for money. Government controls the
supply of money. In that model, the demand for money is related to
income and the interest rate (with the interest rate being the opportu-
nity cost of holding money). But G-S point out that in a modern economy,
money is interest bearing (e.g. money market funds), with the cost of
holding money a matter solely of transactions costs, unrelated
to either monetary policy or the level of economic activity (see Fig-
ure 7.3). Moreover, money is not required for engaging in transactions,

Figure 7.3 The relationship between T-bill rate and money market rate
The two track each other almost perfectly, the difference being largely transac-
tions cost, with no significant cyclical component.

Source: Federal Reserve Economic Database (FRED) [T-bill: https://fred.stlouisfed.org/
series/TB3MS;
BlackRock: www.sec.gov/Archives/edgar/data/65109/000119312516426010/d79618
d497k.htm]

but rather for credit. Even if money were required for transactions,
most transactions are exchanges of assets, and not directly related to
the production of goods and services; hence the demand for money is
related not just to the level of macroeconomic activity ("Y", GDP),
but to other kinds of transactions, and there is no fixed relationship
between these and GDP. There is, in short, no theoretical foundation
underlying the usual theory of interest determination.

Robertson's had earlier proposed an alternative theory of interest
determination, based on the demand and supply of savings. Some
farmers decide not to consume or plant all their seeds, and some wish
to use more than the seeds they have available, and the interest rate
equilibrates the supply and demand of "loanable" seeds (see Green-
wald and Stiglitz 2003). While such a theory may have made sense in
a primitive agriculture economy, it does not describe a modern credit
economy, where banks are central and can create credit within con-
straints imposed by the government. In particular, there is no need for
a bank to have seeds on deposit for it to create credit.
While there is thus a lacuna in the theory of interest rate determination, even were we to have a well-developed theory, with a clear link between the interest rate and monetary policy, there is a further problem: it is not clear that the T-bill rate (so determined) plays the critical role assumed in modern macro and monetary theory. First, as G-S show, the T-bill rate is only loosely related to the lending rate. Moreover, the lending rate is not the only variable affecting macroeconomic activity. With credit rationing (Stiglitz and Weiss 1981), the availability of credit matters too, as do other non-price terms of credit contracts (like collateral requirements; Stiglitz and Weiss 1986). These are endogenous, and while they may be affected by the T-bill rate, they are also affected by other policy and environmental variables. In short, modern macroeconomics has focused on certain substitution effects (e.g. the interest elasticity of consumption), and these may be (and we would suggest typically are) overwhelmed by income, wealth, risk, and other non-price effects, or price effects operating in other ways, for instance through their impact on collateral, self-selection, or incentive compatibility constraints.

The correlation between money and credit

Our analysis emphasizes the role of credit in determining the level of economic activity. For a variety of reasons, data on the money supply [measured somehow] seems more widely available than data on credit, either its “availability” or even the actual level of lending. But these variables are closely related: typically, when a bank lends more, its deposits (or more broadly, the deposits of the banking system) increase (a liability) and so do the bank’s assets – the loan. Thus, money (demand deposits) and credit increase in tandem. So too, if a foreigner were to make a deposit in a country’s bank, the bank would normally have an incentive and ability to increase lending.

But as we explain later, there are times when this normal relationship breaks down, and policies predicated on the normal relationships may be very misguided. If a bank faces a great deal of uncertainty, it may not lend out as much as it could; it has excess reserves. In the East Asia crisis, the IMF became worried when, say, there were large excess reserves in Indonesia. It meant that, suddenly, the banks could start lending, and that would be inflationary. As a precautionary measure, it thought it was wise to “mop up” the excess reserves or to take other actions to eliminate the excess reserves, for example, tighten reserve requirements. The problem was that with the blunt instruments available, even banks that had no excess reserves were typically affected by the tightening. Their customers lost access to credit – deepening the on-going recession. The cost of tightening was palpable; the risk of inflation that the tightening was supposed to reduce was imaginary – there was virtually no realistic scenario in which the banks with excess reserves would turn around and lend so much that inflation would be excessive.

Economies in deep downturns – recessions and depressions – behave differently than those in more normal times, and policies, including especially monetary policies, suitable for one situation may not be suitable for the other (see Stiglitz 2016d). Even if the correlation between money and credit were close in normal times, it is not in deep downturns, as banks are willing to hold on to excess reserves. As we explain later, it is this which gives rise to the modern liquidity trap.

The supply of credit

In standard monetary theory, banks play no role. This is true even for the models used by central banks – ironic since if there were no banks, there would be no central banks. In institution-free Neo-classical economics, one sees underneath the institutions to the underlying economic forces. Thus, as we have noted, in standard models, the (real) interest rate is set at the rate that equilibrates the demand and supply of funds (in Robertsonian monetary theory; in Keynesian theory, the demand and supply of money). Though that model may provide a reasonable if incomplete description of the capital markets on which large enterprises raise funds, small and medium-sized enterprises (SMEs) have to rely on banks, and much of the variability in economic activity is related to investment by such enterprises; and much of that variability is related to credit availability. Interest rates are not set at the intersection of demand and supply curves – there may be credit rationing; but even when there is not credit rationing, the supply curve of funds needs to be derived from the behavior of banks, and when one does that, one gets a very different picture.

Greenwald and Stiglitz (1991, 1993b, 2003) provide a simple model describing bank behavior, showing how lending is related not just to the T-bill rate, but to banks’ net worth, their risk perceptions, their existing portfolio of assets, and the constraints provided by regulators. They describe too how banks adjust not only their lending rate, but the other terms of the contract in response to changes in these variables. Thus, credit (money) supply is determined not just by conventional monetary instruments (open market operations, reserve requirements), but also by macro- and micro-prudential requirements. Indeed, the
two aspects of central bank policy (regulatory and macro-control) cannot and should not be separated.

Their model of banks (combined with their earlier model of the risk-averse firm\textsuperscript{13} facing equity rationing; see Greenwald and Stiglitz 1993b; Greenwald, Stiglitz, and Weiss 1984\textsuperscript{12}) thus shows how changes in economic circumstances today (such as a shock that affects their net worth or even the value or risk of particular assets\textsuperscript{13}) can have large, long-lasting effects. The effects of an economic shock can be persistent. At the same time, they explain why an increase in liquidity – a conventional open market operation, lowering the T-bill or the lending rate – may have little effect on credit availability.\textsuperscript{14}

Banks typically respond to a lower cost of funds by lending more, and lending at lower interest rates (whether they choose to ration credit or not). But there are some circumstances in which they do not, or do not do so to any significant extent. In particular, G-S explain why, if risk perceptions have increased and if the risk of the banks’ existing portfolio has increased – that is, the risk of both new and past loans has increased – the bank may be at a corner solution, where it will not undertake further loans, even when the interest rate is lowered. And this is especially so if, due to asymmetric information, the bank can only divest itself of the risk associated with past loans by taking large capital losses on its loan portfolio.\textsuperscript{15}

This problem is exacerbated by the fact that in severe economic downturns, the value of highly leveraged banks’ net worth is severely decreased, so risk-averse banks are even more overly exposed to risk, unless they cut back severely on lending. (The inability to divest oneself of risk generates an important hysteresis effect. There are, in addition, effects on banks’ optimal portfolio, e.g. shifting away from more risky lending.)

Changes in government (central bank) policy, as desirable as they may be, typically give rise to new risks, which have their own adverse effects even when the intent of the policy change is to stimulate the economy. Thus, a decrease in the interest rate changes asset values of different firms in different ways, depending on their assets. A firm that has outstanding short-term liabilities and long-term assets (with returns fixed at a higher rate) may be much better off, but a firm with a different maturity structure of assets and liabilities could actually be worse off. Lenders may lose which have to have detailed information about all the assets and liabilities of a firm to know precisely how each firm is affected; and in the absence of that information, uncertainty will have increased. Thus, an increase in the interest rate will have a more adverse effect than anticipated, but a lowering of the interest rate will have a smaller effect – or even an effect that is adverse. This is especially so once one takes into account all the general equilibrium effects. A lowering of the interest rate will lower the exchange rate, thus hurting importers and domestic firms that use imported inputs.

With risk aversion, the benefits of the winners from such changes in relative prices do not offset the losses of the losers. The aggregate effect can be negative (Greenwald and Stiglitz 1993b).

The ability and willingness of banks to lend depends not just on what may be called the environmental variables (risk perceptions\textsuperscript{16} and net worth) described in earlier paragraphs, but on all the constraints facing banks today – and the expectations of future constraints. For instance, banks face capital adequacy constraints, specifying, say, net worth relative to outstanding loans. If that constraint is tightened, then the bank must either raise new capital or reduce outstanding loans. But because of capital market imperfections, firms in fact typically face constraints in raising new equity; at the very least, doing so may be very costly to existing shareholders.\textsuperscript{17} Hence, an increase in the capital adequacy ratio – or an increase in defaults on loans that reduces capital – reduces lending. But because a quick reduction in lending may be costly, firms need to anticipate that they might face such a situation, and hence well before the constraints bind, banks may curtail lending and firms may curtail borrowing.

This simply emphasizes that all of the constraints facing a bank – whether binding today or possibly binding in the future – can affect lending and borrowing today. And it is not just the standard instruments (e.g. open market operations or the discount rate) by which central banks affect lending activity.

Banks that focus on lending to SMEs face an additional problem: this lending is typically collateral-based, and the collateral is typically real estate. In a crisis such as that of 2008, the value of this collateral decreases enormously, and thus, given existing rules and constraints, banks should significantly reduce the amount of their exposure to SME risk. The focus of the bank is thus on reducing SME exposure, not making new loans.\textsuperscript{18}

By the same token, severe economic downturns are often associated with increased disparities in judgments (probabilities associated with different contingencies). This increased disparity in judgments may give rise to an increase in trading in existing assets, rather than for newly produced assets, and an increase in the demand for credit to support such trades. To see this, consider the 2008 crisis. Some believed that the market had overshot – real estate prices had fallen excessively. The banks argued that that was the case, not wanting to believe that they
had made massive misjudgments about the real estate market. The more optimistic market participants believed this contention, and were willing to pay a risk premium to get access to funds to buy these depressed assets. The banks agreed with their judgments (for reasons given in the previous sentence). These new borrowers could offer the real estate (at the new low price) as collateral. Thus, from the perspective of the bank, these new real estate loans offered a low-risk (in their calculus), high-return loan—far better than the high-risk loans to real firms. From the perspective of the banks as a group, this lending has a further benefit: it raises real estate prices, improving the value of their existing portfolio.19

In short, in a deep downturn changes in the balance sheet of the bank and in its risk perceptions typically lead to a significant contraction in the supply of funds and an increase in the interest rate that it charges, and to corresponding changes to its non-price terms; the magnitude of these effects overwhelms any ability of the central bank to stimulate lending by lowering interest rates and other actions designed to ease credit availability.

Of course, when we observe a net contraction in lending in a recession it does not necessarily mean that monetary policy has been totally ineffective: it simply means that it was unable to fully counteract the other effects.20 And even when we see an expansion of credit, it does not mean that monetary policy has been effective: the expansion of credit may not have facilitated the purchase of newly produced goods, and thus may not have contributed to an increase in GDP.

Moreover, a lowering of interest rates on T-bills does not translate into a lowering of the lending rate, and it is that rate that matters for firm and consumer behavior. Further, even that rate may not provide an adequate description of the financial market: there may be credit rationing and collateral and other non-price terms.

We have even identified some circumstances in which lowering T-bill interest rates may be counterproductive (we will identify some further circumstances later), because of the increased risk associated with the change in interest rates that increases risk perceptions (in association with the other relative prices effects generated). By the same token, negative interest rates may adversely affect banks' balance sheets, if not carefully designed, and, in doing so, lead to a contraction in lending.21

The demand for and uses of credit

The previous section focused on the determinants of the supply of credit, explaining in particular why an “easing” of monetary policy might not result in a lower lending rate and a greater availability of credit. Here, we explain why the same thing is also true on the demand side: in a severe downturn, risk-averse firms will face an adverse shock to their balance sheets and an adverse increase in their risk perceptions, both of which will lead to a contraction in production and investment. Lowering interest rates at which they can borrow (which is not the same as lowering the T-bill rate, as we explained in the previous section) may lead them to borrow more than they otherwise would have borrowed, but this increase may be small compared to the contraction in investment from the increase in risk and worsening of the balance sheet. Moreover, even when interest rates are lowered— for those who can get loans—credit may be rationed.

In addition, for reasons explained earlier, changes in interest rates by themselves can give rise to an increase in uncertainty, with adverse effects on the demand for credit. Each firm is embedded in a complex general equilibrium system, in which it has an array of assets and liabilities, some explicit, some implicit, in part related to its economic relations with other entities. A marked lowering of interest rates can increase uncertainty and the perception of risk, and firm risk management may entail a corresponding adjustment in its activities, including decreases in production and investment. Later in this chapter, we shall identify some distributional effects of lowering interest rates that may also result in a reduction in the demand for credit as interest rates fall.

While a change in interest rates thus may not be effective in increasing the demand for and use of credit, even when it does, the increases in credit (money) do not necessarily translate into increases in economic activity—greater consumption or increased investment in newly produced capital goods; there is many a slip between the cup and the lip. Increases in credit (money) can go into several uses:

1. Increased purchases of existing assets, and especially land. Indeed, much of increased wealth is an increase in land values—so much so that the ratio of the value of produced capital to GDP is actually declining.22 Of course, when more money goes to the purchases of land, it does not lead to more land, but rather, to an increase in the price of land. This wealth effect can lead to more real spending, but this effect is normally likely to be small—far smaller than that which would have been predicted by any model where it is simply assumed that the increase in money leads to more spending on produced goods.

2. Increased margin to facilitate taking larger speculative positions, for example, in zero-sum bets, such as futures markets.
Increased lending abroad (either for “productive” or non-productive purposes). If the foreign country to which the money goes has an increase in income, it may (slightly) enhance exports, and exports may be further strengthened from the effect on foreign exchange. (Monetary policy in an open economy is discussed briefly further in a later subsection.) It is, accordingly, not surprising that the link between money and economic activity may be much weaker than standard monetary theory assumed.

Limitations on the effectiveness of monetary policy

Liquidity trap and the Zero Lower Bound (ZLB)

The Greenwald-Stiglitz analysis provides an alternative explanation of the “liquidity trap” to that of Keynes. Keynes’ explanation of the inefficacy of monetary policy was that because the demand curve for short-term government bonds becomes horizontal at low interest rates, monetary policy could not push interest rates down below a certain level. Empirically, recent experiences have shown that the interest rate on government bonds can even become negative. Our argument focuses on banks, and their unwillingness to lend more under certain circumstances, no matter how low the T-bill rate is pushed.

So too, our analysis provides a counter to the recent fixation with the constraint on monetary policy imposed by the ZLB. It takes the view that even if the interest rate were lowered below zero, the response would be limited, largely because banks would still not increase their lending, partly because banks would not (fully) pass on the lower cost of funds to their customers, but partly too because the interest elasticity of investment and consumption is low. Of course, if the interest rate became negative enough, to the point where individuals could borrow and effectively never repay, then there would be an increase in economic activity. But that is not what advocates of the ZLB mean.

Elsewhere, I have provided other arguments for why the ZLB argument is questionable: if it were true, there would be other ways of achieving the desired change in intertemporal prices, through investment tax credits and consumption tax rates that change over time. Yet no one is proposing such a scheme. Doing so would provide a test of the hypothesis, and I believe the ZLB theory would be shown to be wanting.

Diversion of credit creation and the creation of instability

Our analysis also provides an additional explanation for the ineffectiveness of monetary policy even short of the ZLB: standard monetary theory assumes that any additional liquidity created goes towards the purchase of produced goods. But, as we have noted earlier, much of the additional liquidity does not go to the purchase of newly produced assets, but rather into existing fixed assets (such as land), helping create credit bubbles, and into institutionally constrained “gambling” transactions in futures markets in which some form of margin has to be put up. This diversion helps explain the regression findings noted in the beginning of this chapter, showing a low correlation between (the change in) money and the (change in the) value of output.

The observation that increases in credit go into increased speculation and an increased value of fixed assets helps explain why a low interest rate environment is often associated with financial instability. (Other reasons are associated with the distortionary effects of monetary policy discussed in the next subsection.) Guzman and Stiglitz have shown, for instance, that increased gambling in futures markets leads to an increase in what they call pseudo-wealth: all of the market participants believe their wealth goes up as they make more of these bets, simply because they expect to win. But the bets are zero sum: the gains of one person occur at the expense of others. Still, the extent of such gambles can change suddenly, as happened in the lead up to and the aftermath of the Great Recession – and thus the amount of pseudo-wealth can change quickly, and so too the level of aggregate demand. If monetary or regulatory policy tightens, then the extent of such gambles may decrease, and so too the value of the pseudo-wealth. The same thing occurs if there are changes in perceptions and/or the willingness to engage in such bets.

Similarly, if credit is used to finance the purchase of fixed assets, like land (and/or there is borrowing on the basis of land as collateral), an increase in credit can give rise to an increase in the price of land, which, if monetary policy is sufficiently accommodating, can lead to more lending, fueling further increases in prices. This credit-collateral spiral can suddenly break, for example, when market participants no longer believe that the price of land will continue to rise – and in fact, it can be shown that it is impossible for prices to continue to rise forever at the rate necessary to satisfy the capital arbitrage equation (giving the same rate of return across all assets). (See Shell and Stiglitz 1967; Hahn 1966; Stiglitz 2015b.)

The problem is not just that additionally provided liquidity goes to these purposes, which do not directly lead to an increase in GDP. It is
also that the proportion of any additional money that actually goes to support GDP is highly variable. Hence, without further constraints, monetary authorities cannot be sure about the link between GDP and money (credit).

Of course, if there were a stable relationship between the nominal or real interest rate and GDP, then it could expand or contract the money supply until it reached the targeted interest rate. But the discussions of preceding sections made clear that the relationships were also highly variable between the T-bill rate and either the supply or demand for money/credit on the one hand, and between the T-bill rate and the level of economic activity on the other.

**Distributive effects**

In a later section we explain how monetary policy may have adverse distributive effects. There are winners and losers – but if the reduction in spending by the losers is greater than the increase in spending by the winners, then the net effect on aggregate demand may be negative, and these distributive effects may again overwhelm the direct interest rate effect leading each to spend more than they otherwise would have. Moreover, the adverse distributive effects may be compounded by the rationing effects described earlier: the losers may be forced to contract their spending, while the gainers may choose to increase their spending only a little; and the lower interest rates may then have no effect on the former.

The argument is parallel to that which has become standard in international economics. There has long been a concern about persistent global imbalances – China’s and Germany’s surpluses, and the US deficit. The worry is that there will be a “disorderly unwinding” of these imbalances – that if global financial markets suddenly stopped being willing to finance the deficits of the deficit countries (Calvo 1998), the contraction of their spending would not be offset by the expansion of the spending of the surplus countries (see Stiglitz 2010).

**Distortionary effects of monetary policy**

Advocates of the use of monetary policy often argue that it is preferable to fiscal policy not only because it can be implemented more quickly, but also because it is less distortionary. That is one of the reasons that so many of those economists supporting the view that monetary policy should bear the brunt of macroeconomic adjustment have been so disturbed by the ineffectiveness of monetary policy in recent years, and why the ZLB argument has become so popular. For it says that their prior view was correct, but that there is a special “regime” under which, when interest rates hit zero, the results are no longer applicable.

But those conclusions are made in the context of highly special models. In this section, we note several reasons why the conclusion that monetary policy should be at the center of macro-stability may be wrong.

**Misparging of risk**

Market participants talk about how the recent low-interest environment leads to a distorted price of risk. They argue that the low-interest environment leads to a distorted price of risk because of the “search for yield”, and that, in this low-interest-rate environment, in order to get “yield” there is excessive demand for risky assets yielding a slight risk premium. That drives up the price of these assets, driving down risk premiums to irrational levels, which eventually get corrected.

The consequences of this mispricing have been severe: funds flow into uses where, with more rational pricing, they would not go. And the later readjustment of prices can itself have severe consequences.

But there is a kind of intellectual inconsistency in this perspective. Financial market participants typically believe in the efficiency of markets. That traditionally has been part of their argument against government regulation. But the entire argument for why there is mispricing is based on behavioral finance: market participants fail to take into account the fact that the irrationally low levels of risk premiums will not be sustained.

There are risks associated with such market irrationality – but market irrationality does not just suddenly appear as interest rates get near zero. Market irrationality is pervasive. And because of this, and because of the macroeconomic externalities that are associated with the consequences of both the excessively low risk premiums and of the corrections that follow, there is a need for much greater market regulation than advocates of unregulated markets claim. They cannot have it both ways: to claim that markets are efficient, but that we need to be wary of low interest rates because they create distortions in the price of risk.

They are, however, perhaps correct in their warning against low interest rates, providing a quite different argument for the limitations of monetary policy than provided by Keynes: it is not that interest rates cannot be lowered (indeed, some central banks have lowered interest rates below zero); nor is it that lowering interest rates will not
have much effect on real economic activity (true, and even, as we argue later, worse than that); but that the consequences of low interest rates mean that central banks should eschew such policies, especially over an extended period of time.

Inter-sectoral misallocations

The aggregate models so beloved by macroeconomists hide a key problem with the excessive reliance on monetary policy: it gives rise to intersectoral distortions. It makes interest-sensitive sectors bear the brunt of adjustment. It may be desirable to make such sectors bear more of the costs of adjustment than others; but there may be (and there typically is) a cost to the reliance on monetary policy.

Optimal macroeconomic policy would distribute the costs of adjustment, which requires both monetary and fiscal policies.

The Ricardo-Barro argument that fiscal policy is ineffective (since it will simply be undone by actions in the private sector) rests on simplistic models. Government spending can be complementary to private spending (either to private consumption or investment) today, and thus affect changes in intertemporal allocations, just as changes in intertemporal prices brought on by monetary authorities can. Even government spending that is complementary to future private spending can elicit more private spending today, for example, because consumers rationally take the future impact on their budget constraints into their current spending (Neary-Stiglitz 1983).

So too, the reliance on using the short-term interest rate for macroeconomic adjustment may lead (even with full rationality) to distortions in intertemporal and risk prices (as we noted), and optimal macroeconomic adjustment may seek to optimize by minimizing the resulting distortions through the use of appropriately designed fiscal policies.

Choice of technique: Creating a jobless recovery

Here, we discuss one piece of evidence that reliance on changing intertemporal prices for equilibrating the economy may not be optimal. There are many alternative theories attempting to explain why the economy fails to attain full employment, including those related to wage and price rigidities (with those rigidities in fact being endogenous in some variants of these theories). Monetary policy attempts to correct for these distortions by controlling the interest rate (usually the short-term interest rate), setting it at a level different from what it otherwise would be. But intuitively, if the source of the distortion is in the labor or product market, it might make far more sense to attempt to correct at least some of the distortion more directly.

The standard argument for monetary policy is that it increases investment (and possibly consumption) leading to higher GDP and thus employment today. But there is another effect: lower interest rates induce firms to invest in more capital-intensive technologies, lowering future demand for labor. It affects the choice of technique. Even if real wages go down in a recession, the decline in the cost of capital is even larger. The original distortion is an excessively high price of labor relative to capital because of wage rigidities; the interest rate policy exacerbates that distortion. We see the consequences: firms replacing unskilled checkout clerks and tellers with machines.

Thus, as the economy recovers, there will be a lower demand for labor than there would otherwise have been – it will take a higher level of GDP to achieve a restoration of full employment.

The problem is that we are asking too much from a single instrument, and in principle, there are more instruments in the government’s tool kit. The government could, for instance, provide a larger investment tax credit for more labor-intensive technologies. But most governments have eschewed using this broader set of instruments. With a more constrained set of instruments, monetary policy may not only have these adverse distributional effects, but also may be less effective, as we shall explain shortly.

Distributive effects of monetary policy

The economist’s focus on aggregative models with a representative agent has shifted attention away from another important set of effects of monetary policy: their implications for the distribution of income. The presumptions have been the following: (1) The focus of monetary policy should be macroeconomic management, and if there are distributive effects, they are likely to be minor and correctable through fiscal/tax and transfer policies. (2) Ensuring that the economy is at full employment is the most important thing that government can do to ensure the well-being of workers. Higher employment helps workers directly and indirectly: Lower unemployment will lead to higher wages and higher GDP will lead to higher tax revenues and greater benefits for ordinary citizens. Recent failures of monetary policy have highlighted, however, that there can be significant adverse distributional effects, and the politics in the US and Europe have shown that the likelihood of any adverse effects on distribution being offset by government are
nil. More generally, research over the past three decades has shown that there are significant costs of such redistributions, and unless the growth benefits are significant, the distributional effects may thus outweigh them.\footnote{31}

Among the distributional effects, two stand out: the first, its role in creating a jobless recovery, was discussed in the previous subsection. The second arises from the fact that better-off individuals disproportionately hold equities, while worse-off individuals hold debt, including government bonds. Lowering the interest rate on government bonds to stimulate the economy hurts bondholders, at the expense of those who own equity, thus leading to more wealth inequality.\footnote{32} Indeed, elderly retirees who have acted prudently, in a risk-averse manner, holding onto government bonds, have been devastated by quantity easing.

The distributive effects may undermine the effectiveness of monetary policy. Since the marginal propensity to consume (out of income or wealth) of those at the top is much lower than at the bottom, any adverse distributive effect lowers aggregate demand.\footnote{33} Moreover, target savers (those saving for retirement, to obtain a down payment on a home, or to finance the education of their children) have to save more to meet their targets.\footnote{34} If the distributive effect is large, and the stimulus to investment is small (which has been the case since the 2008 crisis, with investment as a share of GDP actually lower in 2015 than it was in 2007 in spite of QE\footnote{35}), then the net effect on the economy of lowering interest rates or more accommodative monetary policies (QE) may have been negative.\footnote{36}

Open economy

The previous discussion focusing on a closed economy emphasized that the government, having delegated the allocation of credit to the private sector, with limited restrictions, had relatively little control over the use to which money/credit would be put, and therefore there was at best a loose connection between monetary policy and macroeconomic activity. Matters are even worse in an open economy for two reasons. Now, there is a further use to which the credit created can be put – purchasing assets abroad. This was evidenced in the aftermath of the 2008 crisis, where much of the liquidity created in the US went to purchase assets and make loans in emerging markets – not a surprise given the boom in these economies and the lackluster performance of the US. In short, the money went where it was not wanted and needed; and didn’t go where it was wanted and needed. The stimulus of/to the US economy from this loose monetary policy was thus limited.

There is a second effect: now there is an alternative supply of credit from lenders outside the country. Thus, even if monetary authorities tighten credit, there can be an offsetting effect from a flow of money into the country. Indeed, there has been a regular pattern of exactly this happening: when countries tighten credit, raising interest rates, out of concern about overheating, the higher interest rates attract an inflow of capital partially or fully offsetting the domestic contraction.\footnote{37} Only by controlling the sources and uses of funds carefully can some semblance of control over the macro-economy be achieved.

Summing up

Standard monetary theory has sought a neat set of instruments and targets by which the macro-economy could be well-regulated. There was perhaps as much a political drive for such parsimony as an intellectual one: if one could find a simple variable that could lead to macrostability, the nature of government intervention would be very limited; there would be little need for discretion. It was as if, when it was realized that Adam Smith’s beautiful economic machine didn’t work quite as perfectly as his latter-day descendants believed (though Smith himself was far more aware of these limitations), a slight modification to that machine would ensure its smooth running. Monetarism held that the government should simply expand the money supply at a fixed rate. New theories focused on controlling the interest rate (and indeed, some proposed a simple rule by which that might be done, reflecting inflation, a rule that would work regardless of the source of the disturbance to the economy giving rise to the inflation). Today, these theories are largely discredited (see the various papers in Blanchard et al.\footnote{2012}; Akerlof et al.\footnote{2014}). The discussion here has provided the underlying analytics explaining why we should not be surprised at the failures of these simple theories, and of the broader institutional theory attributed to Tinbergen of assigning to the central bank a single target – inflation – and a single instrument – the short-term interest rate.\footnote{38}

Government only controls the supply of credit very indirectly through the instruments under its control, and it does even a poorer job at controlling that part of credit that goes to purchase newly produced goods – say machine goods, buildings, or consumption – within the country. As a result, monetary policy is a weak instrument for controlling the economy in the sense that the link between the actions taken and the desired effects is uncertain. In certain circumstances, we have explained why it is an ineffectual instrument – it simply may not be able to restore the economy to full employment. Its ineffectiveness goes
well beyond the usual “zero lower bound” argument, a generalization of the Keynesian liquidity trap. Indeed, seemingly more accommodative monetary policies may, under certain circumstances, even have perversely contractionary effects, especially when they are not well designed to take into account likely effects on the banking system and broader distributive consequences. Some evidence of this has been seen in recent forays into negative interest rates.

Earlier IS-LM analysis was largely predicated on a stable demand curve for money. What was variable was the “real” economy; the interest rate at which full employment could be attained. Hence, with an unstable IS curve and a stable LM curve, monetary policy sought to increase $M$, the money supply, to the point where the rate of interest fell, to the level that induced full employment.

It became clear, however, that the LM curve itself was unstable, and this naturally led the government to target the interest rate. In effect, $M$ increased until the desired interest rate was achieved. It was assumed that lowering the (real) interest rate would lead to higher output. Hence, all that needed to be done was to lower the interest rate enough. But then, in the Great Recession, monetary authorities hit the ZLB. Clever economists responded that it was only the nominal interest rate that was constrained. If, somehow, we could raise inflationary expectations, credibly committing to a higher inflation rate, then the real interest rate would fall, and the faith that one could rely on monetary policy to restore the economy to full employment would itself be restored. Putting aside the fact that no one has figured out how to make such a credible commitment, we have argued that this framework is badly flawed.

We have explained that the links between the instruments under the control of monetary authorities and the variables that affect aggregate activity are weak, unstable and uncertain, and even of ambiguous sign. For instance, lowering the T-bill rate may or may not lead to a lowering of the lending rate. Because of distributive effects, we have explained that lowering interest rates may lower aggregate demand. Even when there is a positive elasticity of aggregate demand to changes in interest rates, the interest elasticity may be small, so that the magnitude of the changes in interest rate required will be very large – larger than would be politically acceptable, because such large changes will inevitably have large distributive consequences. We have explained too that changes in interest rates, especially large changes, increase uncertainty; and such increases in uncertainty themselves can have adverse aggregate effects. And this is even more so as monetary authorities stretch themselves, seeing that traditional instruments are failing; there is uncertainty associated with these innovative instruments, even if the policymakers were confident about the effects, if market participants are not, there are adverse effects on the real economy of these risk perceptions.

We have also explained why, though monetary policy has long been held out as the instrument of choice, it is an instrument with some adverse side effects, both on efficiency and distribution.

Indeed, it seems peculiar – and inadvisable – to attempt to correct a deficiency of aggregate demand arising from “shocks” to the economy arising, say, from an increase in uncertainty or an adverse change in the distribution of income or wealth (because of a deflationary shock which increases the real indebtedness of firms and households) by changing the interest rate (intertemporal prices). Even if one could do that, it seems preferable to address the underlying problem. If there is an increase in uncertainty, then government can take a more active role in risk mitigation, for example, by issuing income-or state-contingent loans. If there has been a redistribution of wealth as a result of a large deflationary shock, the government might consider a better system of debt restructuring (e.g. through a homeowners’ chapter 11; see Stiglitz 2010).

To reflect the central refrain of my criticism of the Washington Consensus, there are broader goals (not just price stability, but employment, growth, financial stability, and even distribution) and more instruments (including a broad set of regulatory tools called macroprudential instruments) than those seeking to employ monetary policy should use. Managing a complex economic system in the face of uncertainty requires as many tools as one can manage; the single-minded focus on short-term interest rates narrowly confined what central banks could do, just as the single-minded focus on inflation narrowly confined what central banks should do.

We have attempted to dethrone monetary policy from the pedestal on which it has been placed by some economists seeking to put it at the center of macroeconomic management. But at the same time we have shown how we can make monetary policy more effective than those who have focused on the narrow set of instruments that have traditionally been assigned to monetary authorities.

The simple empirical results in the Appendix show that on average the link between monetary policy and variables and the real economy is very weak – results that are consistent with the numerous schools of thought (such as real business cycle theory) that have argued that the real economy is affected by real variables, and that monetary variables have, at most, second order effects. The results of the Appendix are
especially powerful in discrediting standard formulations, for example, where there is a simple Keynesian demand curve for money that plays a central role in interest rate determination, and this itself is at the center of the transmission mechanism for monetary policy.

But even if monetary policy on average has little impact, that is not what monetary authorities care about: they want to know whether, under the particular circumstances being confronted at this particular moment, monetary policy can be used to stabilize the economy, to stimulate the economy when there is excess capacity, and to constrain aggregate demand when there are inflationary pressures. It should be clear that there have been particular moments in history when monetary policy has mattered. Hopefully, the analysis of this chapter helps us understand better when monetary policy matters, and when it does not.

In the next section of the chapter, we argue that there are more fundamental reforms to the monetary architecture of the economy — to the system of credit and transactions — making use of 21st-century technology, which will enable monetary authorities in the future to do a far better job of macroeconomic management.

Creating a new electronic financial system

The 2008 global financial crisis and the subsequent discussion of financial sector reforms highlighted the failures of financial markets and the enormous consequences of these failures for the economic system. These included the following: excessive volatility in credit creation, with a misallocation of capital and a mismanagement of risk; more credit going to the purchase of fixed assets rather than to the creation of productive assets; excessive and volatile cross-border flows of short-term capital, leading to volatility in exchange rates and trade flows; excessive charges for the running of the payments mechanisms; and an array of socially unproductive practices, from market manipulation to insider trading to predatory lending. Around the world, these financial market dysfunctions have had serious macroeconomic consequences. In the case of Europe, misguided credit flows to the periphery countries created the imbalances from which Europe is still suffering. In the case of the US, predatory lending, securitization (often based on fraudulent practices), and derivatives led to the deepest downturn since the Great Depression.

Modern technology provides the basis of a new and more efficient financial system, one that would simultaneously lead to better macroeconomic regulation of the economy. The following sections briefly describe the key elements of such a system — a low-cost "medium of exchange" for facilitating transactions and a system of credit creation focused on the real economy, managed in a way far more conducive to macroeconomic stability than the current system.

Creating a 21st-century financial transactions system

The banking and monetary system serves multiple purposes. One of them is as a medium of exchange. The world has several times made a change in the prevailing medium of exchange. Gold was once used as a medium of exchange; then, at least in the United States, there was a move to the bimetallic standard, where gold and silver were used. Finally, we moved to paper (or "fiat") money. For years, it has been recognized that it would be far more efficient to move to e-money, away from currency. Our payments mechanism has already changed dramatically. We have gone a long way towards an electronic payments mechanism, and in most of the world we could go far further with an even more efficient one, if it were taken out of the hands of the monopolistic financial system. Electronic transfers are extraordinarily cheap, but banks and credit card companies charge exorbitantly for the service, reaping monopoly profits as a result. Electronic money is more convenient for people on both sides of the transaction, which is why it has become the dominant form of payment. It saves the costs of printing money, which has increased as the sophistication of counterfeiters has increased. It has a further advantage, especially in countries where small businesses predominate; it significantly curtails the extent of tax avoidance.

With electronic money, the money inside a country's banking system can, in effect, be easily "locked in" simply by not allowing the transfer of money out of the country's banking system. But anybody can transfer the money in his bank account to that of anyone else. Thus, everybody has, in effect, almost full use of his money. Money inside the country's banking system (which for convenience, we will call the G-euro) would be just like any other currency, with a well-defined value relative to any other currency.

Most individuals today have accounts; only the very poor are "unbanked", and in recent years governments and nongovernmental organizations, like the Bill & Melinda Gates Foundation, have been making great efforts to bank the unbanked. In most countries, government pension payments are now transferred through bank accounts, partly to reduce the risk of stolen checks, and partly to reduce the outrageous charges that are sometimes charged by check-cashing services.
Thus, the task of implementing an electronic banking system today is clearly manageable.

**Credit: Creating a banking system that serves society**

A big advantage of the use of fiat money was that one could regulate the supply. When gold was used as the medium of exchange, when there was a large discovery of gold — or when the gold supply increased as Spain conquered the New World — there would be inflation, as the price of gold would rise relative to other goods; if there were few gold discoveries, then there would be deflation. Both caused problems. Deflation, for instance, would redistribute income from debtors to creditors, increasing inequality and imposing hardship. America’s election of 1896 was fought on the issue of the money supply. The debtors wanted to increase the money supply by moving from gold to a bimetallic standard of gold and silver.

While the modern financial system based on fiat money doesn’t suffer from the vagaries of gold discoveries, it has sometimes suffered from something else: volatility in the creation of money and credit by the banking system, giving rise to the booms and busts that have characterized the capitalist system.

 Banks effectively increase the supply of money by increasing the supply of credit. In a modern economy, central banks regulate, usually indirectly, banks’ creation of money and credit. They are supposed to do it in *just the right amount*, so that there is a “Goldilocks” economy, neither under- nor over-heated but “just right”. It is apparent that they have often failed to do so. This is partly the result of the often-noted “long and variable lags” associated with monetary policy – with monetary authorities having to base their actions on predictions concerning the future course of the economy. But more importantly, for our purpose, is “instrument uncertainty”, the weak link between what monetary authorities do and the impacts on GDP, since the increased liquidity may go for many uses other than stimulating the economy, as our earlier discussion emphasized.

The traditional view of banking was based on a primitive agriculture economy. Farmers with excess seed — with harvests greater than they wanted to consume or plant the next season – could bring the seed to the bank, which would lend, at interest, the seed to some farmer who wanted more seed than he had, either for consumption (say, because he had a bad harvest that year) or planting. The bank had to have seed deposits in order to lend. Markets on their own equilibrated the demand and supply of seeds, so there was really little need for government intervention.

But if, for some reason, there was, the interest rate provided the natural mechanism: if for some reason, savings (at full employment) – the supply of seeds – exceeded investment (the demand for seed), then lowering interest rates would cause the supply of seeds to fall and the demand for seeds to increase until the two were equilibrated.

But this reasoning again totally misses the nature of credit in the 21st century. In a modern economy, banks effectively create credit out of thin air, backed by general confidence in government, including its ability and willingness to bail out the banks, which is based in part on its power to tax and borrow.

**Targeted regulation of credit creation**

There is a problem in our current system: because the central banks’ control mechanisms are typically very indirect, the economy is often over- or under-heated. Sometimes there is too much credit creation, leading to an excess of aggregate demand, and prices rise: there is inflation. Sometimes there is a lack of demand, and prices fall: there is deflation.

The first part of this chapter has explained some of the key reasons for this failure; while central banks can regulate the supply of credit reasonably well, they can’t (or more accurately don’t) regulate the use to which the credit is put. Much of the credit goes to buying preexisting assets, like land. Some of the credit goes to providing margin for bets (e.g. in futures markets). What determines whether the economy is over- or under-heated is the purchase of new goods and services (whether for consumption or investment). Thus, after the 2008 crisis, there was a massive increase in liquidity, as the Fed pumped money into the economy. But relatively little of this went to buy goods and services in the US, so in spite of the huge expansion of the money supply as conventionally measured, the economy remained weak.

In short, even with fiat money, there may still be a deficiency of domestic aggregate demand – a deficiency that could be easily corrected: there are individuals and firms who would like to spend but cannot get access to credit. A near zero interest rate does not mean businesses can get access to credit at such a rate – or at any rate.

**Restoring domestic control over credit creation**

The electronic payments mechanism allows a country to assert control over the supply of credit and the uses to which it can be put in a way which is far better than the current system. Think of this most directly
as occurring through a government bank. It can add “money” to the payments mechanism by lending money to a small enterprise with a proven reputation that wants to make an investment. The government simply puts more “money” into the bank account of the enterprise, which the enterprise can then use to pay contractors. Of course, in providing credit there is always a risk of non-repayment, and standards must be established for evaluating the likelihood of repayment.

In recent decades, faith in government's ability to make such evaluations has diminished, and confidence has been placed in the private financial system. The 2008 crisis as well as other frequent crises that have marked the last third of a century have shown that that confidence has been misplaced. Not only didn't the banks make good judgments - as evidenced by the massive, repeated bailouts - but they systematically failed to fulfill what they should have seen as their major responsibility: providing credit to businesses to create new jobs. By some accounts, their "real" lending amounts to just 3 percent of their activities, by others, to some 15 percent. But by any account, bank finance has been absorbed in other directions.

There were always obvious problems in delegating the power of credit creation, backed by government, to private institutions: banks could use their power to benefit their owners, through connected lending. Regulations circumscribed this, motivated by the experience of bad lending, perhaps more than by the implicit corruption and inequality to which such lending gives rise.

Circumscribing connected lending didn't address one of the key underlying problems: credit is scarce; giving private banks the right to create credit with government backing gave them enormous "economic rents". Even with connected lending circumscribed, bankers use their economic power to enrich themselves and their friends. Russia provides the quintessential example: those with banking licenses could use that power to buy enormously valuable state assets, especially in natural resources. It was through the banking system that the Russian oligarchs were largely created. In Western countries, matters are done more subtly - but the net result in creating enormous inequality remains (though not of the magnitude of Russia). In many cases, the banks lend money to those whom they "trust" and judge creditworthy, with collateral that they value; in short, the bankers lend money to those who are similar to themselves. Even if banker A can't lend to herself or his relatives, banker A can lend to the relatives of banker B, and banker B can lend to the relatives of banker A. The fallibility of their judgments has been repeatedly demonstrated: overlending to fiber optics at one moment, to fracking at another, to housing in a third.

There is a second danger to the delegation of the power of credit creation to private banks. Throughout history, moneylenders have had a bad reputation because of the ruthlessness with which they exploit the poor - especially at moments of extreme need, where without money their family members might die. At such times, there is an enormous asymmetry in bargaining power, which the moneylenders exploit. Virtually every religion has tried to proscribe such exploitation, prohibiting usury, and in some cases, even interest. Somehow, in the magic of Neo-liberalism, this long history was forgotten: bankers not only didn't suffer from the stigma of being called moneylenders, they were elevated to being the paragons of capitalism. In the enthusiasm over their new virtues, as linchpins in the workings of the capitalist system itself, it was simply assumed that such exploitation would not occur, perhaps in the belief that competition would ensure it couldn't happen, perhaps in the belief that with the new prosperity of workers, ordinary citizens wouldn't let it happen.

All of this was wishful thinking. Freed of constraints, 21st-century moneylenders have shown themselves every bit as ruthless as the moneylenders of the past; in fact, they are in some ways worse, because they have discovered new ways of exploiting both the poor and investors. The financial sector has enriched itself on the back of the government's credibility, without performing the societal functions that banks were supposed to perform. In doing so, the financial sector has become one of the major sources of the increased inequality around the world.

Even given this history, the government may want to delegate responsibility for making credit decisions to private enterprises, but if so, it should develop strong systems of incentives and accountability, such that the financial system actually focuses on lending for job and enterprise creation and so that it does not make excessive profits as it performs these functions and so the government should be adequately compensated for its backing. In effect, in the current system, all the "value" of the underlying government credit guarantee is captured by the private sector.

Credit auctions

Here, we consider one possibility for addressing this issue and providing for greater economic stability. First, the central bank (government) auctions off the rights to issue new credit. The amounts would be added to the "money" that is within the financial system. The magnitude of net credit that it allows to be added each month will be
determined by the country's central bank on the basis of its assessment of the macroeconomic situation — that is, if the economy is weaker, it will provide more credit to stimulate the economy. The winners of the credit auction then allocate this "money" to borrowers on the basis of their judgments about repayment capacity, within the constraints that the central bank may impose (described later).57

Note that in this system, banks cannot create credit out of thin air, and the amount of money being created each month is known with considerable precision. The winners of the credit auction can only transfer money from their accounts to the borrowers' accounts.

Conditions would attach to selling the "rights to lend" to the banks. Minimum percentages of the loans would go to small and medium-size enterprises and to new enterprises or to underserved communities; a maximum would go to real estate lending (perhaps apportioned by location, based on local changes in prices), to purchases of other existing assets, or to those engaged in speculative activities, like hedge funds. None would be allocated to socially proscribed activities, like those contributing to global warming or associated with the promotion of death, such as cigarettes. In short, there would be minimum standards for social responsibility. There would be limits on the interest rates charged. Discriminatory lending practices and other abusive practices by credit card companies would be proscribed. So, too, would connected lending. There would be further restrictions to ensure that the loan portfolio of the bank is safe and sound, and there would be strict supervision by government regulators to ensure compliance with the regulations governing any such program.

If it wished, the monetary authority could target credit even more narrowly, to be used to purchase goods which are in excess supply, or which use labor of types which confront high levels of unemployment. There is always a trade-off: such targeted lending may be subject to political pressure in ways that more broad-based measures may not be.

In a 21st-century banking system, a bank's ability to lend is, in a sense, given only temporarily. It is conditional on compliance with the rules and standards established. The government would allow for entry into the banking system; indeed, separating the depository and lending functions and the open auction of rights to issue credit should make entry easier, and thus competition more vigorous than under current arrangements.

Still, since lending is an information-based activity, and the gathering of information is a fixed cost, one would like stability in the new banking system, and this will require that banks not live on the edge — that is, that they be sufficiently well capitalized and sufficiently profitable.

"Sufficiently profitable" should not be taken to mean the 25 percent return on equity that one of the European banks, Deutsche Bank, famously came to expect as normal. Hence, entry of enterprises with sufficient capital and who also satisfy other conditions that enhance the view that they would be responsible lenders, would be encouraged.58 The system of auctioning of credit would ensure that banks do not earn excessive returns; most of the value of the public's backing to the creation of money/credit would be captured by the public, rather than by the bankers. At the same time, the new system of credit creation ensures that the social functions of finance are more likely fulfilled, at least better than under current arrangements.

This is an example of how to create a 21st-century banking system, responding to the advantages of electronic technology, doing things that would have been far harder to accomplish in earlier decades — a banking system more likely to ensure responsible lending and macroeconomic stability than the current system, and without the huge rents and exploitation that have contributed so much to the inequality that has stalked advanced countries around the world.

But this reform is about more than curbing bankers' exploitation. It is also about enhancing macroeconomic stability. One of the major contributors to macroeconomic instability is the instability in credit supply, and, in particular, to the supply of credit for the purchase of produced goods and services. The 2008 crisis demonstrated that all the advances in markets and our understanding of markets have not led to greater stability in this crucial variable — in fact, quite the opposite. The electronic banking system described here not only enhances stability in this critical variable, it also provides the basis of a virtuous circle leading to an increase in overall stability of the economy. One of the most important reasons that small businesses don't repay loans is macroeconomic fluctuation: loans simply can't be repaid when an economy is in depression. Ensuring greater macro-stability (than under the current regime) would do more than anything else to ensure the viability of the banking system and to encourage a more competitive economy.

Whence bank capital?

The beauty of the modern credit system is that it doesn't really require the same kind of capital as required by banking systems of the past. Recapitalizing a destroyed banking system would not require gold or borrowing to buy seeds as it did in the old days. As we have seen, the government itself can simply create credit.59
The fact that the money created by the government can be used to pay the taxes that are owed to the government, and that the government has the power to levy taxes, ensures the value of the credit it has created. Indeed, because the credit that has been created is electronic money, the movement of which can easily be monitored, the government does not only have the ability to levy taxes—it also enhances the ability to collect taxes.

The only reason for bank capital in this world is as a partial guarantee that the bank has the capacity to repay the credit—the bank’s “purchases” from the government of the right to issue credit are only temporary, and the credit thus created must be repaid to the government. (The fact that the bank will lose its own capital has, in addition, strong incentive effects, incentivizing the bank to make good decisions about whom to give the credit to, and to monitor the loan well.) But if the government is doing an adequate job of bank supervision and has imposed appropriate regulations (e.g., on connected and excessively risky lending), the amount of capital required will be limited. And that fact alone should lead to more competition in the market for the provision of credit—reducing the excessive returns currently received.

**Macro-stability and income-(state-)contingent loans**

To achieve full employment may entail an auction of credit at which the price is negative, that is the only terms at which potential lenders are willing to “accept” the temporary use of funds, to be repaid later, entail a negative interest rate. The auction may entail a provision (unlike the current system) where a negative “bank rate” has to be passed on (at least partially) to borrowers, in the form of a negative lending rate. Presumably there is some negative rate at which the desired credit creation—that viewed as necessary to ensure full employment—related to new spending (investment or consumption) is achieved. But it may be a very negative rate, and the distributive and even allocative consequences of that negative rate may be adverse. Accordingly, it makes sense to look for more effective ways of stimulating the economy. One such way—ensuring a trade surplus—is discussed in the next section.

Here we consider another way: state-contingent loans, whereby the amount the borrower has to repay depends on the state of the economy.

There is a widespread consensus that one of the reasons that consumption and investment are depressed in a deep economic downturn is “lack of confidence”, or, slightly more precisely, uncertainty about the future. Consumers are not sure of their future wages, retirees of the future return on their savings, and producers of the returns on their investment. They worry that if the downturn persists, unemployment may be high, wages low, interest rates low, and sales poor. Traditional monetary policy has tried to compensate for the absence of insurance markets by which individuals might mitigate these risks by changing the intertemporal price. It is, to say the least, a peculiar response: it makes far more sense to try to address the market failure directly, than to increase one distortion in response to another. As we noted in the first part of the chapter, it is not even clear that lowering the interest rate is even an effective response, not just because of the distributive and distortionary effects, but because as the interest rate is lowered, risk perceptions may increase, and the adverse risk effects could overwhelm the intertemporal price effects.

**Managing the current account deficit**

The analysis so far has been for a closed economy. Extending the analysis to an open economy is at least conceptually easy. When a firm exports some good, say a widget, it receives dollars. The dollars could be kept out of the country, say in a dollar account in New York. But the exporter may want to bring the dollars into the country, depositing them into the country’s electronic banking system. The number of, say, G-euros that the exporter would receive in return for the dollars would be market determined; that is, importers may want dollars to buy goods from the US. They thus transfer money in their bank account to the exporter. By the same token, an individual in the country wanting to make an investment abroad, say in the US, might want dollars, and be willing to transfer G-euros in his electronic banking system to someone who is willing to sell him dollars.

These capital flows may, however, be very destabilizing—leading to large fluctuations in exports and imports as the exchange rate changes, leading in turn to macroeconomic instability. The central bank can attempt to offset these effects through the system of credit creation (auctions) described earlier. However, there is another way of regulating trade flows that may be more effective.

**Managing the current account deficit through trade chits**

In this proposal, the government would provide to any exporter a chit, a “token” (in this case, electronically recorded—alternatively called trade chits or Buffett chits), the number in proportion to the value of what was exported. To import a G-euro’s worth of goods, there would
be a requirement to pay, in addition, a G-euro's worth of chits or “trade tokens”. There would be a free market in chits, so the demand and supply of chits would be equal; and by equating the demand and supply of chits, one would automatically balance the current account.

In practice, the value of the chit might normally be very small, at least for a country with a small trade deficit.

This system would be a way of managing the high level of volatility in market economies associated with short-term capital flows. With the free flow of capital, the exchange rate is determined by the vagaries of the market. And those capricious changes in exchange rate then drive exports, imports, the trade deficit, and borrowing, and in doing so, give rise to macroeconomic instability. With the system of trading chits, the trade deficit can be controlled, enhancing overall stability.62

In the previous analysis, where every import needs a chit, there is neither a trade surplus nor a trade balance. The government could use this system to limit the size of the deficit or surplus as well. For instance, if it wants to limit imports to be no more than 20 percent greater than exports, it can issue 1.2 import chits for every G-euro of exports. When there would be an excessive surplus, every import would be granted an “export” chit. Then every export would require a chit. This would automatically bring exports down to the level of imports. By issuing both import and export chits, the trade balance can be kept within any prespecified bounds.

The fact that the country could thus stabilize the size of the trade deficit or surplus has an enormous macroeconomic advantage: it facilitates macroeconomic stabilization itself. It means, for instance, that a small country doesn't have to suffer from the vagaries of its “external balance”, its net export position. These fluctuations impose enormous costs on society, of which the market, in generating them, takes no account.63

But ensuring stability in the trade deficit also engenders longer-term stability, for national indebtedness, built up over many years, can suddenly become unsustainable. The market sees the world through very myopic lenses. It is willing to lend year after year – until it suddenly changes its mind.64 By limiting the trade deficit, a country is in effect limiting national borrowing; this framework thus reduces a key source of instability.65

Moreover, we can see how this system would help strengthen the G-euro. In the absence of the chit system, an increase in the demand by Greeks for imports (i.e. for, say, dollars to buy American cars) would lead to a fall in the price of the Greek-euro. But now, with imports discouraged by the necessity of also paying to purchase a chit, the increased demand for imports would be reflected in an increased price of a chit, rather than a decrease in the value of the Greek-euro. The Greek-euro will be stronger than it otherwise would be.

Economic theory and macro-stability

Some might complain: aren't we interfering with the market? Of course, all monetary policy represents an interference with the market; few believe that interest-rate determination should simply be left to the market.

This proposal entails minimal intervention in the market, and even in doing so, uses market mechanisms. It corrects for a well-recognized externality, the market externality associated with external imbalances. Markets exhibit enormous volatility in both prices and quantities: interest rates demanded of borrowers from different countries have moved violently in different directions, and capital and credit flows have fluctuated in ways that are virtually uncontrollable under current arrangements.

Workers are told that they should simply accept being buffeted by these maelstroms that are not acts of nature but the creations of irrational and inefficient markets. Workers should accept wage cuts and the undercutting of social protections in order for the capital markets to enjoy their “freedom.” The electronic payment system, with credit auctions and trade chits, is intended to bring a modicum of order to this chaos, which has not even produced the higher growth in GDP that was promised – let alone the social benefits that were supposed to accompany this higher GDP.

In the Arrow Debreu world with perfect markets, prices play a critical role in ensuring economic efficiency. But in the real world in which we live, as Martin Weitzman (1974) explained long ago, it is often better not to just rely on prices, but to try, as our proposed framework does, to control the quantity of credit and net exports, and to regulate the uses to which credit is put. There is a large literature showing that under a variety of conditions, when there is a departure from the first-best world, such quantity controls are a better way of regulating the economy.66 However, the management of the economy in our proposed framework relies heavily on the use of prices, but not fully so; there is no micromanagement, but more macro-management than exists today.

Conclusions

Decades ago, we learned that one could not let a market economy manage itself. That is why, for instance, every country has a central bank determining interest rates and regulatory authorities overseeing
banking. Some would like to roll back the clock to a world without central banks and with free banking, with no restraints. Anyone who has read his economic history knows what a disaster that would likely be.

But anyone observing macroeconomic performance in recent years will see that things have not gone well in many countries around the world - even in advanced countries, in Europe and the US, with supposedly well-functioning markets and institutions and well-educated individuals to manage the economy. The framework provided here provides a way of improving matters. These are modest reforms that would not upend the system. But they systematically address some of the major weaknesses of current economic arrangements, some of the major instabilities that have proven so costly to our economies and our societies.

There are, of course, many details to be worked out. The system is surely not perfect. It is not intended to eliminate all speculative activity, and it will not do so. But by restraining the uses of newly issued credit it will curb such activities. But almost as surely, it is better than the current system. This framework could lead to greater economic stability and growth.

Appendix

On the relationship between money, output, and interest rates

In this appendix, we present some new evidence which supports our reformulation of the theory of money and credit and macroeconomic activity. The evidence is only suggestive, but we believe that, at the same time, it persuasively argues against existing formulations, even if it provides only limited support for those advanced here. Each of the models we present here were widely believed among monetary economists, some based on well-articulated theories, others advanced simply as empirical regularities. The data we present for recent years undermines the empirical foundations for these ideas.

Quantity theory of money

We begin with the (now widely discredited) quantity theory of money, which underlay monetarism:

\[(A1) \quad MV = PQ\]

where it was postulated that \(V\), the velocity of circulation, was a constant. In the quantity theory of money, government controls the money supply, so \((A1)\) is an equation determining national income, \(Y\), where

\[(A2) \quad Y = PQ,\]

that is,

\[(A3) \quad Y = VM,\]

an hypothesis which can be directly tested, in the form of \((A3)\), or in the form of

\[(A3') \quad \ln Y = \ln M + \ln V\]
or

\[(A3')\Delta Y = \Delta M\]

or

\[(A3'')\Delta \ln Y = \Delta \ln M.\]

The quantity theory of money implies that the coefficient in the above regression should be unity. If we run the regression in this form, over the period 1959-2016, using quarterly data, for instance, the coefficient is significantly different from 1.\(^{69}\)

We can test the model in other ways, for instance, by running the regression

\[(A4)\Delta Y = \alpha + \beta \Delta M.\]

The hypothesis \((A3'')\) predicts that \(\beta = 1\) and \(\alpha = 0\). We can reject both hypotheses, and the fit of the regression is remarkably poor.\(^{69}\)

We have run the tests against all the specifications, with flexible lags, using the standard definition of money, M2, and with alternative time periods. We have also run the regressions in logarithmic form, a more natural specification in order to remove scale effects:

\[(A4')\Delta \ln Y = \alpha + \beta \Delta \ln M\]

The fit worsens, but the conclusion is unaltered.\(^{70}\)

Often, it makes more sense to express everything in real terms:

\[(A5)\Delta Q = \alpha + \beta \Delta M/P\]

or

\[(A5')\Delta \ln Q = \alpha + \beta \Delta \ln M/P\]

Again, if the quantity theory were true, \(\alpha\) should be zero and \(\beta\) unity. Two striking things stand out from our regressions, done over the period 1959 to 2016\(^{71}\) (and various subsets of that period) with US data (preliminary work suggests similar results for other countries): (1) \(\beta\) is always small, and sometimes even insignificant; and (2) the relationship between a change in Q (or Y) and changes in M/P (or M) has weakened, with a structural break sometime around 1990.\(^{72}\) That there should be some structural break is hardly a surprise, for there were marked changes in both the financial sector (both in regulations and the creation of new instruments, which serve as effect substitutes to money, the money market funds) and in the conduct of monetary policy (the switch from a focus on the money supply itself to the interest rate, with a greater emphasis on inflation).

It is possible, of course, that our measure of money is “wrong”. Moreover, M2 clearly has some element of endogeneity, with banks lending more, creating more money, when income is going up (or is expected to be going up.) Accordingly, we reran our regressions with another notion, base money – the balance sheet of the Fed – which is seemingly a variable more directly under control of monetary authorities than M2. Let B be the size of the Fed’s balance sheet (which we take as a crude measure of “base money”). Then the true money supply (here assumed unobservable) is assumed to be a function of B (observable). We again run these regressions, replacing B (or \(\Delta B\)) for M (or \(\Delta M\)). The results are similar: either no significant effect, or a small effect.\(^{73},^{74}\)

**The demand for money**

These are all models suggesting that some-how money “drives” the economy. But there is an alternative way of writing all of these equations, more modestly, as simply a demand curve for money. The quantity theory can be thought of as being an equilibrium model based on a special case of the Keynesian demand equation for money

\[(A7) M^d = k(r) PQ,\]

where the demand for money is made proportional to income, with the proportionality factor depending on the interest rate.

**Constant velocity**

If \(k' = 0\), then \(V = 1/k\). If we also assume that we are always (nearly) on the demand curve of money (after all, no one forces people to hold money that they don’t want to hold), then

\[(A8) M = kY,\]

with the correlates \(\Delta M = k\Delta Y\), etc. As we shall comment further, testing \((A8)\) simply tests whether the economy is “on” a demand function
of the given form, a seemingly much weaker hypothesis than that M "drives" the economy. We can test it by running the regression

\[(A8') \Delta M = \alpha' + \beta' \Delta Y,\]

or, perhaps better (eliminating scale effects)

\[(A8'') \Delta \ln M = \alpha' + \beta' \Delta \ln Y,\]

with the quantity theory hypothesis being \(\beta' = 1\) and \(\alpha' = 0\).

Our empirical analysis provides convincing evidence rejecting the hypothesis of the quantity theory of money and the related demand theory for money. There are several alternative explanations. The analysis of this chapter has provided one explanation: money is not needed for transactions, and most transactions are not directly related to income \(Y\) (and there is no reason that they would grow in proportion to \(Y\)). There are some difficult econometric problems, which might lead to coefficient bias but are hardly likely to provide a convincing explanation of the results. The most obvious explanation is that \(V\) is not constant.

**Keynesian demand for money**

Keynesian monetary theory hypothesized a stable demand function for money, where a linearized version of \((A7)\) gives a corresponding set of equations, such as

\[(A9) \Delta \ln M = \alpha' + \beta' \Delta \ln Y + \gamma' \Delta r\]

or

\[(A10) \Delta \ln M/P = \alpha' + \beta' \Delta \ln Q + \gamma' \Delta r\]

Testing \((A9)\) is testing whether the economy is on a Keynesian demand function. There are, of course, important econometric problems in testing any of these relations, since all of the variables are, in some sense, endogenous. This is true even when government sets \(M\) and/or \(r\) because it sets these variables based on expectations of \(Y\), or what \(Y\) would be in the absence of intervention, and those expectations may well be correlated with \(Y\) or \(Q\) itself. We shall return to these problems later.

Here, we simply note that in running the regression, \(\alpha'\) should be zero (a value of \(\alpha'\) not equal to zero would suggest that there are economies or diseconomies of scale in the use of money. While inventory theories of money might suggest that there are economies of scale, the presumption has been that these are sufficiently small that \(\alpha'\) would not deviate significantly from zero), \(\beta'\) should be unity, and \(\gamma'\) should be negative (here should be the opportunity cost of holding money). The first two hypotheses can clearly be rejected, and the overall fit is sufficiently poor that it suggests we are not on the Keynesian demand function for money. \(\gamma'\), while of the right sign, is very small. These results should create skepticism about the edifice created on the foundations of Keynesian monetary theory.

**Keynesian equilibrium**

The standard Keynesian model assumes a stable demand curve for money, while the demand for investment and consumption is somewhat volatile. Investment is a function of the (real) rate of interest. The IS-LM curves give the resulting equilibrium, assuming that inflationary expectations are fixed, so that a change in \(r\) translates into a change in \(R\), the real interest rate. The same results hold if, as in more recent models, one assumes consumption is a function of the interest rate. Thus, we write

\[(A11a) \, r = \varphi (Y/M),\]

from the LM curve, and insert this into the stochastic IS curve

\[(A11b) \, Y = \phi (r) = \phi(\varphi(Y/M)) + \varepsilon,\]

which we rewrite in linearized form

\[(A11b) \, Y = \alpha' + \beta' M + \varepsilon.\]

or

\[(A11c) \, \Delta \ln Y = \alpha' + \beta' \Delta \ln M + \varepsilon.\]

In the standard Keynesian model, price is fixed (in the short run), so the more interesting variant predicts that an increase in the money supply increases real income:

\[(A11d) \, \Delta \ln Q = \alpha' + \beta' \Delta \ln M + \varepsilon.\]

\((A11c)\) and \((A11d)\) are, of course, just the "reverse" equation of \((A8)\), and the same as \((A4)\) and \((A5)\), respectively, except now our
hypothesis is only that $\beta' > 0$: An increase in the money supply leads to a lowering of the interest rate, and this should increase nominal and real income. But our naïve regressions show that an increase in money lowers real income ($\beta'$ is negative, though not significantly different from zero), though it increases nominal income – with a significant coefficient. It would seem, on the basis of this crude analysis, that if monetary expansion affects national income, it is more by affecting the price level, contradicting the underlying assumption of the Keynesian model that prices are fixed. There are, of course, important simultaneity problems; in particular, monetary policy is conditional on many variables that could themselves be correlated with $\Delta Y$. We will say a little more about this later.

**Government controls $r$**

As the instability of the LM curve became clear, as we noted in the text, it made sense for monetary authorities to switch to a focus on the variable that they (wrongly) believed to be the key determinant of macroeconomic activity, the interest rate. They could control at least the nominal interest rate directly, and if inflationary expectations were fixed, they could thus control the real interest rate. If inflationary expectations were variable and highly unpredictable, of course, controlling the nominal interest rate would leave much of the real interest rate out of the control of monetary authorities. Since different economic actors may respond to different interest rates and associated variables affecting the cost of capital, it is not just the T-bill rate that matters. Indeed, the thrust of the first part of the chapter was to argue that the most important variable affecting investment (at least by SMEs) is the lending rate, and that this variable is only loosely connected with the T-bill rate. Tobin argued that it was the price of equity, and this may even be more loosely related to the T-bill rate. Some investments may be more related to the long rate than the short, and the maturity structure itself is an endogenous variable – at least it was before monetary authorities sought to affect it through QE.

We ignore these complexities, postulating that

$$(A11d) \quad \Delta Q = \alpha + \gamma \Delta r,$$

or in logarithmic form

$$(A11e) \quad \Delta \ln Q = \alpha + \gamma \Delta r,$$

where it is assumed that the government controls $r$ (possibly through the control of $M$). We run reduced form regressions of the form $(A11d)$ and $(A11e)$, obtaining a coefficient on $\Delta r$ of the wrong sign, with again there being a structural break in the late 80s/early 90s, perhaps corresponding to the adoption of the interest rate as the focal point of monetary policy. The logarithmic form does slightly better, but in both, the $R^2$ is very low. There are many things driving the economy. Changes in interest rates do not appear to be among the more important.

In the formulations so far, $r$ enters through the money demand equation, representing the opportunity cost of holding funds. But standard Neo-classical theory suggests that investment and consumption are affected not by the nominal interest rate, but by the real interest rate, $R$. Changes in nominal interest rates translate into changes in real interest rates if inflationary expectations are constant. But inflationary expectations may well change as $r$ or $Q$ changes. Most simply, assume $\Delta R = f(\Delta Q, \Delta r)$ and $\Delta Q = H(\Delta R)$. We thus obtain in reduced linearized form $(A11d)$ once again. Only the interpretation of the coefficients has changed.

More agnostically, we regress changes in output and the log of output against both changes in nominal and real interest rates, using the interest rate derived from treasury inflation-indexed security, constant maturity, as a proxy for the real interest rate

$$(A12a) \quad \Delta Q = \alpha + \gamma \Delta r + d \Delta R,$$

$$(A12b) \quad \Delta \ln Q = \alpha + \gamma \Delta r + d \Delta R.$$

Several things stand out from these regressions, done over the shorter period 2003–2016 (largely because of data availability) using quarterly data. The $R^2$ is much higher than in any of the other equations. Most significantly, only the coefficient on the nominal interest rate is significant. By the same token, if we run the regression on only nominal or real interest rate, nominal does better than real – the real is not significantly different from zero. We have run regressions with variable lags, and the results remain unchanged.

To be sure, we have not estimated a sophisticated structural model, nor have we massaged the data or engaged in any of the usual data mining. (There are, accordingly, reasons that our estimates may be downward biased.) All that the analysis says – and it may be saying a lot – is that a quick and dirty look at the data doesn't provide the kind of support for many of the maintained hypotheses in conventional
monetary theory and macroeconomics. In particular, most striking are the results on money itself – for the standard hypothesis is that the money demand equation is relatively stable; that the money demand equation is of the general postulated form; and equilibrium analysis requires that the economy be “on” the money demand equation. The data rejects those hypotheses.

An alternative interpretation?

There are other, more complicated interpretations of the results on (A12) (or the earlier simpler versions). If monetary authorities were perfect in predicting what output would be in the absence of changes in monetary policy, and thus changed monetary policy perfectly, to ensure that the economy was always at full employment, there would be no correlation between real output and monetary policy variables. Of course, we know that monetary authorities have been far from perfect in predicting what output would have been in the absence of their intervention, and far from perfect in designing intervention to ensure that output is stable.

Assume that monetary authorities raise interest rates when they expect output to increase next period above trend, and dampen output from what it would have been but not relative to the long-run trend. Then periods of high interest rates would be periods of high growth. Monetary policy may have worked – but the regressions don’t show it.

Fortunately, we have data that expresses the forecasts of the Fed itself and of the general market consensus. Unfortunately, such forecasts embed both forecasts of the underlying disturbances to the economy and the Fed response, and the fact that such forecasts show systematic deviations from full employment implies that forecasters do not believe that the Fed will fully correct any “shock” to the underlying system. On average, one does not believe that Fed policy deliberately moves the economy away from full employment. Thus, it is reasonable to assume that when forecasters see a larger deviation from full employment, they see an adverse shock that will be partially, but not fully, offset by the Fed (consistent with risk aversion combined with instrument uncertainty; see Greenwald and Stiglitz 1989).

We have data on market forecasts of both interest rates (predictions of Fed policy) and output, and thus can ascertain whether there is any systematic effect on outcomes relatives to forecasts of the Fed’s doing something different from what was expected. There does not appear to be. Lest we view what forecasters say about their forecasted interest rate not as reflecting their “true” forecasts, we modeled a forecast of the Fed’s policy based on their expectations of output (or other variables), and ascertained whether deviations of policies from these predicted policies had an effect on output. It did not seem to.

The theoretical analysis of this chapter has provided a set of explanations for why none of these results should come as a surprise. They are not “econometric artifacts”, likely to disappear if we massage the data enough, but rather should be taken to be stylized facts which economic theories need to take account of.

Nor should the fact that our results show no systematic relationship between money and output be taken to mean that money doesn’t matter. Our models explain why it matters, both in normal and abnormal times. In abnormal situations – such as Volker’s sudden change in US monetary policy – the credit constraints and soaring interest rates mattered hugely – they threw the US economy into a recession. Our analysis explains why one shouldn’t look just to the interest rate – and especially to the T-bill rate – to assess the consequences. In more normal times, our analysis explains why other variables, like changes in expectations and perceptions of risk are likely to be as or more important; and again, the effects of monetary policy may be felt more through credit availability than through small changes in interest rates; and changes in credit availability may be far from perfectly correlated with changes in money supply. Thus, while this chapter takes a somewhat nihilistic stand on some of the monetary econometrics, it is quite positive about the relevance of monetary policy – though it argues that its effects cannot be well-captured in a single variable, like M2 or “r”.

Notes

1 All data in this chapter was obtained from the Federal Reserve Economic Data base (FRED), available at https://fred.stlouisfed.org/.
2 This section represents a development of ideas earlier presented in Greenwald and Stiglitz (2003).
3 In the 2008 financial crisis this relationship broke down temporarily. Apart from that, there appears to be no significant cyclical movement in the difference between the T-bill rate and the money market rate.
4 See Robertson (1934) and Ohlin (1937).
5 That is, the spread between the two is endogenous, and can vary with economic conditions and policy.
6 More broadly, with imperfect information, behavior is constrained by collateral, self-selection and incentive compatibility constraints.
7 Effects which may arise from the change in policy (interest rates) itself – some of which we describe in greater detail later – or which may arise simultaneously from other sources.
8 As we have already noted, as influential as Keynes’s work has been, it provides a poor description of a modern credit-based economy. (In the
Appendix, for instance, we provide convincing evidence against the hypothesis that individuals are on a stable demand function of the kind hypothesized by Keynes. But while Robertson’s focus on the demand and supply of funds is more convincing, his analysis is flawed, partly because he failed to recognize the central role of asymmetric information in the provision of credit, partly because he failed to take adequately into account the role of banks in the provision of credit (the subject of the discussion here). In the standard loanable funds theory (without banks), the role of government was limited: It was individual farmers who decide how much seed to supply and demand. Our theory, by contrast, says even here there can be a role through the rules government sets for the functioning of the critical intermediary institutions.

9 It leaves out, for instance, the role of rating agencies, investment analysts, etc. That these markets often do not work well is an understatement, evidenced by the problems in the financial crisis of 2008 and the scandals of the early 2000s. See Stiglitz (2003, 2010).

10 Moreover, ultimately, the supply of funds to large enterprises depends on the funds made available to a variety of intermediaries, which in turn depends on the credit creation mechanisms described here.

11 There are other reasons that firms (including banks) may act in a risk-averse manner. Imperfect information means that there is a separation of ownership and control (Berle and Means 1932; Stiglitz 1985) and firms typically construct incentive arrangements that lead managers to act in a risk-averse manner (Greenwald and Stiglitz 1990).

12 Their analysis also assumes that the risks confronting banks (and other firms in the economy) can neither be insured nor distributed across the economy, for example, because of information asymmetries.

13 In their model, bank assets are not fully tradable, because of information asymmetries. Accordingly, if the perceived risk associated with certain assets the bank holds increases, its willingness to undertake more risks may be adversely affected.

14 Their models also explain amplification: Why a seemingly small shock can have large effects.

15 The inability to diversify oneself of risk generates an important hysteresis effect. Government regulatory policy may exacerbate these problems: When there are, for instance, capital adequacy requirements and banks’ net worth is not evaluated on a mark-to-market basis, then a sale results in the recognition of a loss which is otherwise “hidden”. On the other hand, marking to market forces banks to contract lending (or raise new equity) when there is a (what the bankers believe is a) temporary change in market sentiment against the assets they hold. Of course, the irony is that in other contexts, bankers, as a group, have been the strongest advocates of the “market” and its rationality. But as the 2008 crisis demonstrated, they have shown an impressive level of cognitive dissonance—arguing against subsidies for others (such subsidies would distort markets) but for themselves (without state aid, the whole economy was at risk). See Stiglitz (2010).

16 As we have noted, risk perceptions relate not just to macroeconomic risks, but to risks of particular individuals, firms, and institutions, which in turn have macroeconomic consequences. Thus, it does not suffice to know that the value of say equity has decreased somewhere in the economy. A bank contemplating making a loan to a particular firm wants to know the economic situation of that particular firm. Uncertainties surrounding that are affected both by rules governing transparency and the structure of the economy—the nature of the interlinkages among firms.

We need to distinguish too between structural breaks—the move from agriculture to industry or from industry to services—with shocks to the system that, though large, do not fundamentally alter the structure of the economy. Thus, while Greenwald and Stiglitz (1993a, 1993b, 2003; as well as the large number of papers leading up to those studies and cited there) provided the intellectual foundations for what has since come to be called balance sheet recessions, they have argued that the current economic downturn is not fully described as a balance sheet recession, but rather is best seen as part of a deeper structural transformation. See Deligiorgi et al. (2012, 2016).

17 For a review of the arguments, see Greenwald and Stiglitz (2003).

18 I should emphasize that the significant bank contraction in lending to SMEs is not just a response to conventional rules, and regulations. In 2008 there was a significant increase in risk perceptions, and such changes have a particularly large adverse effect on undercapitalized firms, among which SMEs are heavily represented.

19 This discussion illustrates a more general principle: In markets with asymmetric information, there are marked discrepancies between private and social returns. This can be especially so in the presence of rationing. See Greenwald and Stiglitz (1986).

20 This has been a long standing criticism of Friedman’s criticism of monetary policy in the Great Depression. The fall in the money supply does not necessarily mean that the Fed caused the depression through its contractionary policy. The fall in money holding could be the result of the reduction in (anticipated) economic activity. And the Fed may have been powerless to overcome the exogenous perturbations giving rise to the decline in GDP. Indeed, while it may not have been able to fully offset the underlying forces, it may still have had an unambiguously positive effect: The decline in GDP could have been smaller than it otherwise would have been. See, for example, Tobin (1970).

21 Not surprisingly, there has been enormous controversy over whether the negative interest rates have had a positive or negative effect. Japan’s central bank governor Kuroda tried to design the negative interest rate program in ways which limited the balance sheet effect, while retaining the intertemporal substitution effect. Whether he fully succeeded is part of the debate.


24 As the regressions reported in the Appendix amply illustrate.

25 Again, as evidenced in the regressions described in the Appendix.

26 This effect has been stressed by Jonathan Kreamer in his Ph.D. thesis (Kreamer 2015).

27 There may be a loss of intertemporal welfare from the variability in fiscal expenditures. But if the variability takes the form of infrastructure
investments, and if the investment authority (say an investment bank, like the EIB) were to keep an inventory of good, high return projects, then the flow of “services” from the aggregate stock of public capital would not be highly variable. If the inputs used in public infrastructure investment were highly substitutable with those used in say private construction, and if one of the main sources of variability in aggregate output is private construction, then the social costs of putting the burden of adjustment on public infrastructure investment may be relatively low.

Of course, debt financed government spending may lead to an offsetting effect through the expectation of future taxes, but the conditions under which the adverse consequences of this is fully offsetting are highly restrictive. See Stiglitz (1988).

That is, Ramsey showed that optimal taxation entailed distorting all prices a little from their marginal costs, rather than a single price a lot. Modern monetary policy is based on the hypothesis that government intervention should be limited to interventions only in the short-term interest rate. There is, to my knowledge, no general proof that it is optimal to limit interventions in this way.

For a more extensive discussion of the issues raised here, see Stiglitz (2015a). Even the Fed has begun to recognize the potential importance of these effects. See Yellen (2014).

These are associated with the “repeal” of the second fundamental theorem of welfare economics, implying that issues of distribution and efficiency cannot be separated, as suggested by earlier analyses. See, for example, Stiglitz (1994, 2002a).

See Stiglitz (2015a, 2015c). We should expect such differentials in wealth holdings: Life cycle savers have to be more prudent in their wealth management than wealthy “capitalists”, Giovanni (2014, 2015) provides evidence.

See Stiglitz (2015d) and the references cited there.

This may be especially so if individuals are saving to purchase a home, since the lower interest rate may itself give rise to higher house prices, meaning that the down payment required is also larger.

In 2007, gross domestic investment for the US was 22 percent of GDP, and fell to 20 percent by 2015.

Of course, these numbers do not answer the relevant hypothetical question: What would investment have been but for the lowering of interest rates? Still, the fact that lowering interest rates from 5 to 0 percent has had such small an effect suggests that lowering the interest rate from 0 to minus 2 percent is unlikely to have a large effect.

See Stiglitz (2002b, 2015c) and Guzman and Stiglitz (2013) for a discussion of these issues and the consequent importance of monetary policy coordination.

See, for example, Stiglitz (1998b, 2014).

Much of the argument for an independent central bank is based on enhancing the ability to make such a commitment. If bankers control the central bank, because they benefit from a low inflation rate, it is more credible that the central bank will act in ways which limit inflation. But the crisis of 2008 showed the flipside risks: A central bank captured by the financial sector will do an inadequate job at financial regulation, exposing the economy to the far greater risks associated with financial instability.

Moreover, such large changes give rise to high levels of uncertainty, with strong adverse effects.

The shock does not actually have to be deflationary: All that is required is that the rate of inflation be less than was expected.

Australia has provided income contingent loans for a long time. The US has begun doing so in the case of certain student loans. Stiglitz (2014) and Stiglitz and Yun (2013, 2014, 2016) have proposed doing so for unemployment loans, and Chapman et al. (2014) present a range of other examples of such loans.


This section is adapted from Stiglitz (2016c).

Regulators, legislatures, and courts in antitrust actions have finally begun intervening to curtail the high fees and abusive practices, but the fees remain far higher than what they should be.

Regulators, legislatures, and courts in antitrust actions have finally begun intervening to curtail the high fees and abusive practices, but the fees remain far higher than what they should be.

Cyber security is one of the key problems faced in modern electronic payment mechanisms. The advantages of electronic transactions are, nonetheless, overwhelming, which is why even with monopoly pricing, there has been a shift toward this system.

The major exception, for the purchase of goods and services from abroad, is discussed later.

The evolution of the banking system from the primitive corn economy toward its modern form is interesting and informative. Early banks were really based more on gold deposits than on corn deposits. Those with more gold than they wanted to spend put it in the bank, and the bank lent it out to others. Soon, banks discovered that they could create pieces of paper, claims on gold, that others would accept, and that they could produce more of such pieces of paper than they had gold, in the knowledge that not all holders of these pieces of paper would ask for their money simultaneously. As it gave gold to some who asked for it, it would receive gold from others.

Occasionally, there would be a panic when holders of these pieces of paper worried whether the bank could fulfill its promises, and, of course, when they panicked and all went to the bank to demand their gold, there was not enough to satisfy their demands. The banks would go bankrupt, and the economy could be thrown into a depression.

Deposit insurance was invented to prevent these panics: The government explicitly stood behind the banks’ promises. This gave greater faith that the promises would be honored (so long as there was faith in the government), and this in turn reduced the likelihood of a panic. But if the government was to provide these guarantees, this insurance, it had to make sure that the bank was acting responsibly— for example, lending out money to people who could actually pay it back, and not lending to the owners of the bank and their friends. Gerry Caprio, with whom I worked at the World Bank and who studied government rescues around the world, was fond of saying that there are two kinds of countries— those who have deposit insurance and know it, and those who have deposit insurance and don’t know it. Sweden, before its financial crisis in the 1990s, had no
deposit insurance, but it rescued its banks nonetheless. In the 2008 crisis, suddenly deposit insurance was extended to accounts that had not been fully insured before.

One can understand the government taking on this new role, partially as a result of the magnitude and frequency of the panics and downturns in the market economy in the 19th and early 20th centuries. Moreover, as advanced countries, like the US, transformed themselves from agricultural economies to industrial economies, with an increasing fraction of the population dependent on manufacturing and other nonagricultural jobs, these economic fluctuations took a toll. So long as ordinary citizens had little voice in what government did, so what if so many suffered so much? But with the extension of the franchise and increasing democratic engagement, it became increasingly difficult for government to ignore these mega-failures of the market.

The theory of credit rationing based on information asymmetries provided an explanation for why markets on their own might fail.

63 These are an example of macroeconomic externalities, such as discussed by Anton Korinek, themselves a generalization of the pervasive pecuniary externalities to which Greenwald and Stiglitz (1986) called attention.

64 See Calvo (1998) for a discussion of sudden stops.

65 The experience of Europe and elsewhere has shown that it is not so much government borrowing that gives rise to crises, but national borrowing. In some cases, the national borrowing was government borrowing (Greece), but in many other cases (Ireland and Spain) it was private borrowing. When a crisis hits, the debt quickly moves from the private balance sheet to the public's.

66 See also Dasgupta and Stiglitz (1977).

67 All the data is from the Federal Reserve Economic Database ("FRED") at the Federal Reserve Bank of St. Louis. Contact the author for information on the exact time series used.

68 The 95 percent confidence interval is [0.7287, 0.8811]

69 \(\alpha = 57.9465, \ 95\% CI = \{45.7192, 70.1738\}; \ \beta = 0.3692, \ 95\% CI = \{0.2148, 0.5237\}; R^2 = 0.0895\)

70 \(\alpha = 0.0119, \ 95\% CI = \{0.0091, 0.0146\}; \ \beta = 0.2287, \ 95\% CI = \{0.0796, 0.3777\}; R^2 = 0.0389\)

71 In the case of (A3): \(\alpha = 62.5533, \ 95\% CI = \{52.1177, 72.9949\}; \ \beta = -0.2022, \ 95\% CI = \{-0.4990, 0.0945\}; R^2 = 0.081\) In the case of (A5b): \(\alpha = 0.0069, \ 95\% CI = \{0.0056, 0.0082\}; \ \beta = 0.0897, \ 95\% CI = \{-0.0094, 0.1887\}; R^2 = 0.0141\)

72 To identify the structural break, we used the sup Chow test (Andrews 2003); the tests show that in general there is a lot of parameter instability in the two decades between 1975 and 1995.

73 The only regression in which the coefficient on the assets variable is (barely) significant is a contemporaneous regression of real log output on log real assets – even there it is a meager 0.068. Regressions with lags or leads eliminate significance. Interestingly, there is very strong evidence of a structural break exactly in 2008 in this model.

74 If \(P\) is constant, as in Keynesian theories, then the aforementioned equations give a simple relationship between real output (\(Q\)) and nominal \(M\), for example,

\[
(A6) \Delta Q = \alpha + \beta \Delta M \quad \text{or} \\
(A6') \ln Q = \alpha + \beta \ln M.
\]

75 For (A8b): \(\alpha = 34.7909, \ 95\% CI = \{23.9735, 45.4444\}; \ \beta = 0.2423, \ 95\% CI = \{0.1410, 0.3436\}; R^2 = 0.0895\). For (A8c): \(\alpha = 0.0139, \ 95\% CI = \{0.0118, 0.0159\}; \ \beta = 0.1700, \ 95\% CI = \{0.0592, 0.2808\}; R^2 = 0.0389\). The \(R^2\) is tiny. It is clear that these do not provide (without further massaging) a good description of money demand holdings.

76 See also the later discussion.

77 (A9): \(\alpha' = 0.0127, \ 95\% CI = \{0.0106, 0.0148\}; \ \beta' = 0.2439, \ 95\% CI = \{0.1294, 0.3584\}; Y' = -0.0023, \ 95\% CI = \{-0.0035, -0.0011\}; R^2 = 0.0954\).
(A10): \( \alpha' = 0.0052, 95\% CI = [0.0033, 0.0071]; \beta' = 0.2768, 95\% CI = [0.1015, 0.4521]; \gamma' = -0.0037, 95\% CI = [-0.0053, -0.0020]; R^2 = 0.0924 \\

As we have emphasized repeatedly in this Appendix, we have not engaged in extensive data mining. Our objective is simply to suggest that it is hard to reconcile observed data with the standard hypotheses. Of course, as we argue in the text, the observed money supply may depend on many other variables than income and the interest rate, and controlling for those variables, one should obtain a better fit.

Since the costs of adjusting money balances are small, it might be argued that there is little reason that individuals are not on their money demand curve. Still, one can view money balances as a residual, and if individuals cannot adjust other elements of their spending, then money balances will be off the demand curve in the event of an income (or interest rate) shock. We have estimated the demand function for money, assuming flexible lags, and the Keynesian monetary equation is still rejected.

79 The analysis becomes only slightly more complicated if inflationary expectations depend on the level of output in equilibrium.

80 If inflationary expectations (ie) are a function of the gap between actual and potential output, with potential output fixed for the moment, then ie = H(Y). Writing the IS curve as \( Y = Z(R) = Z(\tau - \pi) = Z(\tau - H(Y)) \), or \( Y = Z(\tau(R)) \).

81 Aggregate demand is a function of the real interest rate, and the IS curve only determines the nominal interest rate (the opportunity cost of holding money). But if expectations about prices or price changes are fixed, a change in nominal interest rates translates directly into a change in real interest rates. A more general case is discussed later.

Figure 7.3 and an associated regression undermine the credibility of the determination of the (nominal or real) interest rate through a Keynesian LM framework.

82 A11d: \( \alpha = 59.5758, 95\% CI = [51.0058, 68.1457]; \gamma = 20.4850, 95\% CI = [11.0189, 29.9511]; R^2 = 0.0745. A11e: \alpha = 0.0075, 95\% CI = [0.0065, 0.0086]; \gamma = 0.0029, 95\% CI = [0.0018, 0.0041]; R^2 = 0.1008. \) In these regressions, the "effective federal funds rate" is used as a proxy for the nominal interest rate.

83 Moreover, the coefficient on the nominal interest rate is of the wrong sign, that is an increase in the nominal interest rate is associated with a higher \( Y \), for a given real interest rate. In fact, this is not as surprising as it seems. Since the nominal interest rate is the real interest rate plus the rate of inflation, an increase in the nominal interest rate \( given \ a \ real \ interest \ rate \) is equivalent to an increase in inflation. We would expect an increase in inflation to be associated with a higher level of real (and nominal) output. The association between the nominal interest rate and real output thus being picked up is the standard Phillips curve. But we are not picking up the hoped for relationship between the interest rate as determined by the government and the level of economic activity.

84 A similar analysis, though, applies if we look beneath the surface, say to the relationship between investment or consumption and (real) interest rates—the channel through which monetary policy is supposed to have much of its effect. Most studies (unmassaged) suggest that income and substitution effects are broadly offsetting.

---

Bibliography


Buffett, Warren. (2003, November 10). “America’s Growing Trade Deficit Is Selling the Nation out from Under Us. Here’s a Way to Fix the Problem: And We Need to Do It Now,” Fortune.


Credit and macroeconomic stability


