NBER WORKING PAPER SERIES
ON
HISTORICAL FACTORS IN LONG RUN GROWTH

THE GREAT DEPRESSION

Peter Temin

Historical Paper No. 62

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
November 1994

Prepared for The Cambridge Economic History of the United States. This paper is part of NBER's research program in the Historical Development of the American Economy. Any opinions expressed are those of the author and not those of the National Bureau of Economic Research.

© 1994 by Peter Temin. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

THE GREAT DEPRESSION

ABSTRACT

This history of the Great Depression was prepared for The Cambridge Economic History of the United States. It describes real and imagined causes of the Depression, bank failures and deflation, the Fed and the gold standard, the start of recovery, the first New Deal, and the second New Deal. I argue that adherence to the gold standard caused the Depression, that abandoning gold started recovery, and that several of the New Deal measures adopted in the recovery lasted in good order for half a century.

Peter Temin
Department of Economics
M.I.T.
50 Memorial Drive
Cambridge, MA 02139
and NBER

The Depression

1. REAL AND IMAGINED CAUSES

The world-wide Depression of the 1930s was an economic event of unprecedented dimensions. There had been no downturn of its magnitude or duration before, and there has been none since. It stands as a unique failure of the industrial economy.

Economic activity in the United States declined from the middle of 1929 through the first few months of 1933. This four-year decline was not smooth, but it was nevertheless an unprecedented and bewildering fall in production. Industrial production declined by 37 percent, prices by 33 percent, and real GNP by 30 percent. Nominal GNP, therefore, fell by over half. Unemployment rose to a peak of 25 percent and stayed above 15 percent for the rest of the 1930s. There were many idle economic resources in America for a full decade. Only with the advent of the Second World War did employment rise enough to absorb the full labor force.

This large event has to be evidence either of a great instability in the economy or of a great shock to it. Traditional scholarship tended to emphasize the former; recent work concentrates on the latter. An older view saw events in the United States in isolation. More recent scholarship insists on the international scope of the Depression and the need to see the United States in an Atlantic if not a world perspective.

The shock that destabilized the world economy was the First World War. More broadly, the shock was the continuing conflict that Churchill called the Second Thirty Years War. This shock

affected both the world economy and the context for policy decisions. Even though the United States emerged from the war as the preeminent industrial economy, it still was part of a world economy. This was nowhere more evident than in the common theoretical basis of economic policy decisions in the United States and Europe.

The war and its associated changes had many effects on the American economy. Three are of primary importance: the changed pattern of international debts and lending, the expansion and collapse of agriculture, and the end of mass immigration.

Before the First World War, Britain had been the primary exporter of capital. The United States, long a recipient of British lending, had only recently begun to reduce its international indebtedness. The British, however, spent much of their foreign portfolio paying for the war. Much of this debt was sold to the United States, which became the world's largest creditor. It went from being a net debtor of at least $3.5 billion in 1914 to a net creditor of over $7 billion three years later. Although there is some double counting in these measures, it is clear that a dramatic change had taken place.

After the war, and after five more years of instability, the gold standard was reestablished. While not precisely the same as before the war, the revived gold standard still mandated deflation rather than devaluation as a remedy for foreign exchange deficits, and placed far more pressure on deficit countries to contract than on surplus countries to expand. The altered international debt structure did not fit well with the

old exchange rates. Reestablishment of the gold standard at (mostly) prewar exchange rates therefore meant that imbalances would proliferate. Britain and Germany would find themselves at the start of the 1930s in financial trouble and without adequate policy tools to deal with the trouble.

American agriculture had been very prosperous during the war, exporting to a Europe hungry for food and fiber. Other countries not directly in the conflict also expanded their capacity, further increasing the world supply of primary products. When peace came, the military demand for these products fell at the same time that European supplies reappeared on the market. The result was falling prices and agricultural distress throughout the 1920s. The effects of the fall in demand were compounded by the post-war deflation, which left farmers with debts high in relation to their incomes.

The problems of American farmers were compounded by overexpansion into marginal lands that proved unsuited to crops in the longer run. Erosion, not prosperity, was the result. The problem of debt was acute since the demand for American farm products had been high for several years and farmers had borrowed to take advantage of high prices. Low prices meant extreme difficulty for farmers who had extended themselves both geographically and financially.

Not all farmers were in trouble. Technical change--particularly in grain production--was rapid in the 1920s. Gasoline tractors began to alter the demands for labor. Large-scale farming began to change the face of the plains. Many wheat

farmers consequently could prosper despite low prices. But cotton farmers, particularly tenant farmers on small farms, were impoverished and even displaced by the combination of low prices and the new technology. Black farmers from the South, finding themselves in this position, migrated to Northern cities in search of work.

Immigration virtually ceased after the war in response to the laws restricting immigration. While not nominally part of the war, the restrictive laws reflected the same hostility that intensified the war. The immigration laws were important politically and socially, but they did not have a large immediate effect on the economy. The rate of population growth had been falling slowly even before the war; ending immigration therefore just accentuated an existing trend. The decline of immigration also was offset in part by the movement of Blacks from the South to the North, replacing the immigrants who might have come in the absence of restrictive legislation. The economic effects of immigration limitation therefore are hard to see.

The distribution of income worsened in the 1920s. In fact, inequality reached its peak just at the start of the Great Depression. This has given rise to the idea that workers could not afford to buy the products of industry in the late 1920s, that "underconsumption" was the cause of the Depression. This view has received some support from observations that housing investment had started to decline before the industrial decline and that purchases of automobiles fell precipitously once the Depression began.

The evidence is not persuasive. Profits rose as share of National Income in the 1920s. The rise was about five percent of National Income. If the propensity to consume was ten percent lower among capitalists than among workers, then the decline in consumption caused by the shift of income was only .5 percent of National Income. This is far too small a decline to have been a potent factor in the Depression; consumption fell by ten percent in 1930 alone. Housing construction also frequently moves to its own rhythm, and the rapid fall in automobile sales is consistent with almost any story of the Depression. "Underconsumption," and its converse, "overproduction," are not useful concepts in the investigation of the Great Depression.

Industrial production began to decline in 1929. This decline did not appear to be the start of a great depression; it was a downturn similar in appearance to the sharp, but brief, downturns in 1907 and 1921. It was caused by contractionary monetary policy in 1928 and 1929. This credit contraction was not the result of international strains; the United States and France had accumulated the bulk of the world's gold reserves. It was an attempt by the Federal Reserve to arrest what the Fed considered a speculative boom in stock prices. Economists have debated ever since whether the dramatic rise in stock prices at the end of the 1920s was indeed a speculative bubble. The jury is still out.

The tightness of credit was severe enough to explain most of the fall in production and prices during the first phase of the Depression. Although the Fed believed that it could restrict

credit to Wall Street without harming the rest of the economy, it was mistaken. The Bank of England thought it could use monetary policy to preserve the value of the Pound without affecting the domestic economy. It too was wrong.

The initial shock to the economy was not, however, strong enough to cause a deep and protracted depression. There is no sign that the economy was so fragile that interest rates of six percent could cause an economic tailspin. If the economy had been that fragile, then the Depression should have started with the short, sharp decline in 1921.

Instead, there were additional shocks during the economic downturn that continued and even accelerated the contraction. Four events from the fall of 1929 to the end of 1930 have been accorded prominent roles in the propagation of the Depression. The four events are the stock-market crash in New York, the Smoot-Hawley tariff of 1930, the "first banking crisis" described by Milton Friedman and Anna Schwartz, and the worldwide collapse of commodity prices.

Time has not been kind to the school of thought that blames the Depression on the stock-market crash. The stock market has gone up and down many times since then without producing a similar movement in income. The most obvious parallel was in the fall of 1987. The isomorphism was uncanny. The stock market fell almost exactly the same amount on almost exactly the same days of the year.

If the crash of 1929 was an important independent shock to the economy, then the crash of 1987 should have been equally

disastrous. The stock market had grown in the intervening half century, and news of the stock market was pervasive. Many more people owned stocks in 1987, even though stocks probably were a smaller part of personal wealth than in 1929. There were strains on the international economy to rival those of the 1920s, centering on American rather than German borrowing. And stock markets around the world were much more closely synchronized in 1987 than in the late 1920s.

Despite a flurry of speculation in the popular press, the world economy did not turn down in the fall of 1987. The boom in production that had been under way for five years continued apace. It follows that a stock-market crash is not a big enough event on its own to initiate a depression.

In neither case was the change cataclysmic. Stocks retained the major part of their values after each crash. The effects of the change in value therefore were minimal. The stock market crash in 1929 helped communicate the Fed's tight monetary policy throughout the economy. But it was not a strong or independent force of its own. The crash of 1987 reflected nervousness about the Reagan fiscal policy but, like its earlier cousin, had little effect on expenditures.

That is not to say that the crash of 1929 had no effect. As a part of the propagation mechanism, the stock market crash had several effects. It reduced private wealth by about ten percent. It increased consumers' leverage, that is, the ratio of their debts to their assets. And it no doubt increased consumers' uncertainty about what the future would bring. Each of these

effects tended to depress consumer expenditures, particularly the demand for consumer durables. The American economy experienced a fall in consumption in 1930 that was too large to be explained easily. These influences comprise part of an explanation.

The idea that the Smoot-Hawley tariff was a major cause of the Depression is an enduring conviction. It was stated at the time, reiterated after the Second World War, and has found its way into popular discussion and general histories. Despite its popularity, however, this argument fails on both theoretical and historical grounds.

A tariff, like a devaluation, is an expansionary policy. It diverts demand from foreign to home producers. It may thereby create inefficiencies, but this is a second-order effect. The Smoot-Hawley tariff also may have hurt countries that exported to the United States. The popular argument, however, is that the tariff caused the American Depression. The argument has to be that the tariff reduced the demand for American _exports_ by inducing retaliatory foreign tariffs.

Exports were 7 percent of GNP in 1929. They fell by 1.5 percent of 1929 GNP in the next two years. Given the fall in world demand in these years, not all of this fall can be ascribed to retaliation from the Smoot-Hawley tariff. Even if it is, real GNP fell over 15 percent in these same years. With any reasonable multiplier, the fall in export demand can only be a small part of the story. And the decline in export demand was partially offset by the rise in domestic demand from the tariff. Any net contractionary effect of the tariff was small.

2. BANK FAILURES AND DEFLATION

The primary propagating mechanism in the American Depression identified by Friedman and Schwartz in their classic Monetary History of the United States revolved around banking panics. They identified the first of three banking crises in December 1930 with the failure of the Bank of United States. Had the banks responded to panic by restricting payments (a nineteenth-century practice), Friedman and Schwartz claimed, the Depression need never have happened. They argued that restriction in 1893 and 1907 had quickly ended bank suspensions and promoted economic recovery.

The events after the restriction of payments in 1893 and 1907 show that the American economy of the time was very stable. A restriction of payments is defined as a refusal on the part of banks to honor their commitment to exchange deposits for currency at par. When a single bank refused to redeem its obligations at par, it was legally bankrupt. But when banks acted in concert, there was an effective devaluation of deposits against currency.

The price of deposits was determined, like all prices, by the forces of supply and demand. People who were afraid that the price of deposits would decline wanted to sell, driving down the price. People who thought that the price of deposits had already fallen and was due to rise back toward par wanted to buy, driving up the price. The market price was where the supply from the former group just matched the demand from the latter. The currency premium in 1893 and 1907 was never more than 4 percent; it had fallen to almost nothing in a month, even though full

resumption came somewhat later. Most people, in other words, expected the banks to resume payments at par speedily. They did not anticipate a major depression or further bank crises. They did not rush to sell discounted deposits.

Friedman and Schwartz therefore adopted an inconsistent position toward the banking crisis of 1930. On the one hand, they said that the economy was unstable, that a small event set off the Great Depression. In fact they traced the cause of the Depression back to the death of Benjamin Strong in 1928, even though their main story starts with the banking crisis in 1930. On the other hand, they implied that the economy was very stable, that a restriction of payments would have resulted in only a tiny change in the price of deposits--like the 2 or 3 percent seen in 1893 and 1907--and that this change would have brought the economy back onto an even keel. They cannot have it both ways. Either there was an impulse more powerful than the death of the head of the New York Fed or the economy was far less stable in 1930 than in 1893 and 1907 (and a suspension of bank payments would have had only limited impact). As noted above, the former position is taken here.

Friedman and Schwartz argued that the banking failures in December 1930 reduced the supply of money by increasing the banks' demand for reserves and the public's demand for currency. This in turn depressed spending. If it happened this way the monetary restriction should have affected income through the financial markets. Even if the progress of the Depression eventually led to lowered demand for money and low interest

rates, we still should observe a rise in interest rates at the time of the banking crisis--before any effects of the banking failures had run their course. No such credit stringency is observed at the start of 1931.

There was an increase in bank failures in November and December of 1930. But much of the rise of liabilities in failed banks was due to the failure of just two banks. Caldwell and Company failed in Tennessee, and the Bank of United States failed in New York City. Both of these banks had undergone reckless expansion in the late 1920s, and their overblown empires collapsed under the pressure of the emerging Depression.

If the liabilities of these two banks are subtracted from the total liabilities in failed banks in those months, it emerges that the rise in other bank failures was clearly noticeable, but not of the same scale as the rise of bank failures in the summer and fall of 1931. The level of bank failures also returned to its earlier level at the end of 1930, where it stayed for four months. There was no reaction in the markets for short-term credit, aside from a temporary rise in rates in Tennessee. There was no fall in the stock of money at the end of 1930. There was no shock to the quantity of money that could have produced a large macroeconomic effect. There was no direct effect of the "first banking crisis."

Instead, there was the beginning of a movement to increase currency in the hands of the public. This movement was small relative to the other events of the time. The change in the rate of growth of the money supply from the 1930 "banking crisis,"

therefore, was swamped by changes from other causes. As a result there was no reason to expect interest rates to react to such a change.

Alternative mechanisms have been proposed for the effects of banking crises. The most popular recent view, due to Ben Bernanke, argues that the effect of banking panics operated through credit rationing. Credit became harder to get for many borrowing firms, who had to shop around for loans or do without. Published interest rates did not reflect this added cost because they were the cost of loans granted, not loans refused.

Any lender had imperfect knowledge of the comparative risks of different firms. Banks specialized in making the best use of the available data. They acquired most of the loan business because they were the low-cost intermediaries. When banks failed, they no longer could extend credit, and other banks switched into more liquid loans to protect themselves. This reduced the supply of the most efficient intermediation services and raised their cost and consequently the cost of loans to borrowers.

This hypothesis typically is tested by time-series regressions explaining the movements of industrial production. A more direct test examines the progress of different industries. Bernanke noted explicitly that the rising cost of credit intermediation hurt households and small firms much more than large firms. Bank failures then should have hurt industries populated by family firms and other small businesses more than those composed of large, well-established firms.

But the presence of large firms is positively related to the fall in production, not negatively as the credit rationing hypothesis predicts. Comparison with 1937-38 reveals that the cross-sectional pattern of industrial decline in the Great Depression was not unusual. Despite the banking crises, the pattern of industrial decline--as opposed to its magnitude and duration--was unexceptional. There is no evidence that the pattern of industrial decline was rendered unusual by the dramatic collapse of the banking system.

We need to take care here not to throw the baby out with the bath water. The American financial system was being battered at the end of the 1920s by the stock-market decline, business failures, bank failures, and international events. After the stock-market crash, firms shifted their new offerings from stocks to bonds. Net new stock offerings fell by $2.5 billion from 1929 to 1930, while net new bond offerings rose by $1.4 billion The price of lower grade industrial bonds then began to decline in late 1930. The increased supply of bonds lowered their price. Business and bank failures decreased the demand for bonds by increasing their perceived risk.

A gap opened up between the cost of bank loans to firms that could borrow at the prime rate (falling steadily in 1929 and 1930) and the cost of industrial bonds for smaller firms. This is the kind of premium that Bernanke was talking about, although market prices reflected this premium rather well. The spread between the prime rate and other interest rates is a good indicator of monetary pressure even without bank failures. In

addition, since bonds were being reclassified to show their increased risk at this time, the return on risky bonds was rising for two reasons: bonds of a given riskiness were worth less, and any given bond was becoming more risky. The largest firms had access to credit at costs far lower than smaller firms. The cross-sectional pattern of industrial decline shows, however, that access to credit did not determine which industries declined.

Bank failures undoubtedly accentuated the Depression. International comparisons of countries with and without banking difficulties suggests that banking difficulties in general were harmful. But the mechanism by which bank failures had their effects is not clear. As a result, their importance in the American contraction is still a matter of dispute.

At about the same time as the stock-market crash, the prices of raw materials and agricultural goods--which had already been tending slowly downward--began to fall precipitously. Charles Kindleberger identified the fall in commodity prices as one of the primary channels through which deflation spread, from "stock prices to commodity prices to the reduced value of imports." Although a change in prices only reallocates income, he argued, the effect is asymmetric. The losers found their budgets curtailed and were forced to cut spending; the winners did not correspondingly increase theirs.

The prices of agricultural products and raw materials had been falling in the 1920s as a result of the overexpansion of production during and after the First World War. Various

attempts to prop them up through tariffs or purchases had proved ineffective. Inventories accumulated as the production of many raw materials exceeded demand at the market price. The costs of holding these stocks and conducting orderly marketing rose as credit conditions were tightened at the end of the 1920s. In the credit squeeze that always came to the United States in the fall, many owners of these inventories failed in 1929. Further price declines were of course in store as the demand for raw materials contracted.

The effects of the price declines on different groups need to be distinguished. For countries whose agricultural or mineral products were the main source of foreign currency, the fall in price was a disaster. Devaluations were the frequent response. But for importing countries, the decline in product prices was a plus. Even if Kindleberger is right and the price decline did not cause spending to rise, it allowed greater monetary ease. (It reduced any inflationary pressure, and it increased the real money supply.) The United States experienced both effects. Farmers suffered, while the rest of the economy gained. The net effect of the initial fall in commodity prices in the United States therefore probably was positive, since there were many more consumers than producers of these commodities in the United States.

The gain was limited, however, as prices in general began to decline in 1930. The more pervasive deflation cannot be attributed to the breakdown of cartels, and it was not closely correlated with the stock market. It was a reflection of the

falling aggregate demand that came from the preceding credit stringency. Both the stock-market crash and the collapse of raw materials prices were part of the propagating mechanism by which this tightness affected economic activity, but they were only part of a complex picture.

There are two effects of a general deflation, static and dynamic. The static effect, known sometimes as the Keynes effect, is to increase monetary ease. A given nominal stock of money buys more goods; real balances rise. The fall in aggregate demand affects prices more than production. The deflation <u>substitutes</u> for depression.

The dynamic effect, known sometimes as the Mundell effect, works through expectations. If people expect the deflation to continue, they anticipate that prices will be even lower in the future than they are now. They hold off on purchases to take advantage of the expected lower prices. They are reluctant to borrow at any nominal interest rate because they will have to pay back the loan in dollars that are worth more when prices are lower than they are now. In short, the real interest rate rises above the nominal rate. The deflation <u>causes</u> depression.

To distinguish between these two effects, we need to know when people began to anticipate continuation of the deflation. It is always very hard to discover expectations since they are not directly observed. Current research suggests that people did not anticipate the Depression or even a large deflation at the time of the stock-market crash. It seems most likely that expectations began to change near the start of 1931 when the

economy failed to recover quickly, as it had in 1907 and 1921. At that time, the Keynes effect was overwhelmed by the Mundell effect; the deflation became destabilizing.

By the summer of 1931, therefore, the United States was in the grip of a severe depression. But if recovery had come then, the downturn would have still been within the historical range of business fluctuations. It would have been a hard time, but not the disaster of the 1930s.

3. THE FED AND THE GOLD STANDARD

The growing depression was turned into the Great Depression by the Federal Reserve in the fall of 1931. A series of currency crises hit Europe in the summer of that year. The Credit Anstalt, the largest bank in Austria, failed in May, leading to a run on the schilling. This was followed by a run on the German mark in June and July. Depositors drew down their deposits in the large Berlin banks, which then replenished their cash by selling bills to the Reichsbank. But the Reichsbank ran out of cash with which to monetize the banks' reserves, and it was not able to borrow from other central banks on acceptable terms. The German government instituted currency controls over the mark to arrest the outflow of funds.

The pressure on the Reichsmark was contained by exchange controls, and the international panic spread to the pound. The Bank of England was unwilling to raise Bank Rate, which it kept relatively low throughout the crisis. It then had to support the pound by direct intervention, that is, by buying pounds from whomever wanted to sell. The Bank of England needed reserves to

make these purchases, which it borrowed from the United States and France. The borrowed reserves, like the Bank's own reserves, were quickly spent. On September 20, 1931, the Bank of England threw in the towel and announced the suspension of the gold standard.

Germany and Britain therefore both abandoned the gold standard, albeit in different ways. The Germans preserved the price of the mark, but restricted the sale of gold. The British continued to sell gold, but no longer at a fixed price. Neither country made immediate use of its new freedom from international pressures. The Germans continued to deflate, and the British waited for six months before expanding.

When the pound was devalued, investors figured the dollar was next. They rushed to sell dollars before the United States devalued. But the Fed was not about to yield to this international pressure; it chose to preserve the value of the dollar. It raised interest rates and accelerated the decline in the money supply. The result was that interest rates in the United States rose sharply in the fourth quarter of 1931, and credit became harder to get. Industrial production--which had paused briefly in its descent in the spring of 1931--continued to fall. The Depression in the United States intensified.

Unlike the "first banking crisis," the effect of the Fed's response to Britain's devaluation is clearly visible in the growth of the money supply. The rate of monetary growth fell to its lowest level in the Depression in October, just after the British devaluation.

The Fed's open market purchases of 1932 were in part a response to the clamor for expansion in response to the monetary contraction of late 1931. The purchases succeeded in restoring the rate of money growth only to the low levels prevailing before the summer of 1931, and they were abandoned by midyear. As interest rates fell, the lower rates reduced earnings of banks holding bills and threatened their already precarious solvency. The Fed's objectives as overseer of the nation's banks and of the national economy came into conflict. In addition some Federal Reserve banks were running out of "free gold," that is, excess reserves on their currency. The Federal Reserve banks were unwilling to pool their reserves by interbank borrowing, and the effective reserve of the system was set by the weakest banks. The French and then the British began to fear eventual devaluation and to withdraw their dollar balances in New York. The open market purchases of 1932 were abandoned under this pressure. They were a temporary aberration in Hoover's deflationary policy, not the start of a new, expansionary policy.

The Fed's contraction to save the dollar is often regarded as an isolated act of foolishness. But it was not that at all. It was part of a concerted effort to preserve the gold standard--even as it was collapsing in Europe. The Fed acted consistently, if misguidedly, throughout the contraction. It interpreted the lack of excess reserves in the banking system as a sign of monetary ease. It did not see its job as the restoration of full employment by monetary expansion. In fact, it did not see its way clear to try for this goal because to do

so would threaten the value of the dollar. The failure of the open market purchases of 1932 confirmed the view that the Fed was severely limited by the gold standard. No one in the Hoover Administration seems to have questioned the premise that the gold standard itself was worth saving.

4. THE START OF RECOVERY

There appear to have been two low points in industrial production, in 1932 and 1933. Looking at the monthly indexes themselves, it is just as likely that the abortive recovery of 1932 was part of the way down as part of the way up. Sustained recovery, however, started only in 1933. The Federal Reserve's open market purchases of 1932 were halted after only a few months; they failed to provide an impulse strong enough to arrest the economic decline. As Irving Fisher (who was better at understanding than at predicting) observed at the time, "Those who imagine that Roosevelt's avowed reflation is not the cause of our recovery but that we had reached the bottom anyway are very much mistaken."

Far from ending, the Depression seemed to be irresistible in 1932. Business was bad everywhere. Hardly anyone expected to make money from new investments, and new investments consequently were few. Few jobs were secure, and many workers were getting used to unemployment as a way of life. There did not seem to be any effective antidote.

This view, however, was wrong. The Depression only seemed to have a momentum of its own. The downward spiral was perpetuated and accelerated by the policy stance of governments

and central banks in the major industrial countries. Contracts and investments had been made in the expectation of further deflation. But activities only reflected these expectations because government policies warranted these expectations.

Investors and workers were not responding to isolated government actions. They were acting in accord with the underlying policy regime, that is, the systematic and predicable part of all decisions. The policy regime is the thread that runs though the individual choices that governments and central banks have to make. It is visible even though there inevitably will be some loose ends, that is, some decisions that do not fit the general pattern. These isolated actions have little impact because they represent exceptions to the policy rule, not new policy regimes.

It was not a trivial task to change the direction of the economy. People were locked into their bargains in the short run. More important, they had expectations about the policy regime that had to be changed. They regarded actions that departed from the deflationary policy regime initially as aberrations, individual actions that had no implications for the regime as a whole. They needed to be convinced that the regime had changed, not simply that the policy process was uneven.

There needed to be a dramatic and highly visible change in policy. There needed to be symbols of the change that could be widely understood and that would be hard for policymakers adhering to the old regime to send. But changing expectations alone was not enough to turn an economy around. The new

expectations needed to be supplemented by effective macroeconomic policies.

The primary thread running through the deflationary policies of the early 1930s was adherence to the gold standard. Devaluation--going off gold in the parlance of the day--was therefore a good signal of a changed policy regime. It was not an infallible indicator, as was shown by the British experience of 1931, but it was the best one available.

Devaluation also had direct effects. The stimuli from relative prices and monetary ease were added to the effects of a new policy regime. In fact the interaction was beneficial. Devaluation speeded the change in expectations by showing a tangible sign of the altered regime. And the changed expectations that came from the initiation of a new policy regime amplified the effects of the devaluation.

The change in policy regime can be seen clearly in the Federal government. The Hoover administration followed a policy that became more orthodox over time. It was highly traditional in its support for the gold standard and its focus on efforts to bolster the credit markets rather than the economy directly. Although not initially deflationary, Hoover drew exactly the wrong lesson from the currency crisis of 1931 and became a strong deflationist.

The Reconstruction Finance Corporation (RFC) is an exception that proves this rule. Hoover's most forceful expansionary effort, the RFC was strictly limited in its goals. Hoover wanted the RFC to promote investment, but he limited the RFC to an

agency function, making its finance "off-budget" and emphasizing the "soundness" and "bankable" quality of supported projects. The RFC in addition was directed at the relief of financial institutions; two-thirds of its 1932 loans went to them. The expansionary aspect of the RFC therefore was designed to be a mild exception to the prevailing deflationary regime, not the first step in a new direction.

The Federal Reserve maintained a passive stance in the early stages of the Depression, which was replaced by active contraction in response to the run on the dollar in 1931. The Federal Reserve's steps toward expansion in March to July of 1932 were halted when the open market purchases alarmed other central banks and threatened the precarious health of member banks by lowering the returns on bank portfolios. The Glass-Steagall Act of 1932 reiterated support for the gold standard.

It was not clear during the presidential campaign of 1932 that Roosevelt would implement a change of policy regime. He had recently raised taxes in New York to balance the state budget, and he emphasized a balanced federal budget as well. He strongly criticized Wall Street, business, and utilities during the campaign and employed a generally anti-business rhetoric. These were not features of a candidate one would expect to help the business environment.

The first sign that a new policy regime was on the way came after the election, in December 1932, when Roosevelt torpedoed Hoover's efforts to settle war debts and reparations multilaterally, signifying his opposition to continuation of the

existing international financial cooperation. A change in regime became more tangible in February 1933, when the President-elect began a serious discussion of devaluation as part of an effort to raise commodity prices. This talk led to a run on the dollar and helped cause the Bank Holiday in March. The New York Fed found its gold supplies running dangerously low at the start of March. It appealed to the Chicago Fed for help. But the midwestern bank refused to extend a loan to its New York cousin. Its different view of the world echoed the contrast between German and French attitudes when the Reichsbank appealed for a similar loan in July 1931. The New York Fed appealed to Roosevelt to shut down the entire national banking system, a draconian way to force cooperation among the Federal Reserve banks.

Once inaugurated, Roosevelt declared the Bank Holiday. He also imposed controls over all foreign exchange trading and gold exports. He ended private gold ownership and took control over the sale of all domestic gold production. The Bank Holiday was a failure of economic policy, but the controls introduced in the Holiday allowed Roosevelt to avoid speculative disequilibrium when he began to devalue the dollar.

Roosevelt effectively devalued the dollar on April 18, when he announced that he would support the Thomas amendment to the Emergency Farm Mortgage Act of 1933, which allowed him to set the price of gold. At the same time he prohibited the private export of gold by executive order. The dollar, freed from its official value, began to fall. It dropped steadily until July, when it had declined between 30 and 45 percent against the pound.

Barry Eichengreen has shown that a devaluation not only has a favorable terms-of-trade effect, but that it also frees domestic macroeconomic policy to expand the economy. If this opportunity is taken, then devaluation need not be a beggar-thy-neighbor policy. And if all countries devalue, then monetary and fiscal policies could ease all over the world. By 1933, virtually all countries except the die-hard members of the gold bloc had devalued, and recovery could begin.

The clarity of Roosevelt's change in policy was unmistakable. The United States was under no market pressure to devalue. Despite the momentary pressure on the New York Fed, the United States held one-third of the world's gold reserves, ran a chronic foreign trade surplus, and dominated world trade in modern manufactures like automobiles, refrigerators, sewing machines, and other consumer durables. The devaluation was a purely strategic decision that appeared without precedent. Orthodox financial opinion recognized it as such and condemned it. Senator Carter Glass called it an act of "national repudiation." Winthrop Aldrich, the new chairman of the Chase National Bank, thought devaluation was "an act of economic destruction of fearful magnitude."

Devaluation was only one dimension of a multifaceted new policy regime. During Roosevelt's First Hundred Days, the passive, deflationary policy of Hoover was replaced by an aggressive, interventionist, expansionary approach. The New Deal has been widely criticized for internal inconsistency. There was, however, a steadily expansionary bias in policy that added

up to a marked change from the Hoover administration.

A major step toward a compatible monetary policy was taken when Eugene Meyer resigned as chairman of the Federal Reserve Board. Meyer, an orthodox Wall Street financier with a strong international orientation, was replaced by Eugene Black, governor of the Atlanta Federal Reserve Bank, who was compliant to the wishes of the administration. The Federal Reserve cut the discount rate in both April and May, from 3.5 to 2.5 percent, and its holdings of U.S. Treasury securities rose from $1.8 to $2.4 billion between April and October. The change in monetary regime initiated by devaluation was extended by reforms of the Federal Reserve System that initiated what contemporary observers labeled a new monetary system.

Devaluation received wide, although not (as we have seen) universal, support. J. P. Morgan told reporters, "I welcomed the reported action of the President and the Secretary of the Treasury in placing an embargo on gold exports." Keynes advised a client that, "President Roosevelt's programme is to be taken most seriously as a means not only of American but of world recovery . . . . [H]is drastic policies have had the result of turning the tide in the direction of better activity". Congress easily passed the New Deal measures. The business and farm community welcomed the possibility of "reflation."

The reaction to Roosevelt's new policy regime was immediate. The stock market rose as the value of the dollar fell, signifying the business community's favorable reception of the new regime. Stock prices, which had been bouncing around at a low level in

1932, almost doubled in the second quarter of 1933. Farm prices--or at least the prices of those products such as cotton and grain that were traded on international markets--rose sharply as well.

Recovery, however, was not instantaneous. The direction of change had been reversed. People were no longer in the grip of deflationary expectations. But business remained bad, and unemployment remained high. The national product grew rapidly after 1933. Looked at in isolation, the recovery appears strong. But unemployment remained above 15 percent until 1940. The United States was "in the Depression" throughout the 1930s.

The U. S. was depressed despite a veritable flood of anti-depression activity from the Roosevelt Administration. The New Deal, as Roosevelt labeled it, was a multifaceted program reaching into almost every corner of economic life. But while the New Deal transformed American government and life, it did not lead to a full recovery.

5. THE FIRST NEW DEAL

The New Deal consisted of three primary initiatives: reform of the banking system, increasing government control of production, and initiation of a social "safety net." The first two of these were begun in the famous "100 days" of 1933. Roosevelt bombarded Congress with myriad bills in the second quarter of 1933 that sparked the recovery and reshaped the American economy. The third initiative came later, in Roosevelt's second term. The "Second New Deal" was an effort to extend the benefits of recovery to the whole population.

The financial system was in a state of collapse when Roosevelt took office. The Bank Holiday was a clumsy response to a problem created by Roosevelt's loose talk of devaluation and tension within the federal structure of the Federal Reserve System. It represented yet another demonstration of the banking system's inability to deal with the financial strains of the Depression.

Had the economy continued to decline, the Bank Holiday would have been only the worst crisis to that time. But the economy began to recover as Roosevelt unveiled his new policy--and carried out his threat to devalue the dollar. The Bank Holiday therefore stands at the threshold of recovery. It has been regarded even as the first step in recovery, as a clearing of the air or a cleansing of the banking system.

This romantic view is wrong. The Bank Holiday was yet another symptom to the Depression disease. It was a desperate bid for time to think on the part of the new administration. By itself, it was part of the problem, not part of the cure.

But the breathing space acquired during and after the Bank Holiday was used, as noted above, to announce and implement a new macroeconomic policy. A key part of the new policy had to be reform of the banking system. In June, Congress passed and Roosevelt signed the Glass-Steagall Act of 1933, known also as the Banking Act of 1933.

The aim of the Glass-Steagall Act was to reduce instability in the banking system. To that end it disallowed the combination of investment and commercial banking that had characterized the

large banks before the Depression. One motive for this divorce was the belief that banks' activities in the securities markets had increased their vulnerability in the recent years of economic decline. This was a reasonable hypothesis, but it appears to have been wrong. Banks with integrated securities departments in fact fared better than other banks in the decline.

The reason is clear in light of modern research, although it would not have been then. The returns to a portfolio of financial assets depends on the variation of the price of each asset and on the correlation between the movements of different assets. If the prices of all assets move together, then the portfolio's price will move too. But if the prices of the individual assets move independently, then the price of the portfolio may move less, even dramatically less, than the price of any asset within it. Each asset may act as a hedge for each other. Even though stock prices declined in the early 1930s, stock market movements were not closely correlated with financial problems. Integrated banks, as a result, had less trouble with banking crises than unintegrated banks.

Another reason to divide commercial and investment banking was to reduce the power of the "money trust." Congressional hearings on banking held by Congressman Pecora exposed banker arrogance and--to some--a banking conspiracy against the people in addition. The ability to sell securities through bank branches, pioneered in the 1920s by the National City Bank, had enlarged the resources available to the "money trust." Congress chose to eliminate that source of funds to reduce the strength of

the investment bankers.

The "money trust" has appeared to be elusive to later investigators. Investment bankers, to be sure, were wealthy men who had little use for mere mortals and particularly for Congressmen. They clearly were paid well for their banking services. But their pay is only part of the question; the rest is whether the rest of us were made better off or worse off by the actions of the investment bankers. Pecora looked only at the possibility of monopoly profits. Historians have looked also for the benefits to the economy of powerful and integrated banks. While no theory has emerged to clarify this point, the examples of Germany and Japan, whose industrial growth is generally thought to have been aided by their integrated banks, are suggestive.

In addition to separating commercial and investment banking, the Glass-Steagall Act also introduced federal deposit insurance. The act mandated the formation of the Federal Deposit Insurance Corporation (FDIC) that would insure deposits in member banks of the Federal Reserve system. The FDIC was to begin operations in 1934, but its opening was delayed for a year, until July 1, 1935.

The immediate effect of federal deposit insurance, therefore, was virtually nil. Despite its announcement at the depth of the Depression, the FDIC did not begin operations until well after devaluation had occurred and recovery had begun.

In the longer run, deposit insurance clearly increased the stability of the banking system. It prevented the kind of cumulative banking runs that had characterized the early 1930s.

Fears for a single bank led depositors to rush to withdraw their deposits before the bank failed and their deposits were lost. To acquire reserves to pay depositors, the troubled bank called in its outstanding loans and borrowed from other banks. Holders of these loans went to their banks to get funds, spreading the pressure. Banks previously doing well found themselves in trouble, particularly if they had loaned to less fortunate banks. Each bank failure intensified the pressure on all other banks.

This cumulative movement was short-circuited by deposit insurance. Depositors did not need to fear for loss of their deposits, although they could experience some inconvenience as the FDIC took over. And troubled banks did not need to borrow from other banks. The Bank of New England, to cite a recent example, failed in early 1991. Depositors lined up in classic fashion to withdraw their funds at the end of one week. But the FDIC stepped in over the weekend and announced that it would pay all insured deposits (up to the legal limit). There were no lines on Monday, and no other bank in the region was "infected" by fear.

This stability, however, was not achieved without cost. As deposit insurance spread, both by the expansion of FDIC coverage and the formation of similar insurers for other types of financial intermediaries, the need for depositors to scrutinize their banks declined. Instead of inquiring whether a potential recipient of your savings was sound, you asked if their deposits were insured. Banks were left to their own devices under increasingly loose supervision, a condition of "moral hazard."

The problem came to light at the end of the 1970s. After a decade of inflation, banks which held fixed return securities such as mortgages were in bad shape, even insolvent. Congress tried to rescue the situation by allowing banks more freedom to invest, hoping that the banks would pull themselves up by their own bootstraps. But without monitoring, banks undertook risky--even foolish--investments. If they were successful, the bank was saved. If not, the FDIC would pick up the pieces.

By the end of the 1980s, the problem had grown to huge proportions. The FDIC was running out of funds, and Congress was debating how much money it needed to inject into the banking system to prevent a collapse reminiscent of the Depression. The problem, as even this capsule history makes clear, was due both to deposit insurance and to the subsequent relaxation of bank regulation. The existence of the FDIC created a moral hazard. This problem was contained up to 1980 by bank regulation; it surfaced only when bank regulation was eased. To achieve stability, we need either to reimpose bank regulation or sharply curtail the FDIC.

The Glass-Steagall Act did not end the Depression, nor did it ensure banking tranquility ever after. It did provide a setting in which banks were stable for over half a century during a great expansion of the American economy. That is a fine accomplishment. We should not forget it, even as we consider revising or repealing the act itself.

The second strand of the New Deal began a half-century of social democratic policies in the United States. The government

asserted its control over many parts of the economy, substituting political control for the apparently misleading signals of the market. This ideology was embodied chiefly in two important bills: the National Industrial Recovery Act (NIRA), which created the National Recovery Administration (NRA), and the Agricultural Adjustment Act (AAA).

The NIRA was passed on June 16, 1933. It induced employers and employees to get together and make agreements on hours of labor, wages, and other conditions of employment. As long as these agreements were in accord with codes drawn up by the government, they were exempt from the antitrust laws. In fact, the government tended to approve codes drawn up by industry trade associations because it proved too difficult for the federal bureaucracy to formulate the needed codes. Despite this partial delegation of power to employers, the government had introduced itself into the very bowels of employment contracts.

The codes widely mandated shorter hours of work in an attempt to spread the available work over more people. They also included sharp wage increases. The wage gains would have been impressive in the best of times; they were unprecedented at a time of mass unemployment. The employers agreed to raise wages because they in turn were allowed to raise prices. The effect of the NIRA, therefore, was to raise both wages and prices.

Contemporary thought was focussed on the aggregate price level. The NIRA was part of Roosevelt's program of "reflation." The price rise was designed to mark the end of the old deflationary policies, revive expectations of a recovery, and

promote investment. It succeeded only in part.

The NIRA was an important part of Roosevelt's new policy regime. Devaluation had freed economic policy from the need to define its objectives in accord with international economic conditions. Policy could be set for domestic needs, and the exchange rate would adjust. Roosevelt clearly signaled his intention to look inward by his sabotage of the World Economic Conference in July, 1933. The NIRA gave substance to this intent, assuring investors that Roosevelt would exploit the opportunity he had created.

The rise in prices lowered the expected real interest rate. If people expected deflation to continue in the absence of the NIRA, this was an important change. But if people assumed that the devaluation had ended the deflation, then the NIRA was not as big a change. Nominal interest rates were very, very low by 1933. In the absence of deflationary expectations, any sound investment could earn the needed interest.

Offsetting this beneficial effect were two deleterious effects. First, as noted above, monetary policy had turned from passively declining to actively expanding. The rise in prices under the NIRA absorbed much of the initial increase in the money supply. The expansion of nominal income induced by easy money went more into higher prices than higher employment.

Second, wages rose more than prices. This was considered a gain by the federal administration and by labor, but they did not think through the effects on employment decisions. For if wages rise relative to the cost of products, employers will reduce the

number of employees they hire. As labor becomes an expensive factor of production, employers minimizing costs will substitute other inputs for the more expensive labor. The rise in real wages therefore acted to preserve unemployment--not to reduce it.

This paradoxical conclusion has generated research into the dynamics of this peculiar labor market. How can real wages rise in the presence of massive unemployment? In fact, why didn't wages continue to fall during the 1930s?

Two hypotheses have been proposed. Some historians have argued that firms were paying "efficiency wages." In other words, employers consciously raised wages above the market-clearing level in order to attract good workers to their firm and to induce workers to put effort into their jobs. Since the efficiency wages were higher than those available elsewhere, this argument goes, workers would vie to get and hold jobs at these wages.

This appealing story is not much use in explaining events in the 1930s. People worked hard at jobs in the Depression because the alternative was not another job at lower pay; it was a high probability of no other job at all. The efficiency wage theory presumes that other jobs are freely available, which was hardly true in the Depression. Extending this line of reasoning, the theory also says that the wage premium for efficiency wages should be high when employment is high and low when employment is low. The efficiency wage in 1934 therefore should have been extremely low. It cannot provide an explanation for the sharp jump in wages under the NIRA.

An alternate hypothesis emphasizes the process of bargaining over the industry codes. The "hysteresis theory" notes that only employed workers got to bargain with employers over wages. If these "insiders" were concerned only about preserving their jobs, not in lowering their wages to employ more "outsiders," then they would have sought wages higher than the market-clearing level. In fact, the level of unemployment would not be relevant to their desires. The hysteresis theory therefore removes the paradox of rising wages in the presence of high unemployment by asserting that the former was not a function of the latter. It also provides an economic interpretation of the process of wage bargaining under the National Recovery Administration.

As a short-run measure, the NIRA was a failure. What it gave by improving expectations, it took away by raising nominal and real wages. The net effect was to restrict rather than to expand employment. As a long-run measure, however, the NIRA led to a substantial improvement in the conditions of labor.

The NIRA prevented employers from interfering with organizations of labor and collective bargaining. New unions were formed and grew in this receptive atmosphere. But the NIRA itself did not last long. In the "sick-chicken case" of 1935, the Supreme Court ruled that the NIRA was an unlawful delegation of legislative power to the NRA and an unlawful extension of federal power into activities within states. The NRA was dissolved, but the labor provisions of the NIRA were not forgotten. Senator Wagner introduced the National Labor Relations Act of 1935 that reestablished the rights of labor

under the NIRA. This narrower bill was upheld by the Supreme Court, and the National Labor Relations Board still oversees union activity and wage bargaining today. The law placed strong restrictions on the means used by employers to fight unions with the result that unionization of the labor force increased rapidly. At the peak of unionization, around 1950, fully one-third of the non-agricultural labor force belonged to unions.

The New Deal did not restrict its attention to industry. Farmers had been complaining about poor farm prices even before the Depression, and Roosevelt actually turned his attention to agriculture before industry. The Agricultural Adjustment Act was passed in May, 1933, before the NIRA. The philosophy of the two acts was the same. The AAA allowed the government to control production of agricultural commodities. By restricting production, policy makers hoped to increase the price.

Farmers could agree with the government to restrict production and be compensated for the land left unplanted. The payments were made from a processing tax that was in turn paid out of the difference between the current price of a commodity and the price resulting from lower production. The tax therefore was designed to be a redistributive one within agriculture; it was to be collected from farmers in proportion to the amount they marketed and paid out to farmers in proportion to the amount they did not market. The program's overall goal was to raise agricultural prices to a level that would provide the same purchasing power in 1933 that they had done before the World War, in 1914. The prewar conditions were adopted as "parity," against

which all current arrangements were judged.

The AAA got off to a slow start because the act was passed after many crops had been planted. The government contracted with cotton growers to destroy part of their crop, but prices did not rise as far as desired. Subsidies for destroying the crop should have been--but weren't--paid before the processing tax was collected. Farmers decided that the government was more interested in industry than in agriculture, particularly as the NRA approved higher prices for goods farmers bought.

Farm unrest was increased when the Supreme Court ruled that the AAA was unconstitutional at the start of 1935. As with the NIRA, the Court ruled that the federal government had trespassed on areas reserved to the states. And as with industry, Congress moved rapidly to salvage what it could of the AAA. The task was harder or Congress was more ambitious, because it was not until 1938 that a satisfactory replacement for the AAA was passed. The new law set up granaries to protect against drought and to allow the government to control prices through its inventory policies. The law also mandated support programs for specified crops and provided for acreage allotments and marketing quotas to be used as the means to this end.

The AAA and its successor programs did not do much to alleviate the agricultural depression in the mid-1930s. They did, however, create the framework for farm supports after the Second World War. The government attempted to raise agricultural prices by limiting production. But acreage limitations led to increased production per acre rather than lower production. The

government accumulated surpluses as it attempted to restrict the flow of agricultural products to the market.

6. THE SECOND NEW DEAL

The recovery in the 1930s has a dual aspect. Measured in terms of income growth, it is very impressive. GNP rose by one-third from 1933 to 1937. But measured by the reduction of unemployment, it was an anemic recovery. Unemployment remained well above ten percent throughout the 1930s. This is true even if workers employed by the government on various relief projects are counted as employed. Since these jobs were not paid market wages, traditional analysis views the workers holding them as unemployed. But since these workers were not idle, others have argued that they should be considered employed, albeit at a low wage.

If workers were willing to take jobs at these wages, then why didn't market wages fall to this level? As noted above, "hysteresis" in wage setting can prevent the real wage from falling enough to restore full employment. Wages at private firms were set to preserve the jobs of those people already employed, not to move others out of unemployment. The government promoted bargaining between associations of employed workers and their employers. It did not require unions to think about potential members who might be employed if wages were lower. Unions appear to have set their goals in terms of their actual members, that is, in terms of workers employed at the time of the bargain. There was as a result no force lowering wages to clear the labor market.

The involvement of government in banking, industry, agriculture and wage setting reveals the New Deal as a socialist policy regime. The New Deal was not national socialism or communism, but it did try to manage the economy directly in order to promote recovery. Instead of promoting a Keynesian expansion--the government refused to increase its deficits--the New Deal injected government into the management of economic activity. It was the precursor of postwar democratic socialism.

The primary aim of this socialist policy was economic recovery. Another aim was the distribution of income to everyone in the economy. If wages were set low enough to provide full employment, then the redistributive impulse could be subsumed under the goal of employing all workers. But if the government set wages higher than this, if it accepted or encouraged wage setting to benefit the already employed, then the redistributive goal of socialism had to be solved by different means.

The Second New Deal of 1935 was Roosevelt's response to this challenge. Turning from measures to revive the economy, Roosevelt extended the government's control over the economy to spread its output more evenly. The organization of labor under the NRA was institutionalized by the National Labor Relations (Wagner) Act and the creation of the National Labor Relations Board when the NIRA was declared unconstitutional. This board was only one of the many regulatory bodies established to oversee and control the economy. Utilities, in particular, were subject to regulation on a new scale.

Various measures--rural electrification, a moratorium on

farm foreclosures--extended the government's helping hand into the countryside. The Social Security Act initiated a program that would end up with the government supporting directly a major part of the population. Unable to pass legislation offering aid to the poor, the program's proponents seized on aid to the elderly as a way of getting the socialist camel's nose into the policy tent.

Once started, Social Security was expanded over the years to include more and more of the population. It has become a major way in which intergenerational transfers of income are made in America. Even though the Social Security system was set up along the lines of private insurance, the actual payments are made from contemporaneous taxes, not from an accumulating individual balance. The result was a windfall gain for the first generation covered by Social Security, that is, the generation that was working during the Depression and receiving benefits soon after World War Two.

Modern drug regulation in the United States also dates from the late 1930s. One of the last acts of the Second New Deal greatly expanded the federal Food and Drug Administration's (FDA) powers. The act was hardly the result of an organized plan to reform medical care; it was only passed at all because of a tragedy that killed a hundred people. Despite this weak beginning, drug regulation has been extended and strengthened in the postwar period to substitute administrative decisions by the FDA for the actions of the private market.

The recovery from the Depression was neither smooth nor

complete during the 1930s. The lack of full recovery has been discussed; it is now time to examine the recovery that did take place. It was rapid by historical standards, although not rapid enough to lead to full employment. What accounted for the rapidity of economic growth from 1933 to 1937?

Fiscal policy deserves none of the credit. The government budget changed from year to year, but the cumulative impact was virtually nil. Fiscal policy did not work in the 1930s because it was not tried. Despite the vast increase in government activity during the New Deal that changed forever the role of the federal government in economic life, the government deficit did not rise. It consequently could not have an expansionary effect on the economy.

Monetary policy deserves no more credit. The Fed was reformed, but it remained as passive after 1933 as it had been before. Monetary expansion, as distinct from monetary policy, was nonetheless critical to the recovery. The monetary base (High powered money) grew extremely rapidly after 1933 as European gold fled to America. The Fed did not sterilize this inflow as it had sterilized the inflow in the 1920s. The result was an extremely high rate of growth of the money supply.

It has been a commonplace of macroeconomics that this expansion did not affect the recovery. You "cannot push on a string," and monetary policy cannot work when interest rates are very low. This traditional view may well be wrong; it ignores the difference between nominal and real interest rates. Real interest rates were high during the later stages of the deflation

as people expected the deflation to continue. Roosevelt's devaluation and the NIRA, in fact, the whole New Deal, changed the course of prices and with them people's expectations. Real interest rates fell, and spending on consumer durables and investment rose. To the extent that monetary expansion was inflationary, the anticipated inflation also reduced real interest rates. Monetary expansion was a factor in the recovery.

It must be emphasized that the policies of the New Deal did not always support each other. For example, the NIRA raised prices and wages at the same time that the money supply was beginning its expansion. If we ignore expectations and look at the Keynes effect, then the policies were in conflict. The NRA codes channeled the increasing monetary ease into a rise in prices instead of a rise in production. If we look at the Mundell effect, the two policies seem to be working together. But there is another problem. For if the NIRA changed expectations and lowered real interest rates, then the monetary expansion was not as important as it looks by itself. And if it was the monetary expansion that lowered interest rates, then the NIRA had little positive effect. The evaluation of these policy combinations therefore depends on precise research on expectations.

After the rapid recovery in 1933-37, the economy experienced a renewed although short contraction. The 1937 recession was clearly caused by government policies. The high-employment government surplus, that is, the expenditures minus the taxes that would have been collected at high employment, rose

dramatically in 1937. A large veterans' bonus had been paid in 1936, echoing one paid in 1931, and the surplus rose after the payments were concluded. There was a fiscal contraction.

At the same time, the Federal Reserve became alarmed at the amount of excess reserves in banks. The Fed thought it was losing control over monetary policy since the banks had such a large cushion to fall back on in times of trouble. In order to mop up these excess reserves, the Fed doubled the reserve requirements in 1936. No macroeconomic effect of this policy was expected, since only excess reserves would diminish. But banks were not indifferent to the size of their excess reserves; they contracted to rebuild them in the uncertain economic environment. There was a monetary contraction.

Historians have disputed which policy was more effective, with the current laurels going to the monetary contraction. But the division is less important than the dependence of the economy on government policy. As in the great contraction of the early 1930s, the government demonstrated its power to contract the economy yet again in 1937.

Unemployment rose sharply in 1938. The recession delayed the return of full employment for several years. The record of the 1930s looks so dismal partly because there was a reprise of the Depression in the late 1930s. This echo may show how little had been learned in the Depression; Keynes' General Theory was only published in 1936 and not accepted widely for many years thereafter. Or it may show that full recovery was not the primary aim of economic policy. The record of the 1930s clearly

shows the presence of multiple goals, from maintaining the external value of the dollar to distributing the fruits of recovery more widely.

The 1937 recession was both sharp and short. Production, which fell rapidly in 1938, recovered in 1939, and unemployment fell. The recovery after the recession was even faster than before. It absorbed the labor force that had remained idle during the 1930s. The Second World War clearly provided the demand to pump up the economy. But the expansion started well before the United States entered the war and even before American production was turned toward Hitler's defeat. A renewed gold inflow, stimulated by rapidly growing fears of Nazi aggression, caused the money supply to resume and even exceed its previous rate of growth. This monetary expansion provided the final push needed to get the United States out of the Great Depression.

Bibliographic Notes

The modern analysis of the causes of the Depression starts with Milton Friedman and Anna J. Schwartz, A Monetary History of the United States, 1867-1960 (Princeton: Princeton University Press, 1963). Their view was disputed by Charles P. Kindleberger, The World in Depression, 1929-39 (London: Allen Lane Penguin, First Edition, 1973) and Peter Temin, Did Monetary Forces Cause the Great Depression? (New York: Norton, 1976). The view expressed here was initiated by Barry Eichengreen and Jeffrey Sachs, "Competitive Devaluation in the Great Depression: A Theoretical Reassessment," Economic Letters 22: 67-71 (June, 1986). It is expounded more fully in Peter Temin, Lessons from the Great Depression (Cambridge: MIT Press, 1989) and Barry Eichengreen, Golden Fetters (New York: Oxford University Press, 1991).

The start of the recovery is analyzed in Peter Temin and Barrie Wigmore, "The End of One Big Deflation, 1933," Explorations in Economic History 27: 483-502 (October, 1990); banking reform, in Carter H. Golembe, "The Deposit Insurance Legislation of 1933: An examination of its antecedents and its purposes," Political Science Quarterly 76: 189-95 (June, 1960), Thomas F. Huertas, "An Economic Brief Against Glass-Steagall," Journal of Banking Research 15: 148-59 (1984) amd Eugene N. White, "Before the Glass-Steagall Act: An analysis of the instextment banking activities of national banks," Explorations in Economics History 23: 33-55 (January, 1986); the NIRA, in

Ellis Hawley, The New Deal and the Problem of Monopolies: A Study in Economic Ambivalence (Princeton: Princeton University Press, 1966) and Michael M. Weinstein, Recovery and Redistribution under the NIRA (Amsterdam: North Holland, 1980); the AAA, in Van L. Perkins, Crisis in Agriculture: The Adgricultural Adjustment Administration and the New Deal (Berkeley: University of California Press, 1969) and Theodore Saloutos, The American Farmer and the New Deal (Ames, Iowa: Iowa University Press, 1982).