German Monetary History in the First Half of the Twentieth Century

Robert L. Hetzel

At the end of 1998, the German Bundesbank turned over the administration of monetary policy to the European Central Bank (ECB). In the years between World War I and 1998, the Bundesbank had come to embody the modern central bank. What history did Germany traverse to make possible the creation of such an institution? And how does that history help us define a modern central bank?

Today, a central bank chooses one of two objectives. It may target either the exchange rate or domestic economic conditions, including the inflation rate. In either case, the central bank is the unique institution charged with controlling the chosen objective. Such control relies exclusively on the central bank's management of its own balance sheet. In particular, the central bank controls its liabilities (the monetary base) through its asset acquisition.[1]

Conversely, a country with a modern central bank does not rely on government intervention in specific markets to achieve either price-level or exchange-rate objectives. If the central bank targets the exchange rate, the country does not rely on exchange controls, multiple exchange rates, tariffs, quotas, or other

This history will be continued in a related article, to appear in a future issue of the *Economic Quarterly.* It will consider how Germany came to define stability of the mark in terms of its internal value (price level) rather than external value (exchange rate). It will also explain the reasons for the creation of the European Monetary System and its successor, the European Monetary Union, and review the political process that led to the creation of the ECB. The author gratefully acknowledges helpful comments from Michael Dotsey, Martin M. Fase, Andreas Hornstein, Thomas Humphrey, Jeffrey Lacker, Joachim Scheide, and Alexander Wolman. The views expressed in this article are those of the author and should not be attributed to the Federal Reserve Bank of Richmond or to the Federal Reserve System.

[1] Central banks exercise this control indirectly through use of an overnight bank rate as a policy variable. This rate is a market rate. The need to control money creation imposed by the need to control inflation disciplines the central bank to respect the role the interest rate plays as a price in the price system.

administrative measures. If the central bank targets the inflation rate, the country does not rely on wage and price controls, guideposts, antitrust actions, or special intervention into the wage and price-setting decisions of firms. A modern central bank does not in general allocate credit either through subsidized lending at the discount window or quotas on the credit that individual banks can extend.

This article reviews German monetary history in the first half of the twentieth century, employing the theme that the evolution of the concept of a modern central bank required popular support for a free market.[2] It summarizes three episodes: hyperinflation in the twenties, deflation in the early thirties, and the currency reform of 1948. Inflation and deflation accompanied the economic instability of the first and second episodes, respectively. In each case, free enterprise lost public support. The third episode inaugurated a period of economic and monetary stability, during which free enterprise again became acceptable.

1. HYPERINFLATION IN THE WEIMAR REPUBLIC

In 1913, total currency in Germany amounted to just 6 billion marks. In November 1923 in Berlin, a loaf of bread cost 428 billion marks and a kilogram of butter almost 6,000 billion marks. From the end of World War I until 1924, the price level rose almost one trillionfold.[3] The economic cause of this hyperinflation was the monetization of public and private debt by Germany's central bank, the Reichsbank.[4] The political cause lay in the inability of a fragile democracy to impose the taxes necessary to pay war reparations.[5]

Reparations and Budget Deficits

Germany entered World War I believing that the war would be like the Franco-Prussian War of 1870–1871 and that the government would be able to finance a short war by issuing bonds, which a defeated France would redeem in gold (Marsh 1992, 77). In fact, the combatants devoted half of their economic output to the fighting. The central government in Germany, which did not impose income taxes, financed the war almost completely by issuing debt.

---

[2] An intellectually kindred exercise is Humphrey (1998). He traces the historical relationship between proponents of the quantity theory and free markets on one hand and anti-quantity theorists and mercantilists on the other. The quantity theory explains how governments can control the price level or trade balance without direct intervention in markets.

[3] The figures are from Webb (1989, 3) and Bresciani-Turroni (1937, 25).

[4] Among others, Bresciani-Turoni (1937), Keynes (1923), Cagan (1956), and Webb (1989) present this quantity theory view.

[5] Among others, Holtfrerich (1986) and Webb (1989) present this view. The earliest criticism of the punitive character of the Versailles treaty by an economist is Keynes (1919).

With the deficits that followed the end of the war, the Reich's debt amounted to half of national wealth. Interest on the debt amounted to four times the Reich's 1913 revenues.[6]

At Versailles, the victorious Allies imposed a punitive settlement on Germany. They stripped Germany of its colonies and Alsace-Lorraine. The Versailles treaty required that Germany pay for the damages caused by the war without stipulating an upper limit. France in particular demanded heavy reparations, embittered by the appalling human cost of retaking Alsace-Lorraine.

In May 1921, in the London Ultimatum, the Allies set an aggregate amount for reparations of 132 billion gold marks. However, the Ultimatum allowed the Reparations Commission to demand interest on the unpaid amount when it judged that German finances had recovered. Uncertainty about the total reparations payments and the disincentive to run fiscal surpluses that uncertainty created for Germany probably weighed even more heavily than the huge magnitude of the total. Foreign lenders then found it difficult to assess Germany's credit worthiness (Holtfrerich 1986, 143, 145, 154).

Because of differing valuations placed on payments in kind, it is difficult to measure the reparations Germany actually paid. Holtfrerich (1986, 151) compares various estimates and concludes that for the years 1919 through 1922, Germany paid 10 percent of its national income in reparations. Webb (1989, 106) arrives at a similar number, which amounted to 80 percent of Germany's exports. Holtfrerich (1986, 153) points out that, as a fraction of national income, reparations equalled the amount of government expenditure at all levels in the prewar period.

Holtfrerich (1986, 153) argues that Germany could not have raised through direct taxation the amounts necessary "to effect a foreign transfer regarded from the outset as beyond fulfillment, unjust and indeed morally reprehensible by almost the entire population." He explains the resort to an inflation tax by quoting Friedrich Bendixen, a Hamburg bank director:

> Only in taxation do people discern the arbitrary incursions of the state; the movement of prices, on the other hand, seems to them sometimes the outcome of traders' sordid machinations, more often a dispensation which, like frost and hail, mankind must simply accept. The statesman's opportunity lies in appreciating this mental disposition. (153)

Unable to cover its expenditures through explicit taxes, the German government ran deficits exceeding 50 percent of its expenditures from 1919 through 1923 (Holtfrerich 1986, 173). Reichsbank purchases of government debt made the printing press the ultimate source for funding these deficits.[7]

---

[6] These figures are from Holtfrerich (1986, 102, 109, 126).

[7] Keynes (1923, Chapter 2) explains inflation as a tax.

Holtfrerich (1986, 152) reproduces figures of Arnd Jessen showing that, as a proportion of government expenditures, the yield of the inflation tax also amounted to about 50 percent. In the years 1919, 1920, 1921, and 1922, revenue from the inflation tax respectively amounted to 62, 53, 43, and 43 percent of government revenue.

A Chronology of Inflation

Although money creation ultimately caused German inflation, expectations about the ability of the government to achieve ultimate fiscal balance determined inflation's timing (Webb 1989). The foreign and domestic public willingly purchased new debt issues when it believed that the government could run future surpluses to offset contemporaneous deficits. When it did not, the debt presented to the Reichsbank rose. Foreign speculative capital inflows ceased, the exchange rate depreciated, and inflation rose.

Inflation exacerbated the government deficit by reducing tax revenue. Because the government levied taxes in nominal terms, the elapsed time between assessment and collection destroyed their real value (Bresciani-Turroni 1937, 66; Sargent 1993, 69). Figure 1 reveals the pattern of an inflation driven by fiscal fears. It shows the wholesale price index with the periods demarcated as by Webb (1989, 5).

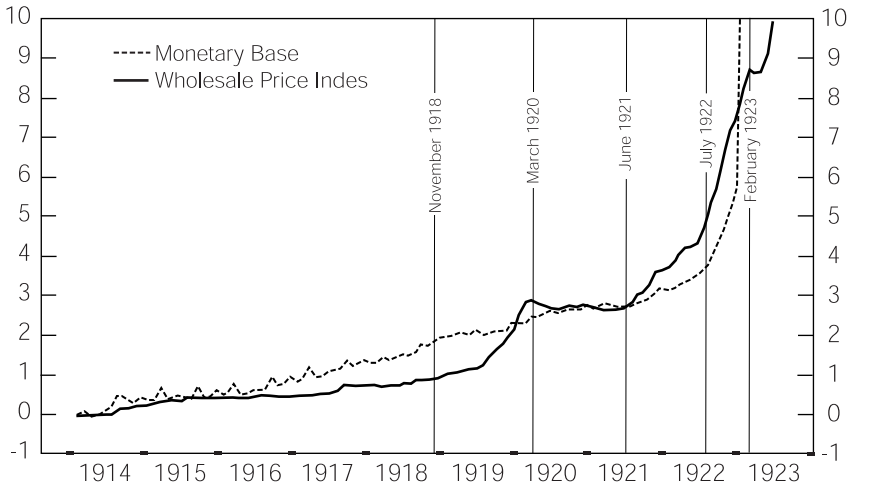
At the end of the war, the fiscal situation was gloomy. In November 1918, a worker revolution overthrew the Kaisertum. Widespread, paralyzing strikes occurred. The new government issued debt to obtain the food supplies and funds for demobilization that were necessary to limit social unrest (Marsh 1992, 79). The government subsidized food, coal, and employment in the railway and postal systems to prevent a second worker revolt.

The new government's concern was valid, for it was indeed possible that fiscal restraint would raise unemployment and set off a Bolshevik revolution (Ferguson and Granville 2000, 1063–65; James 1999, 18). The price level rose from the end of the war until early 1920. By February 1920, the wholesale price level had risen to 17 times its 1913 level. According to Webb (1989, 52), this rise reduced the real value of the nominal government debt outstanding in October 1919 (172 billion marks) to a value consistent with future budget balance. The political situation stabilized in March 1920 with the failure of the Kapp Putsch.

The price level steadied after March 1920.[8] Minister of Finance Matthias Erzberger implemented tax reforms that brought Germany close to achieving future budget balance. Although the budget remained in deficit, tax revenues

[8] Regarding this paragraph and the next, see Webb (1989, 54–60) and Holtfrerich (1986, 301–11).

Figure 1 Monetary Base and Wholesale Price Index



Notes: Data normalized with 1913 equal to 1. Observations are the natural logarithm. The figure uses the data in Diagram 4 in Holtfrerich (1986). The monetary base is cash in circulation plus commercial bank deposits at the Reichsbank.

were rising steadily. Given stable real expenditures, growth in the economy would have increased revenues and balanced the budget. These calculations depended upon maintenance of the current level of reparations, which amounted to 2.24 billion marks in the 12-month period ending 21 June 1920.

However, in the London Schedule of May 1921, the Allies threatened to occupy the Ruhr unless Germany transferred 4 billion marks annually and made additional payments as its economy grew. The Reichstag refused to impose additional taxes, and inflation rose when the prospect of ultimate budget balance receded. As inflation rose, collection lags in the tax system reduced real revenues. In October 1921, the Allies further weakened the political standing of the German government by annexing Upper Silesia to Poland.

Only the United States was in a position to broker a compromise, for it could have forgiven war debts owed it by France and Britain in return for moderation of their reparations demands. But the United States retreated into isolationism. Contradictorily, Allied governments made it hard for Germany to run the surplus on its external trade account that was necessary to pay reparations by imposing high duties on its exports. Many Germans adopted the fatalistic attitude that German economic ruin would be necessary to demonstrate the injustice of reparations. They contended that only when the Allies

scaled back demands for reparations could Germany bring order to its domestic finances.

To support their case, they argued that the reparations caused inflation: The purchase of foreign exchange to make reparations payments depreciated the mark.[9] In turn, the depreciation of the mark raised internal prices. An end to the monetization of government debt without a settlement of the reparations issue would leave unchecked both the depreciation of the mark and domestic inflation. Because the government would still have to maintain subsidies, it would become bankrupt. Rudolf von Havenstein, the head of the Reichsbank, said, "So long as the reparations burden remains, there is no other means to procure the necessary means for the Reich than the discounting of Reich Treasury notes at the Reichsbank" (Feldman 1997, 445).[10]

Inflation became hyperinflation with the assassination of foreign minister Walther Rathenau in July 1922 by right-wing reactionaries. Capital flight spurred mark depreciation, which exacerbated domestic price increases. The resulting fall in the purchasing power of the mark created a liquidity crisis. To deal with this crisis, industrialists argued for the rediscounting of bills of exchange at the Reichsbank. Georg Bernhard, an influential newspaper editor and member of the Reich Economic Council, argued that "There is only one source of money in Germany left, the Reichsbank. . . . It is thus absolutely correct that we create commercial bills; then the Reichsbank can issue money" (Feldman 1927, 450). In the second half of 1922, the Reichsbank began to discount significant numbers of private bills (Webb 1989, 28).

Karl Helfferich, finance minister during World War I, said that if the Reichsbank ceased "the printing of notes. . . all national and economic life would be stopped."[11] Hjalmar Schacht (1927), later the Reichsbank president, wrote:

> In 1923 there were engaged on the production of notes for the Reichsbank. . . 1,783 machines. . . . [E]ven with assistance on so vast a scale the

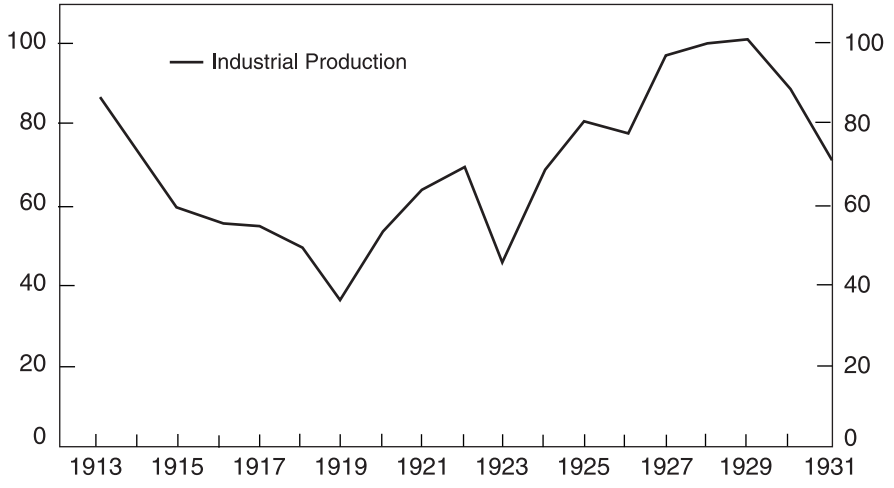
[9] See references in Bresciani-Turroni (1937, Chapter II), especially pages 77 to 82.

[10] The argument is a mixture of economic fallacy and political insight. As a matter of positive economics, a depreciation of the mark on the foreign exchanges could not in and of itself produce a significant rise in the domestic price level. A rise in the price level would have reduced the real money holdings of the German public. The attempt by the German public to restore its monetary purchasing power through a reduction in expenditure would have reversed the rise in the domestic price level (apart perhaps from an amount relecting a reduction in desired real money holdings as a consequence of feeling poorer due to the mark depreciation).

From a political perspective, however, without an end to the reparations, the problems of achieving fiscal balance were, if not insupurable, at least extremely difficult. If the Reichsbank had not monetized debt, the outstanding debt would have grown more rapidly. The effective default on existing debt that came through inflation might then have occurred sooner and explicitly through an actual default. In this sense, the reparations forced the Reichsbank into a situation where inflationary finance appeared to be the only politically feasible option.

[11] Cited in Bresciani-Turroni (1937, 81).

Figure 2 Industrial Production



Notes: Data normalized with 1928 equal to 100. Observations are the natural logarithm. The figure reproduces Diagram 4 in Holtfrerich (1986).

> Bank was not in a position to supply the business world with a sufficiency of notes. (105)

In fact, the high rate of inflation made holding money extraordinarily costly. German moneyholders responded rationally by reducing the real amount of marks they held.[12] The minimal real purchasing power of marks made money appear scarce. That scarcity rationalized the demands of the industrial class for the Reichsbank to continue discounting its commercial paper at a discount rate of 6 percent. In this way, industrialists obtained costlessly the revenues from the taxing powers (seigniorage) that money creation granted to a central bank.

France never confronted the inherent contradictions in its policy toward Germany. It wanted a weak German economy incapable of supporting remilitarization, and it wanted the payment of reparations, which required a strong German economy. On 11 January 1923, France occupied the Ruhr when Germany failed to make in kind deliveries of coal. Germany responded with a policy of passive resistance. With government support, workers in the Ruhr went on strike to prevent France from obtaining the region's coal and steel.

[12] They achieved the reduction in the purchasing power of the mark by increasing the price level in excess of the amount of paper marks issued.

Without coal, German railroads could not run, and without railroads the German economy could not run. Figure 2 shows the decline in output precipitated by the Ruhr occupation. Germany had to import coal. The government also had to pay striking workers, in part, it believed, to prevent them from joining the Communist movement and starting a Bolshevik revolution. As the government deficit widened to 22 percent of net national product, the money stock soared (Ferguson and Granville 2000, 1068).[13] By year-end 1922, the mark-dollar exchange rate had fallen from its prewar level of 4.2 to 1 to 1,500 to 1. By the end of November 1923, it had fallen to 4,200,000,000,000 to 1.

End of the Hyperinflation

On 20 November 1923, Germany ended inflation by pegging the mark's foreign exchange value at its prevailing value of 4,200 billion marks to the dollar.[14] What made monetary reform credible to the German public?

As background, note a characteristic of twentieth-century monetary reforms that ended hyperinflations (for example, in Argentina in 1991).[15] At the time of a reform, the economy in question is using a stable currency as the standard of value, often the dollar, rather than the domestic currency. In hyperinflations, individuals set prices in dollars and then use the dollar exchange rate to convert prices to domestic currency equivalents. The domestic money serves only as a token currency for small transactions. One prerequisite for a successful monetary reform is to make credible the maintenance of a fixed dollar exchange rate, which will reestablish individuals' faith in the purchasing power of domestic currency.

In 1923, "German society was moving massively to disown the paper mark" (Feldman 1997, 691). The German economy largely indexed ("valorized") transactions to maintain their real value. One example was a bank proposed by Hans Luther and Rudolf Hilferding (Food and Finance Ministers, respectively) that would issue a "rye" mark, that is, a deposit redeemable in rye. The bank came into existence on 15 November 1923 as the Rentenbank, but with a rentenmark deposit convertible to gold at the prewar value. The

[13] Webb (1989, Chapters 2 and 3) provides a careful account of the interaction between the fiscal system and inflation. Inflation reduced the revenue that government raised because it vitiated the real value of taxes, which the government set in nominal terms but collected with a delay. If Germany had had a period of price stability or if it had indexed its tax system, it could have balanced its budget. (The wealthy opposed indexing because it would have prevented inflation from reducing the real value of their tax liability.) In the absence of inflation, the tax reforms the government passed might have raised sufficient revenue to cover domestic expenditure and reparations payments. However, the government faced an incessant set of domestic and external pressures (such as French demands to pay immediate cash reparations). As a result, it never had an interval of time during which it could have stopped inflation by meeting its expenditures without printing money (Webb 1989, 42–43).

[14] For a brief account of the November 1923 monetary reform, see Humphrey (1980).

[15] For other examples, see Bruno et al. (1988).

new bank had strict limits on the credit it could extend to the private sector and to the government and met with "astounding success" (Holtfrerich 1986, 316).

The actual breakdown of German economic life came about because of interventions by the German government to maintain the paper mark as the medium of exchange. Holtfrerich (1986, 313) writes of hyperinflation Germany, "The economy had already largely turned over to a foreign, hard-currency standard. . . . The crisis arose out of the reluctance of the Reich to permit business to employ foreign means of payment in domestic transactions as desired; indeed the Reich could not permit the practice. . . as long as inflation remained as a 'tax' source."

Although the government ignored the price setting of the large industrial cartels, it imposed price controls on professionals and the retail trades to prevent "profiteering." Rent control destroyed the wealth of small property owners. Because farmers could not borrow using an indexed value of land as collateral, they could not obtain fertilizer (Feldman 1997, 681–84). "Because the farmers were already refusing to accept the currency. . . Germany faced the imminent danger of hunger revolts" (Feldman 1997, 693).

Bresciani-Turroni offers two observations helpful for understanding the German reform. First, "At the beginning of the inflation. . . the public still did not understand the phenomenon of monetary depreciation" (Bresciani-Turroni 1937, 430). However, by the end, the public associated inflation with the money issues of the Reichsbank. Everyone knew at that point that an end to inflation would require the Reichsbank to limit monetary emission to whatever was needed to maintain the new mark-dollar exchange rate. Second, hyperinflation threatened imminent economic and social collapse. Holtfrerich (1986, 312), citing Schacht, writes "Plundering and riots were a daily occurrence," and Bresciani-Turroni (1937, 336) cites Luther, "The effective starving of the towns and the impossibility of continuing economic activities on the basis of the paper mark was so obvious in the days preceding November 16th that a dissolution of the social order must have been expected almost from hour to hour."

The 1923 reform worked because there was political consensus that it had to work. The economic disruption produced by the combination of hyperinflation and government attempts to force continued use of the mark had pushed Germany to the edge of social disintegration. Revolts, including Hitler's beer-hall Putsch and attempted March on Berlin on 9 November 1923, challenged the survival of the government. A social and political consensus emerged that the Reichsbank had to maintain the dollar-mark exchange rate. Faced with chaos, Germany took steps to restore order.

In this changed environment, the Reichsbank ceased monetizing government debt. The amount of treasury bills held by the Reichsbank went from

190,000,000 trillion marks in mid-November to zero by year-end (Marsh 1992, 83). Finance minister Hans Luther balanced the budget through emergency tax decrees and budget cuts. Germany declared an end to the policy of passive resistance opposing the Ruhr occupation and ceased payments to striking workers in the Ruhr. Negotiations between Germany and the Allies began that led to the 1924 Dawes Plan rescheduling of reparations (Ferguson and Granville 2000, 1078; Yeager 1981, 57).

The final test that established the credibility of the Reichsbank occurred in April 1924. When the dollar value of the mark weakened, the Reichsbank drastically restricted credit. S. Parker Gilbert, Agent General for Reparations Payments, later cabled George Harrison, governor of the Federal Reserve Bank of New York:[16]

> [The Reichsbank's] policy resulted in [a] check to [the] increase of Reichsbank credit and circulation, development of excess of exports over imports, liquidation of heavy commodity stocks accumulated during inflation, decline in commodity prices, and large failures of many firms established during inflation. Rates for month[ly] money rose from 30 to 80 percent but later declined rapidly.

Discrediting Capitalism

Ferguson and Granville (2000, 1084) write, "By discrediting free markets, the rule of law, parliamentary institutions, and international economic openness, the Weimar inflation proved the perfect seedbed for national(ist) socialism." The Weimar inflation produced arbitrary redistributions of income that discredited the market economy. As a class, the wealthy were the main losers because they held most of the mark-denominated financial wealth. Pensioners, bondholders, and rentiers lost everything. Laws against indexation and profiteering hurt merchants. Workers who were protected by labor unions preserved their real wages.[17]

Keynes has described how inflation destroys the social foundation of a market economy:

> By a continuing process of inflation, governments can confiscate, secretly and unobserved, an important part of the wealth of their citizens. By this method they not only confiscate, but they confiscate arbitrarily; and,

[16] S. Parker Gilbert memo to George Harrison, "Reichsbank Credit Policy," 22 July 1931, Harrison Papers. The Harrison Papers are at the Federal Reserve Bank of New York and Columbia University. Gilbert refers to the description of the episode on page 162 of Schacht's book, *Stabilization of the Mark*.

[17] These generalizations come from Holtfrerich (1986, Chapter 8) and Webb (1989, Chapter 5).

> while the process impoverishes many, it actually enriches some. . . . Those to whom the system brings windfalls. . . become profiteers. (1919, 148–49)
>
> To convert the business man into a profiteer is to strike a blow at capitalism, because it destroys the psychological equilibrium which permits the perpetuance of unequal rewards. (1923, 24)
>
> Lenin was certainly right. There is no subtler, no surer means of overturning the existing basis of society than to debauch the currency. The process engages all the hidden forces of economic law on the side of destruction, and does it in a manner which not one man in a million is able to diagnose. . . . By combining a popular hatred of the class of entrepreneurs with the blow already given to social security by the violent and arbitrary disturbance of contract. . . governments are fast rendering impossible a continuance of the social and economic order of the nineteenth century. (1919, 149–50)

Popular resentment concentrated on speculators. When Germans sold their family heirlooms to survive, they blamed the middlemen who organized the sales. A fictional character called Raffke, created by cabaret songwriter Kart Tucholsky, embodied the culturally ignorant person made rich from profiteering (Feldman 1997, 553). Bresciani-Turroni (1937, 328) would later write of "the poverty of certain German classes during the inflation which contrasted with the foolish extravagance and provocative ostentation of inflation profiteers."

Nationalist resentment targeted all foreigners but most especially the minority within reach—Jews. Hitler railed against the "Jewification of the economy" (Feldman 1997, 780, 575). On 5 November 1923, the government raised the price of bread to 140 billion marks, and in response, crowds plundered stores and attacked Jews. In July 1922, the British Consul in Frankfurt wrote:

> [T]he educated classes, deprived, in a great majority of cases, of the right to live and bring up their families in decency, are becoming more and more hostile to the Republic and open in their adhesion to the forces of reaction. Coupled with this movement, a strong and virulent growth of anti-Semitism is manifest. (Feldman 1997, 449)

Germans wanted a world where wealth resulted from hard work, not financial transactions. According to Feldman (1997), Germans desired a return to a world where

> the public good should take precedence over private gain. . . . It was not only Hitler who appealed to these sentiments. . . . [I]nflation. . . caused the Republic to be identified with. . . violations of law, equity and good faith. . . . No less offensive. . . was the sense that there had been a misappropriation of spiritual values. (657–58)

2. WORLD DEPRESSION

Germany ended hyperinflation and restored social order with its commitment to the gold standard. The November 1923 stabilization program committed Germany to exchange 1,392 reichsmarks for a pound of gold. However, German economic stability then became dependent upon the stability of the international gold standard. Starting in 1928, the deflationary monetary policies of two of the largest adherents to the gold standard, France and the United States, forced deflation and economic depression on Germany. Short-run salvation led to longer-run doom. The following section explains the fragility of the reconstructed gold standard.

Reviving the International Gold Standard

In 1920, Britain legislated a return to the gold standard at the prewar parity to take effect at the end of a five-year period. Britain based its decision in part on the assumption that gold flows to the United States would raise price levels there and limit the domestic deflation needed to reestablish the prewar parity (Rothbard 1996, 8). In fact, the United States sterilized gold inflows to prevent a rise in domestic prices. In the 1920s, the Federal Reserve held almost twice the amount of gold required to back its note issue (Yeager 1976, 333). Britain then had to deflate to return to gold at the prewar parity.[18]

After the war, France had counted unrealistically on German reparations to balance its budget. When they did not materialize, it used inflation as a tax to finance expenditures. In 1926, France pulled back from the brink of hyperinflation. Unlike Britain, in France inflation had put the old parity hopelessly out of reach. As a consequence, France returned to gold at a parity that undervalued the franc. Scarred by its experience with inflation, France sterilized gold inflows to prevent a rise in prices.

Allied war debts and reparations added to the inherent fragility of an international gold standard programmed for deflation. They required the transfer of resources from Germany to France and England and then from these countries to the United States. To accomplish these transfers, Germany would have had to run a trade surplus toward France and Britain. In turn, France and Britain would have had to run a trade surplus toward the United States. In the protectionist environment of the 1920s, that trade pattern was politically unacceptable. Only capital outflows from the United States made the system work (Yeager 1976, 333; Holtfrerich 1989, 151).[19]

---

[18] The prewar dollar-pound exchange rate was 4.86. After the war, in November 1920, the value of the pound fell to a low of 3.44. The British price level then had to fall commensurately to validate the former gold standard exchange rate of 4.86.

[19] Eichengreen (1995, 224) calculated that most of the $2 billion in reparations paid by Germany between 1924 and 1929 went to the United States for payment of Allied war debts.

World Deflation

By the end of 1927, it appeared that Europe had successfully returned to the gold standard. However, in 1928, the Federal Reserve initiated a restrictive monetary policy to stop stock market speculation. In 1920 and 1921, a floating exchange rate had insulated Germany from deflationary U.S. monetary policy. In those years, German industrial production rose 46 and 20 percent, respectively. In contrast, in Britain, whose commitment to return to the gold standard at the prewar parity overvalued its exchange rate, industrial production fell 32 percent in 1920.[20] At the end of the decade, a revived international gold standard transmitted U.S. deflation to Germany.

In the 1920s, capital had flowed into Germany. That is, Germany exported not only goods, but also IOUs. When the Federal Reserve System began raising interest rates in 1928, those capital inflows lessened. Germany then could not use funds gained from capital inflows to pay its reparations, but instead had to run a trade surplus to gain the needed funds. The price level in Germany had to fall to make its exports more attractive to the rest of the world. By the last half of 1929, foreign debt issued in New York was less than a third of its 1927 level (Chandler 1958, 456). "Net portfolio lending by the United States declined from more than $1000 million in 1927 to less than $700 million in 1928 and turned negative in 1929" (Eichengreen 1995, 226).[21]

As a result, in 1928 U.S. financial markets began attracting gold from Europe.[22] Foreign central banks had to raise their domestic interest rates to offset gold losses. The *Federal Reserve Bulletin* (November 1930, 655) talked about "Money rates abroad, which had been kept up largely to protect the reserves of foreign countries against the attraction of speculative and high-money conditions in the United States." George Harrison, governor of the Federal Reserve Bank of New York, informed Secretary of the Treasury Andrew Mellon that "our high money rates... continue to act as a pressure upon all the European bank reserves."[23] At the same time, France, with its undervalued franc, also absorbed gold from the rest of the world. In 1928 and the first half of 1929, France absorbed 3 percent of global gold reserves (Eichengreen 1995, 216).

America then returned the funds to Germany through capital flows. Not until 1929 did reparations payments exceed those capital flows.

[20] The figures are from Holtfrerich (1986, Table 38).

[21] Eichengreen (1995, 226) also documents the corresponding decline in German bond flotations in New York.

[22] From August through December 1928, the United States imported $44.5 million in gold; in 1929, $120 million; in 1930, $278 million; and from January through September 1931, $336 million (valued at $20.67 per fine ounce). Data are from Table 156, "Analysis of Changes in Gold Stock of United States, Monthly, 1914–1941," Board of Governors (1976). Also in 1929, Germany's scheduled reparations payments rose.

[23] Harrison memo, "Conversation over the telephone with Secretary Mellon," 29 April 1929, Harrison Papers.

To reverse their gold outflows, other countries had to run a trade surplus with the United States and France. Because the Fed and the Banque de France sterilized gold inflows, those other countries had to achieve trade surpluses through deflation. That is, to make their goods cheaper on international markets, their price levels had to fall. By creating obstacles to trade, protectionism exacerbated the extent of the required deflation.[24] The gold standard became an engine of worldwide deflation. The most visible signs of the stress of deflation were financial panic and widespread bank failures as depositors withdrew gold and currency from banks.

Both international and domestic considerations compelled Germany to deflate rather than abandon the gold standard. The Banking Act, created in 1924 as part of the Dawes reparations plan, required Germany to back its currency with gold and foreign exchange reserves equal to 40 percent of its currency. A foreign member of the General Council, set up to oversee the Reichsbank, could stop note issue if he believed gold convertibility was threatened (James 1999, 25).

The foreign loans from the Morgan syndicate required Germany to stay on gold. Those loans allowed Germany to finance its reparations payments. Germany also believed that adherence to the gold standard would provide it with a reputation for financial conservatism that would make credible its efforts to renegotiate reparations obligations. Finally, and most important, seared by the memory of hyperinflation, German public opinion supported the gold standard. Germans associated abandonment of the gold standard with inflation (Bresciani-Turroni 1937, 402; Eichengreen 1995, 270; James 1999, 25). Politicians had difficulty supporting a possibly inflationary policy without appearing to favor large industry and agriculture, which as debtors had profited from the earlier inflation (Feldman 1997, 853).

The 1931 German Financial Panic

As noted above, the Reichsbank had to maintain at least a 40 percent gold cover, that is, a gold-reserves-to-note circulation ratio of 40 percent. In April 1929, with the near collapse of reparations negotiations in Paris, reserve outflows threatened the gold cover. In May 1929, the successful resumption of the negotiations reestablished calm and led to the Young Plan, signed in June. In September 1930, when elections gave Hitler's party the second largest majority in the Reichstag, gold outflows from the Reichsbank resumed; aid from an international consortium relieved the crisis. In May 1931, the gold cover had risen to a comfortable 60 percent.[25]

[24] The U.S. Congress passed the Smoot-Hawley tariff on 13 June 1930.

[25] This paragraph summarizes Goedde (2000, 16–17).

Any lasting restoration of investor confidence in the mark's gold parity would require settlement of the reparations issue.[26] Investors worried whether Germany could finance reparations payments in the absence of continued capital inflows. They also worried about the political instability within Germany caused by reparations. Right-wing parties demanded that Germany renounce reparations. The Bruening government had been unable to form a parliamentary majority since July 1930. Chancellor Bruening dismissed the Reichstag and governed without it.

At the beginning of June 1931, the Reichsbank again began to lose gold.[27] This worsening in the German balance of payments ultimately occurred at this time because the deflationary pressure of U.S. monetary policy intensified. The U.S. money stock M2 had declined only 2 percent in 1930. In 1931Q1 and 1931Q2, it fell at an annualized rate of 6.3 and 6.7 percent, respectively. In 1931Q3, M2 declined at an annualized rate of 11.3 percent (Friedman and Schwartz 1963).[28]

Bank failures in Austria and Hungary were the immediate source of the financial panic in Germany. The first of these was the failure of the Austrian Credit-Anstalt bank in mid-May 1931, after which Austria suspended gold convertibility and allowed its currency to depreciate. Because foreigners held half of German bank deposits, financial stability required that foreign investors retain confidence in the maintenance by Germany of its gold parity (Eichengreen 1995, 272). Only American leadership could have achieved that result.

On 20 June 1931, President Hoover proposed a one-year moratorium on reparations and Allied debt payments. Financial markets worldwide responded positively (Eichengreen 1995, 277). However, French ill will toward Germany delayed the negotiations until a debt moratorium was too late to help. French reluctance to agree to the Hoover moratorium made investors nervous. That nervousness caused the financial panics in Austria and Hungary to jump the border to Germany.[29]

---

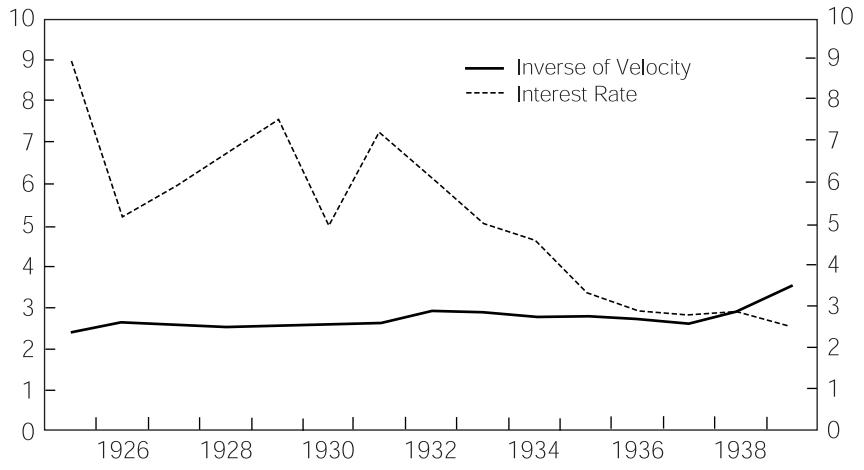
[26] See Norman telephone conversation with Harrison, 18 June 1931, Harrison Papers. See also Shepard Morgan, "Memorandum on German Short-Dated Debt Reduction," 6 July 1931, Harrison Papers.

[27] Eichengreen (1995, Chapter 9) contains a chronological account of the 1931 financial crisis.

[28] An alternative view is that the 1931 financial crisis originated as a banking crisis. See references in Balderston (1994). According to this view, depositors came to see German banks as insolvent. When they withdrew their funds, a foreign exchange crisis developed. On the basis of an examination of deposit withdrawals at German banks, Balderston (1994) disputes this view. (See also Goedde [2000, 22].) He argues that the banking crisis emerged from a foreign exchange crisis. In his view, foreign depositors withdrew funds from German banks as political developments made a negotiated settlement of reparations less likely. Pontzen (1999, 79) backs Balderston's view. He points out that capital outflows abroad set in before withdrawals by domestic depositors. He also points to foreign investor nervousness caused by a more aggressive stance toward reparations by the German government.

[29] Cablegram Dreyse to Harrison, 11 July 1931, Harrison Papers.

Figure 3 The Inverse of Velocity and the Interest Rate 1925–1939



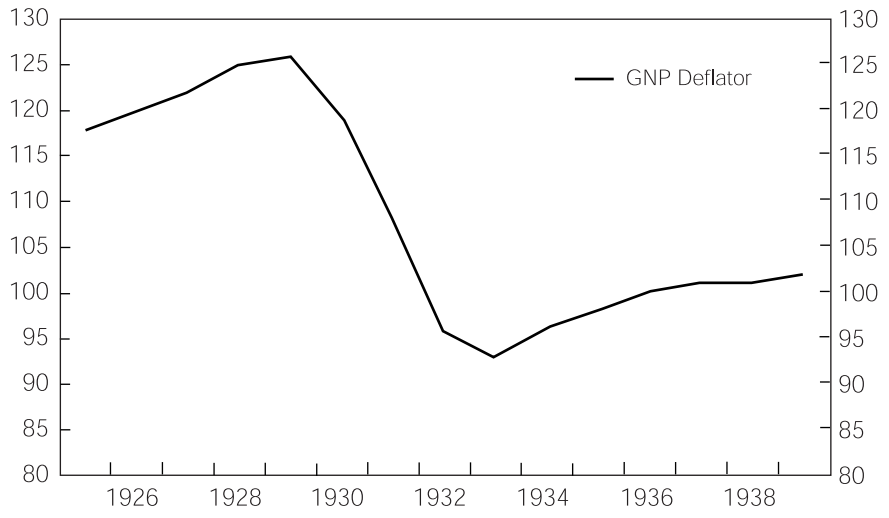
Notes: The interest rate is "Day-to-day money" in "Table 172—Money Rates in Selected Foreign Countries" in Board of Governors *Banking and Monetary Statistics* (1943). Velocity is nominal GNP (the product of Real GNP and the GNP Deflator) divided by money. See Table 2.

The appendix, "The Federal Reserve Bank of New York and the Reichsbank during the 1931 Crisis," provides an account of these events recorded in memos by Governor Harrison of the New York Fed. Central banks had cooperated to maintain the gold standard in the twenties. Walter Bagehot, a British economist, had expounded the lender-of-last-resort principle of lending freely at a high interest rate in a financial crisis (Bagehot 1873). These precedents suggested that the New York Fed, the Bank of England, and the Banque de France would lend to the Reichsbank. Had they done so, it is possible that monetary contraction and the fall in the German price level necessary to improve Germany's trade balance could have occurred in an orderly fashion. However, as a condition for a loan, the New York Fed required that the Bundesbank cease lending to commercial banks. The Reichsbank's compliance precipitated the collapse of its banking system, and the New York Fed then failed to come through with a loan.

Between 1926 and 1932, the Reichsbank maintained short-term interest rates at around 6 percent (Figure 3).[30] Figure 4 shows that significant deflation

[30] The initial high value of 9 percent in 1925 is a carryover from the very high rates needed in 1924 (45 percent in April 1924) to maintain the new gold parity for the mark (James 1999, 27).

Figure 4 GNP Deflator



Notes: See Table 2.

began in 1929 and continued through 1932. Despite this deflation, the Reichsbank kept market interest rates high. Table 1 shows the realized real rate of interest—the market rate minus actual inflation. This series rose moderately through 1929 and dramatically in the years 1930, 1931, and 1932.[31]

Monetary contraction allowed the Reichsbank to maintain high real rates of interest. Given the stability of monetary velocity (Figure 3), monetary contraction required a decline in nominal output (Figure 5).[32] Real output peaked in 1928 and then fell 16 percent through 1932. Only in 1933 did output begin a modest recovery (Figure 6 and Table 2).[33]

[31] The relevant series is the real rate of interest, which is the market rate of interest minus expected inflation. For the United States, Hamilton (1992, 172) estimates that in the Great Depression the public forecast about half of actual inflation. If the public in Germany forecast any significant fraction of the actual deflation, then real rates of interest were extraordinarily high in the years 1930, 1931, and 1932.

[32] Figures 3 and 5 show that after 1937 money rose with no corresponding rise in prices (velocity fell). This divergence is probably an artifact due to price controls. Official price indices did not measure the deterioration in the quality of goods (James 1999, 35).

[33] For the view that the Great Depression resulted from nonmonetary rather than monetary causes see Fisher and Hornstein (2001).

Table 1

Year	Overnight Rate of Interest	Inflation	Realized Real Interest Rate	Growth Rate of Real GNP	Money Growth
1926	5.2	1.8	3.4	2.8	15.1
1927	5.9	1.6	4.3	9.9	8.9
1928	6.7	2.5	4.2	4.4	4.3
1929	7.5	0.7	6.8	−0.4	1.5
1930	5.0	−5.4	10.4	−1.4	−6.1
1931	7.2	−9.3	16.5	−7.7	−15.3
1932	6.1	−11.2	17.3	−7.5	−9.7
1933	5.0	−3.0	8.0	6.3	2.0
1934	4.6	3.2	1.4	8.6	7.8

Notes: The overnight interest rate is an annual average of monthly figures of "Day-to-day money" in Table 172, "Money Rates in Selected Foreign Countries," Board of Governors (1976). The other series are from Deutsche Bundesbank, "Ausgewählte Daten zun Wirtschaftsentwicklung 1879–1974," Table 1.01, "Monetäre Entwicklung 1924–1974," and Table 1.02 "Allgemeine Wirtschaftsentwicklung 1924–1974" in *Geld und Bankwesen* 1876–1975, Frankfurt (1976). The inflation series is the GNP deflator series, "Preis-index des Sozial-produkts." Money is the sum of currency, "Bargeld umlauf," and bank demand and time deposits, "Sicht und Termin gelder." The realized real interest rate is the overnight rate of interest minus inflation.

Discrediting Capitalism

Given its adherence to the gold standard, Germany had to prevent capital outflows to avoid deflation and use fiscal authority to persuade international creditors of its credit worthiness. As German recession deepened, the government cut the social programs instituted after the war. The public therefore associated the market discipline of capitalism with government neglect of social needs.[34]

"The Reichsbank thus succeeded in cutting back the social programs started after the [1918] revolution. . . . The public perceived a dramatic failure of free market capitalism. They fell all the harder for the economic populism of the National Socialists" (Webb 1989, 127). Capitalism in the twentieth century did not appear to be a system that enriched the masses. Instead, it

[34] Compare the German financial crisis with that of Argentina in 2001. In Argentina, capital outflows and a currency board forced deflation. A street protester commented, "We want a government that represents the people, and not the capitalist model that is interested in money in the pockets of the few" ("Argentine Protests" 2001).

Table 2 German Historical Data

Year	Money*	Unemployment	Real GNP	GNP-Deflator	WPI	CPI	Unemployment Rate
1924		978			137.3	130.8	
1925	17106	636	59.7	117.9	141.8	141.8	
1926	19683	2010	61.4	120.0	134.4	142.1	
1927	21438	1327	67.5	121.9	137.6	147.9	
1928	22369	1391	70.5	125.0	140.0	151.7	6.7
1929	22694	1899	70.2	125.9	137.2	154.0	9.0
1930	21304	3076	69.2	119.1	124.6	148.1	14.6
1931	18042	4520	63.9	108.0	110.9	136.1	22.3
1932	16288	5575	59.1	95.9	96.5	120.6	28.1
1933	16608	4804	62.8	93.0	93.3	118.0	24.4
1934	17897	2718	68.2	96.0	98.4	121.1	13.8
1935	20001	2151	74.6	98.0	101.8	123.0	10.7
1936	21609	1593	81.2	100.0	104.1	124.5	7.6
1937	23309	912	90.0	101.0	105.9	125.1	4.2
1938	28490	429	99.2	101.0	105.7	125.6	1.9
1939	37910	119	107.2	102.0	106.9	126.2	0.5
1940	48640	52			110.0	130.1	0.2

*Equals the sum of currency and demand and time deposits.

Notes: Deutsche Bundesbank, ed. *Geld und Bankwesen 1876–1975* (1976). The unemployment rate is from Bundesarbeitsblatt 7–8 (1997), Bundesanstalt fur Arbeit, Bundesministerium fur Arbeit und Sozialordnung.

appeared to be a system that allowed the strong to exploit the weak. The depression discredited capitalism.

Harold James (1986) writes:

> It was the outbreak of the banking crisis in the summer of 1931 that made the German depression so severe.... [T]he collapse of the banks in central Europe had a major social, psychological and political impact. Capitalism appeared to have crashed with the banks, and this helped to discredit existing political systems. (283–84)

The years following the 1923 stabilization had offered the promise of a return to stability. The Young Plan for German reparations, adopted in principle at The Hague in August 1929, promised an end to reparations in 1988. Right-wing German parties rejected reparations and the "war guilt lie" they represented (Nicholls 1968, 137). Nevertheless, in early 1930 these parties, including the Nazi party, were marginal. Their marginal status changed with

the spread of the depression. Under the stress of the deflation that began in 1929, Germany could not keep together a political coalition capable of maintaining the democratic institutions of the Weimar Republic. Hitler was the clear beneficiary of the nationalist resentments revived by rising unemployment.[35] The unemployment rate averaged 28.1 percent in 1932 (Table 2). Hitler became chancellor of the government in January 1933.[36]

When Hitler came to power, the Reichsbank began to discount bills to finance his public works and rearmament program. In 1936, Hitler imposed a price freeze to control inflation. He said, "Inflation is a lack of discipline. . . . I'll see to it that prices remain stable. That's what my storm troopers are for" (Feldman 1997, 855). And again, "The first cause of the stability of the currency is the concentration camp" (James 1999, 35). Schacht, who became Reichsbank president for the second time in March 1933, maintained the mark price of gold by imposing foreign exchange controls and barter arrangements for foreign trade. "Germans who settled foreign debts directly with their creditors were threatened with the death penalty" (Pringle 1998, 71). Germans rejected the arbitrary redistribution of wealth produced by hyperinflation and the unemployment produced by deflation. With the discrediting of capitalism, they turned to monetary arrangements that required the detailed control of individual behavior by the state.

3. THE BIRTH OF THE D-MARK

Once capital controls effectively ended the gold standard, the Reichsbank was able to finance Germany's rearmament and war expenditures by printing money. Price controls created suppressed inflation. After the war ended, the Allied occupation forces maintained the price controls, and inflation continued to be suppressed.

The currency reform of 1948 ended price controls and introduced the deutsche mark (DM). The ensuing period of strong economic growth (the Wirtschaftswunder) and the resulting monetary stability contrasted with the

---

[35] Johnson (1998, 36) writes: "Thus Hitler and the Nazis were enabled in no small measure to seize power by demagogically exploiting popular—and populist—outrage at the banking system, the Depression, and capitalism in general. . . . Hitler's Nazis never enjoyed significant electoral support—not even after hyperinflation—until the onset of the Great Depression in 1929."

[36] I am not saying that a deflationary U.S. monetary policy propagated to the world by the gold standard "caused" the rise of Hitler. An extraordinary number of factors had to fall into place to make possible Hitler's ascendancy. Today, nothing can undo the horrors of the first half of the twentieth century. However, one can try to understand the near collapse of Western civilization and the near triumph of totalitarianism. After World War I, central banks did not understand their responsibility for control of the price level in a fiat money regime. That ignorance allowed the hyperinflations following World War I. Also, the newly created Federal Reserve System did not understand how its deviation from the rules of the international gold standard could create worldwide deflation and depression. An understanding of how central banks contributed to the economic instability that characterized the first half of the twentieth century helps explain what went wrong.

hyperinflation and subsequent deflation of the Weimar Republic. That contrast bred the presumption that economic and social stability required monetary stability, and this widespread presumption later created popular support for the "stability" policy of the Bundesbank.

Price Controls and Suppressed Inflation

Despite the devastation of World War II, Germany was in a position to recover economically after the war. Buchheim (1999) writes:

> The conditions for resuming production in western Germany were actually very good. Despite the war dead, the influx of refugees and expellees saw a sharp rise in the general population. . . . [M]any of the newcomers had been employed in industrial production occupations and had the appropriate skills. . . . [N]ew investments during the war had far exceeded the plant facilities destroyed. For this reason, West German industrial assets in 1945 were not only greater than before the war, but also more modern. (57–58)[37]

However, in 1947, two years after the end of the war, industrial production in Germany amounted to only 40 percent of its 1936 level. Price controls throttled German economic recovery. In other Western European countries, which also suffered war damage but did not have price controls, output exceeded the prewar level by 1947 (Yeager 1976, 388).

From 1936 through 1944, money (measured by currency in circulation plus total bank deposits) rose somewhat more than sixfold (Table 2). Despite this rise in money, price controls restrained the rise in the official consumer price index to only 14 percent from 1936 through 1944. Germany therefore ended the war with suppressed inflation. The Allies kept Hitler's price freeze in effect during the postwar occupation. Goods traded on the black market or through barter because no one wanted to exchange goods for marks at the artificially low price level.[38]

Germans used nylon stockings, American cigarettes, and Parker pens for currency. For example, in 1945, ten cigarettes could be exchanged for 1,500 grams of bread and two pairs of stockings for 1.5 pounds of butter (Haus der Geschichte).[39] American soldiers could obtain those items from the PX. Germans resented the privileges that this commodity money gave Americans. General Clay, the American military governor, responded to this resentment

[37] Streit (1999, 644) expresses the same view.

[38] Kaiserstrasse in Frankfurt was a center for the black market. Ironically, it is now the home of the ECB, the new symbol of a valued currency.

[39] References to Haus der Geschichte are to the exhibits in the museum of modern German history in Bonn.

by strictly enforcing the law against black market transactions (Smyser 1999, 44).

Allied enforcement of price controls paralyzed the German economy. For example, farmers would not bring crops to market. Children died of malnourishment and food riots broke out (Smyser 1999, 31–32). Controls discouraged work effort. With little effort, workers could earn enough marks to buy all the rations allowed them by their ration cards. At the same time, the artificially low wage rates did not permit workers to work long enough to buy the goods that were available at black market prices (Buchheim 1999, 60; Haus der Geschichte).

Paradoxically, the depressing effect of the controls on economic activity seemed to assure their survival. Germans overwhelmingly saw the end of controls as relinquishing the certain for a frightening uncertainty. Most believed that the disappearance of the ration cards and the minimum of sustenance that they assured would push the impoverished over the line into destitution (Bark and Gress 1989, 197).

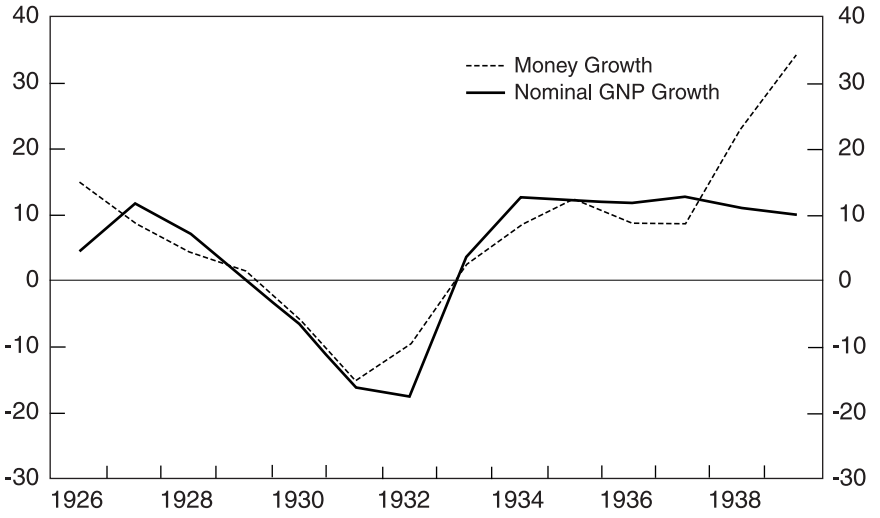
Currency Reform

By early 1947, U.S. Secretary of State George Marshall understood fully the desperate condition of Europe. He had come to believe that Stalin would not accept a unified Germany except under Russian control. Marshall and Ernest Bevin, the British foreign minister, concluded that the West had to reconstruct Germany without the Soviet Union (Smyser 1989, 55–56). On 5 June 1947, Marshall proposed what later became known as the Marshall Plan. The United States wanted an economically viable Western Europe, which required a prosperous Germany. Currency reform was a vital ingredient of economic reconstruction. Germany needed a universally accepted currency to be able to trade freely both within and across its borders.

The American economic advisers to General Clay prepared a plan for currency reform based on reforms in Russia and Czechoslovakia (Clay 1950, 209). The reform preserved the relatively low level of controlled prices by converting wage rates and pension payments one-for-one from reichsmarks to the new DMs. However, it eliminated the monetary overhang inherited from the Third Reich by limiting the quantity of reichsmarks that could be exchanged for DMs. Under the terms of the reform, which came into effect 20 June 1948, West Germans could turn in 60 reichsmarks on a one-to-one basis. Further exchanges occurred at a ratio of ten to one. The monetary authority converted assets and savings deposits at a sharply reduced rate.[40]

[40] This destruction of savings from inflation for a second time in 30 years seared into Germany's collective memory an aversion to inflation that would later provide popular support for the independence of the Bundesbank.

Figure 5 Money and Nominal GNP Growth



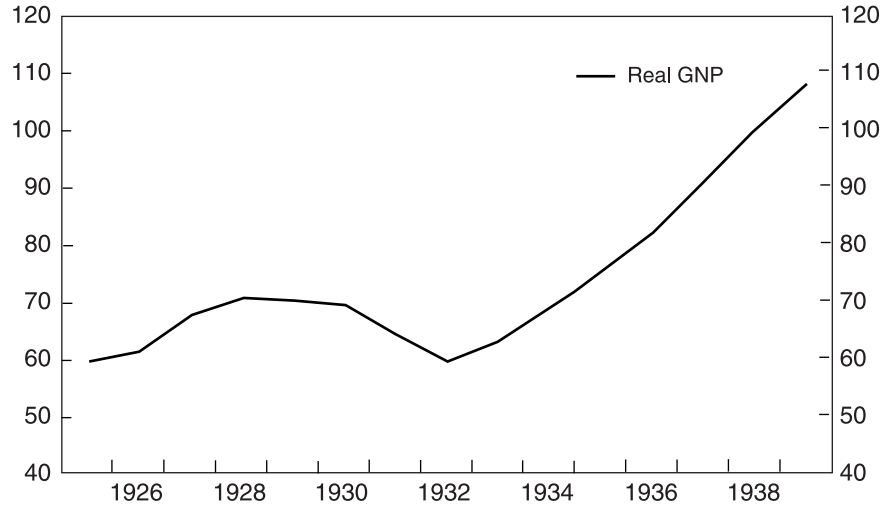
Notes: See Table 2.

Clay (1950, 210) feared that the reform would favor the German "trend toward socialism" by destroying the financial "savings of the little man" while rewarding the "black market operators who had invested their huge gains in real estate." Therefore, the original plan included provisions for a capital levy to compensate the losers from the ten-to-one write-down of the reichsmark. However, Marshall insisted on limiting the reform to the currency exchange.

The German economic miracle began with the currency reform of June 1948. When the currency reform was announced, Ludwig Erhard, economic director of the joint American-British Bizonal Economic Administration, decreed the end of price controls. The day the price controls ended, goods returned to store shelves. With the end of controls, "euphoria engulfed most Germans at the sight of goods and food items they could only dream about in the past. Bakeries miraculously produced and displayed delicious cakes; vegetables, butter and eggs appeared in abundance" (Bark and Gress 1989, 201).

Industrial production in West Germany rose 25 percent within two months and 50 percent within six months (Buchheim 1993, 72; Marsh 1992, 173). With prices that made crops worth selling, West Germany again had enough food (Smyser 1999, 50). From 1948 through 1950, GDP growth in Germany remained above 15 percent per year. In the 1950s, Germany grew at an average annualized rate of 8 percent (Giersch et al. 1992, 2, Figure 1).

Figure 6 Real GNP



Notes: See Table 2.

Nevertheless, immediately after their introduction, the fate of the Erhard reforms remained in question. The currency reform carried the opprobrium of destroying the financial wealth of ordinary Germans while enhancing the tangible wealth of those holding productive capital. It thus restored the industrialists' economic position, which had created envy in the Weimar Republic and Third Reich. Many Germans wanted a reintroduction of price controls, believing they would make basic goods and food affordable. In November 1948, more than 9 million German workers engaged in a general strike (Merkl 1963, 107).

Restoring the Appeal of Capitalism

The ultimate success of the Erhard reforms constitutes a striking example of how ideas can influence events. Erhard initiated the reforms in a social and intellectual environment uniformly predisposed toward central planning. Bark and Grass (1989) write:

> Most Europeans in 1945 regarded socialism, or at least some form of extensive state-controlled economy, as so obviously necessary as to be beyond argument. (193)

> In 1947 many Germans still saw capitalism as largely responsible for the soulless materialism of the modern age and for the alienation of man from his spiritual beliefs and from true religious values. This materialism and this alienation were, according to this view, the main reasons for the success of National Socialism. . . . [B]oth the main parties of West Germany in 1947–8 saw a free market system as an impossibility. (197)

The success of Erhard's free market reforms dramatically altered the postwar intellectual and political consensus. Gottfried Haberler and Thomas Willett (1968) write:

> Ludwig Erhard put Germany on the road to recovery through currency and economic reforms that swept away internal and external restrictions. . . . The impact of the German "economic miracle" through. . . the demonstration of what can be achieved by liberal [free enterprise] economic policies even under adverse conditions can hardly be exaggerated. (2)

Hitler boasted that the stability of the DM was a symbol of the discipline provided by the Nazi dictatorship (Marsh 1992, 23). However, he used price controls and repression to enforce that stability. Suppressed inflation then produced shortages and economic disruption. A new, democratic Germany combined price stability with free markets.

Through the vagaries of politics, in March 1948 Ludwig Erhard became head of the Economic Administration.[41] The United States and Britain had formed the joint American-British Bizonal Economic Administration in January 1947 to allow German administration of economic matters. Initially, the Social Democratic Party (SPD) controlled its Executive Committee (Clay 1950, 200). The SPD wanted a centrally planned, socialist economy.[42] The rival party, the Christian Democratic Union (CDU), desired the nationalization of basic industries and significant economic planning.

The Economic Administration had an advisory council of academic economists composed of "ordoliberals."[43] The ordoliberals disagreed with the SPD's belief that only socialism could assure the success of democracy (Bark and Grass 1989, 194). Instead, they argued that the gradual replacement of a laissez-faire economy with a corporatist economy that started in the nineteenth century had made Nazi totalitarianism possible through a concentration of economic and political power. They concluded that for a free society to

---

[41] The Allies had forced his predecessor to resign over an English translation of the word *Hühnerfutter* as "chicken feed," which was contained in a description of U.S. shipments of corn to Germany. (Germans fed corn primarily to chickens.) Although the translation was literal, the Americans found it to have a perjorative meaning.

[42] This paragraph and the paragraph following the next one draw on Giersch et al. (1992, 30ff).

[43] Walter Eucken, the economist from Freiburg University who coined the term, was a leading ordoliberal. He was one of the few German economists during the hyperinflation to argue that the inflation was a monetary phenomenon (Laidler and Stadler 1998, 823).

be maintained, the concentration of power must be prevented. Dispersal of power required the separation of powers in government and free markets in the economy. Only a freely functioning price system could allocate resources in a complex, highly specialized economy.

Erhard accepted the free market ideas of the ordoliberals. However, the American military administration hesitated to remove the price controls because they feared that the destruction of financial wealth combined with an increase in the price of basic commodities would create a backlash (Eisenberg 1993, 446). Erhard seized the initiative by unilaterally ending price controls on the day after the currency reform took effect. He also ended central planning by eliminating the use of rationing to allocate goods made necessary by the price controls. However, only the Allied command possessed the necessary authority to take these actions. Smyser (1999) gives the following account:

> Clay called in Erhard to lecture him. He told Erhard that U.S. experts had warned that his policies would fail. Erhard, puffing on his ever-present cigar, replied laconically: "My experts tell me the same thing." Clay, who had long since learned to distrust experts, thereupon backed Erhard to the hilt. (77)

The desperation produced by the postwar economic crisis made the economic reforms promoted by the left more appealing to German workers. Furthermore, the black market produced by the price controls had begun to recreate the politics of scapegoating. Impoverished Germans resented the black marketers, whom they blamed for the disappearance of goods from stores. One poster showed a young boy denouncing a black marketer dealing behind a store with empty windows. Government posters denouncing profiteers carried a picture of a skull and the inscription "The black market is death." Political parties appealed for votes on a platform of getting "goods out of the black market and into the kitchens of the needy housewives" (Haus der Geschichte).

However, by the middle of 1949 the general availability of food and clothing, aided by Marshall Plan deliveries, lessened criticism of Erhard's free market policies (Buchheim 1993, 76). The U.S. military reported:

> The currency reform has created a psychological as well as a material revolution in German life. Psychologically it has introduced the hope of better times and improved conditions. Cheer and optimism are taking the place of the skepticism and pessimism which previously prevailed. (OMGUS report cited in Bark and Gress, 202)

The restoration of free markets aligned individuals' hopes for a better future with their own initiative. When individuals no longer saw their fate as determined by external, incomprehensible forces, they turned from collectivist solutions (Bark and Gress, 203).

The disappearance of the black market and the abolition of controls eliminated the politics of scapegoating by realigning socially useful activity with legal activity. And the association of Erhard's free market reform with economic prosperity shifted the dominant intellectual consensus in favor of free markets. The end of price controls also produced a controlled experiment in the efficacy of economic institutions. Although both the British-U.S. Bizone and the French zone introduced the new currency, the French retained price controls. Thereafter, output in the former began to grow but languished in the latter (Giersch 1992, 41–42). On an even more monumental scale, West Germany prospered while East Germany stagnated.

The Berlin Blockade

Currency reform and geopolitical events intertwined. In the immediate post–World War II period, East-West politics turned on whether to unify Germany and, if so, under what conditions. Stalin wanted a united Germany reconstructed with a socialist social order and a sympathetic leftist government that would prevent future aggression toward the Soviet Union (Smyser 1999). In March 1948, the Six-Power London Conference recommended that the United States, Britain, and France authorize their occupied sectors to form a provisional government. The Soviets maintained that such a government would violate the Potsdam Protocol that stipulated a united Germany.

A currency reform confined to the West signaled the partition of Germany and frustrated Stalin's goal of establishing a unified Germany under Soviet influence (Bark and Gress 1989, 211). The Western powers had been negotiating with the Soviets to introduce a single currency to be used throughout occupied Germany. However, against the backdrop of deteriorating American-Soviet relations, Washington looked for an excuse to end Soviet involvement both in currency reform and more generally in the reconstruction of Germany (Eisenberg 1996, 380–82).

General Vassily Sokolovsky was the Russian commander in charge of the Soviet occupation zone and the Russian representative in the Allied group that was set up to consider quadripartite currency reform. General Clay notified him that the Western powers were going to proceed with a currency reform. Sokolovsky replied angrily that "by your unilateral illegal decision you. . . effect a separate currency reform in Western Germany, whereby you liquidate the unity of money and complete the split of Germany."[44] The Soviets then had to introduce their own currency reform to avoid a flood of worthless old reichsmarks into their zone. On 23 June, General Sokolovsky introduced the ostmark—an old reichsmark to which a coupon was glued. Clay then

[44] The quote and material in this paragraph are from Eisenberg (1993, 410).

openly provoked the Soviets by introducing the new DMs in Berlin. On 24 June, the Soviets responded with a blockade of the Western sector of Berlin.

During the blockade, Berlin had two competing currencies, the DM and the ostmark. West Berliners effectively voted for the DM by using it in transactions. Use of the DM also helped defeat the blockade. Its general acceptance allowed West Berliners to attract market goods from the areas around West Berlin. On 20 March 1949, when the DM became the sole legal tender in the Western sector, the city became part of West Germany.

The Berlin blockade pushed West Germans toward acceptance of the Western ideals of free markets and democracy. That choice supported Chancellor Adenauer's program of free market reforms, European integration, and military strength opposed to the Soviet Union (Merkl 1963, 108). In the first elections to the federal parliament, the political parties that prevailed favored a free market over a planned economy.

4. CONCLUDING COMMENT

The end of World War II had left Germany without national institutions. The DM became the first national symbol of the new Germany. West Germany had the DM even before it had a flag. Capie and Wood (1999, 459) refer to the DM as "at once a cause and a symbol of Germany's recovery." Germans prized the stability of the mark as a symbol of its social stability and economic prosperity. The DM symbolized everything that Germany did right after the War.

APPENDIX: THE FEDERAL RESERVE BANK OF NEW YORK AND THE REICHSBANK DURING THE 1931 CRISIS

In June and July 1931, Governor Harrison, head of the Federal Reserve Bank of New York, maintained an almost daily record of his responses to the Reichsbank's pleas for assistance. The detailed character of the memoranda that he preserved implies that he wanted the record to show that he made all reasonable efforts to aid the Reichsbank so that no one could hold him responsible for the collapse of Germany's gold standard and banking system. Ironically, the record creates the impression that Harrison never in fact intended to aid the Reichsbank. Maintenance of investor confidence in the ability of the Reichsbank to maintain the gold parity required the New York Fed to organize central bank and private bank support of the Reichsbank.

The unwillingness of Harrison to exercise leadership doomed the international cooperation that could have maintained that confidence.

The record shows that Harrison had regular communication with Montagu Norman, governor of the Bank of England. On 20 June 1931, Norman informed Harrison that the Reichsbank had suffered a loss of 70 million reichsmarks and was near its legal limit on gold cover. The Reichsbank wanted a loan from the Bank of England and the New York Fed.[45] On 23 June Norman informed Harrison that "the situation was so critical and dangerous" that the Bank of England and the New York Fed should lend the Reichsbank $100 million. Harrison objected, arguing that if France did not participate, it would have no incentive to go along with the Hoover debt moratorium; Norman argued that Harrison's objection was irrelevant.[46] However, the Banque de France did agree to participate and the Reichsbank received the loan, with the Fed's share limited to $35 million.

The international political situation continued to deteriorate and with it investor confidence as the French continued to attempt to use the proposed Hoover debt moratorium to extract strategic concessions from Germany.[47] Norman commented that "the French can afford to wait, but Berlin cannot, for it is being bled to death."[48] The French wanted Germany to abandon construction of a naval battleship and plans for a customs union with Austria.[49] Such concessions would have brought down the German government.

Shepard Morgan, a partner in J. P. Morgan, wrote a memorandum placed in Harrison's files arguing that the United States should prevent U.S. Banks from withdrawing their deposits in German banks. The only purpose of allowing such a runoff would be to initiate an economic collapse that would "oblige the French to step in, either for the purpose of saving Germany from Hitlerism or communism or to protect their reparations revenues."[50] Morgan wrote:

> German unity dates only from 1871, and forces of disruption latent in tradition and the German character have always to be reckoned with. If disintegration should occur, the main objective of the French foreign policy would be served, that is, the advance of national security. Germany in a condition corresponding to that at the close of the Thirty Years' War would not threaten French security for a generation or more.[51]

---

[45] Harrison memo, subject "Germany," 20 June 1931, Harrison Papers.

[46] Harrison memo, subject "Germany," 24 June 1931, Harrison Papers.

[47] Memo to George Murnane, partner Lee, Higginson & Company, from Berlin, sent to Harrison, 6 July 1931, Harrison Papers; Cable from Berlin to Murnane, 6 July 1931, Harrison Papers.

[48] Norman telephone conversation with Harrison, 1 July 1931, Harrison Papers.

[49] Cablegram Moret to Harrison, 12 July 1931, Harrison Papers.

[50] Shepard Morgan, "Memorandum on German Short-Dated Debt Reduction," 6 July 1931, Harrison Papers.

[51] Ibid.

Morgan wanted Harrison to coordinate the responses of individual banks to the crisis to stop the first-out incentive. That is, each bank individually had an incentive to be the first to withdraw its funds from Germany although collectively banks would be better off if they maintained their deposits. Morgan advised Harrison that "the banks, for their part, should agree not to make further efforts to withdraw credits from Germany except after learning from the Federal Reserve Bank that such action will not threaten the position of the Reichsbank."[52]

On 8 July, Hans Luther, president of the Reichsbank, cabled Harrison that he could not "see another way out but the granting of long-termed large rediscount credit to the Reichsbank to prevent Germany from heavy distress. This idea is based on the expectation that such a credit will—because of its large amount and period of time—finally stop the mistrust against Germany without it really being used."[53] On 9 July, Norman informed Harrison that Luther had said without the credit "Germany will collapse" and that he agreed.[54]

Harrison replied that he "was very skeptical about the idea of a credit [and that] the chief difficulty was a flight from the reichsmark by German nationals and that the Reichsbank should resort to much more domestic credit control." Norman responded that "he thought Luther was now rationing credit very strictly." Harrison "insisted that these credit restrictions should be adopted. . . before we could fairly be asked for a sizable new credit."[55]

Harrison demanded that the Reichsbank reduce its discount window lending as a condition for a loan.[56] Despite Luther's concern expressed earlier for "the great danger for German business and trade in their state of bloodlessness," Luther acceded to New York's demand that the Reichsbank cease discounting commercial paper.[57] The Reichsbank then had to allow the reserves and credit structure of its banking system to collapse as gold flowed out.[58] On 30 May 1931, the Reichsbank held 2,577 million reichsmarks in reserves. On 31 July 1931, it held only 1,610 million reichsmarks—a decline

---

[52] Ibid.

[53] Luther had replaced Schacht as president in March 1930.

[54] Telephone call Harrison to Norman, 9 July 1931, Harrison Papers.

[55] Ibid. Harrison repeated his demand to Governor Moret of the Banque de France and McGarrah, head of the Bank for International Settlements, which organized discussion among central bank heads. Cablegrams Harrison to Moret and Harrison to McGarrah, 9 July 1931, Harrison Papers.

[56] Cablegram Harrison to Luther, 10 July 1931, Harrison Papers. See also telephone call Dr. Dreyse, Reichsbank vice president, to Harrison, 10 July 1931, Harrison Papers.

[57] Cablegram Luther to Harrison, 4 July 1931, Harrison Papers.

[58] On the restriction of discount window lending, see cable from Berlin to Murnane, 6 July 1931, Harrison Papers.

of 37.5 percent in two months.[59] Closing its discount window to new discounts was the Reichsbank's last desperate attempt to meet the ever-changing conditions that Harrison set for a loan to replace the gold outflows.

On Saturday, 11 July, Dreyse, the Reichsbank vice president, informed Harrison by cablegram that "the fear of complete cessation of credit expansion on the part of the Reichsbank" risked "the imminent collapse of essential parts of the economic structure."[60] He informed Harrison that credit restriction had reduced reichsmark circulation from 6,300 million a year ago to 5,600 million. In a separate telephone conversation, Dreyse also advised that without prompt help, "the Bruening government would fall."[61]

Also on 11 July, Luther informed Harrison that "a breakdown on Monday, both political and financial, was probably inevitable unless some help were forthcoming."[62] Harrison told Dr. Kiep, the German Consul General in New York, that the New York Fed could participate in a credit only if the Bank of England and the Banque de France participated. Harrison knew that France was imposing political conditions, "the unqualified acceptance of which would mean the immediate downfall of the present government."[63] On 12 July, Harrison cabled Luther about the latter's cable of 9 July mentioning "the urgent need of a large rediscount credit."[64] Incredibly, Harrison argued that New York could do nothing "in view of the fact that this bank had never been asked directly or formally by the Reichsbank to grant a credit."[65]

Dr. Kiep relayed a report from Germany informing Harrison that the Reichsbank had lost 100 million marks on Friday and that a run on German banks "ha[d] set in with great force" and would resume Monday.[66] On Monday, 13 July, Harrison telephoned Norman in Basle, where Norman was attending an emergency meeting of central bank governors at the Bank for International Settlements (BIS). Harrison later reported their conversation in a memo:

> He [Norman] had come out of the meeting to answer my call. . . . Norman then said that he was leaving in an hour to go back to London and that the board was going to adjourn. I said, if that is true, what is going to happen? He said nothing is going to happen—absolutely nothing. I said

---

[59] Table No. 167, "Reichsbank-Principal Assets and Liabilities, Monthly, 1924–1941," in Board of Governors (1976).

[60] Cablegram Dreyse to Harrison, 11 July 1931, Harrison Papers.

[61] Harrison memo, "Germany. Credit to Reichsbank," 11 July 1931, Harrison Papers.

[62] Harrison memo, "Reichsbank," 12 July 1931, Harrison Papers.

[63] Telegraphic report from Berlin, Dr. O. C. Kiep, 11 July 1931, Harrison Papers.

[64] Cablegram Harrison to Luther, 12 July 1931, Harrison Papers.

[65] J. E. Crane memo, "The German Situation," 14 July 1931, Harrison Papers.

[66] Dr. O. C. Kiep, "Telegraphic report from Berlin," 11 July 1931, Harrison Papers.

> I thought it was a great pity that they should adjourn before something was done. He said: "Well, what have you to suggest?"[67]

Harrison replied irrelevantly that "the private banks in the various markets of the world" should make a credit. Harrison concluded his memo by saying that Norman "seemed tired, disgruntled and discouraged."

McGarrah, head of the BIS, told Harrison that "in his judgment it was a pity to let it smash."[68] On Monday, 13 July, the bank runs resumed and Germany declared a bank holiday. On 15 July, the gold reserve ratio fell below the legal minimum of 40 percent. On 15 and 18 July, Germany abandoned the gold standard by imposing foreign exchange controls. George Murnane, head of the New York discount house Lee Higginson & Co., cabled Harrison that his office in Berlin expected "within [the] shortest time. . . grave disturbances. . . [from an] immediate enormous increase [in] unemployment."[69]

On 11 July, the German government appealed to the U.S. government to urge the New York Fed to grant a loan and make a "reassuring statement to the press about conditions in Germany."[70] On 13 July, Dreyse urged Harrison "to exercise at this very last hour your great far-reaching influence."[71] Shepard Morgan informed Harrison that "hope for a peaceful and orderly Germany centers on the maintenance of this [Bruening] government. . . . [E]ither of the extreme parties coming into power would bring complete disorder."[72]

REFERENCES

"Argentine Protests on Austerity Ebb, and Stocks Rise." 2001. *New York Times* (9 August): W1.

Bagehot, Walter. 1873. *Lombard Street: A Description of the Money Market.* London: H. S. King.

Balderston, Theo. 1994. "The Banks and the Gold Standard in the German

[67] Harrison memo, "Credit to the Reichsbank," 14 July 1931, Harrison Papers.

[68] Ibid.

[69] "Cable received by Mr. Murnane from Berlin," 22 July 1931, relayed to Harrison by the former's secretary, Harrison Papers.

The number of unemployed in Germany rose from 1.391 and 1.899 million, respectively, in 1928 and 1929 to 4.520 and 5.575 million, respectively, in 1931 and 1932. "Arbeitslose, Allgemeine Wirtschaftsentwicklung 1924–1974, Ausgewahlte Daten zur Wirtschaftsentwicklung 1876–1974," Deutsches Geld und Bankwesen 1876–1975, Deutsche Bundesbank, Frankfurt 1976.

[70] J. E. Crane memo, "The German Situation," 14 July 1931, Harrison Papers.

[71] Cablegram Dreyse to Harrison, 13 July 31, Harrison Papers.

[72] Memo to Governor Harrison from W. Randolph Burgess, 13 July 1931, Harrison Papers.

Financial Crisis of 1931." Financial History Review 1 (April): 43–68.

Bark, Dennis L., and David R. Gress. 1989. *From Shadow to Substance: 1945–1963.* Cambridge, Mass.: Basil Blackwell.

Bernholz, Peter. 1999. "The Bundesbank and the Process of European Monetary Integration." In *Fifty Years of the Deutsche Mark*, ed. Deutsche Bundesbank. Oxford: Oxford University Press.

Board of Governors of the Federal Reserve System. 1930. "Review of the Month." *Federal Reserve Bulletin* 16 (November): 655–58.

_____. 1976. *Banking and Monetary Statistics 1914–1941*.

Bresciani-Turroni, Constantino. 1937. *The Economics of Inflation. A Study of Currency Depreciation in Post-War Germany.* London: G. Allen & Unwin Ltd.

Bruno, Michael, Guido Di Tella, Rudiger Dornbusch, and Stanley Fischer. 1988. *Inflation Stabilization: The Experience of Israel, Argentina, Brazil, Bolivia, and Mexico.* Cambridge, Mass.: MIT Press.

Buchheim, Christoph. 1993. "Marshall Plan and Currency Reform." In *American Policy and the Reconstruction of West Germany, 1945–1955*, ed. Jeffry M. Diefendorf, Axel Frohn, and Hermann-Josef Rupieper. Cambridge: Cambridge University Press: 69–83.

_____. 1999. "The Establishment of the Bank Deutscher Lander and the West German Currency Reform." In *Fifty Years of the Deutsche Mark,* ed. Deutsche Bundesbank. Oxford: Oxford University Press.

Cagan, Phillip. 1956. "The Monetary Dynamics of Hyperinflation." In *Studies in the Quantity Theory of Money,* ed. Milton Friedman. Chicago: University of Chicago Press.

Capie, Forrest, and Geoffrey Wood. 2001. "The Birth, Life and Demise of a Currency: 50 Years of the Deutsche Mark." *Economic Journal* 111 (June): 449–61.

Chandler, Lester. 1958. *Benjamin Strong, Central Banker.* Washington: The Brookings Institution.

Clay, Lucius D. 1950. *Decision in Germany.* New York: Doubleday.

Eichengreen, Barry. 1992. *Golden Fetters: The Gold Standard and the Great Depression, 1919–1939.* Oxford: Oxford University Press.

Eisenberg, Carolyn Woods. 1996. *Drawing the Line: The American Decision to Divide Germany, 1944–1949.* Cambridge: Cambridge University Press.

Feldman, Gerald D. 1997. *The Great Disorder: Politics, Economics, and Society in the German Inflation, 1914–1924.* Oxford: Oxford University Press.

Ferguson, Niall, and Brigitte Granville. 2000. "Weimar on the Volga: Causes and Consequences of Inflation in 1990s Russia Compared with 1920s Germany." *Journal of Economic History* 60 (December): 1061–87.

Fisher, Jonas D. M., and Andreas Hornstein. 2001. "The Role of Real Wages, Productivity and Fiscal Policy in Germany's Great Depression 1928–37." Manuscript, Federal Reserve Bank of Chicago and Federal Reserve Bank of Richmond.

Friedman, Milton, and Anna J. Schwartz. 1963. *A Monetary History of the United States, 1867–1960*. Princeton: Princeton University Press.

Giersch, Herbert, Karl-Heinz Paqu, and Holger Schmieding. 1992. *The Fading Miracle: Four Decades of Market Economy in Germany.* Cambridge: Cambridge University Press.

Goedde, Isabel. 2000. "The German Twin Crisis of 1931." Manuscript, Department of Economics, University of Mannheim.

Graham, Frank D. 1930. *Exchange, Prices, and Production in Hyper-Inflation Germany, 1920–1923.* Princeton: Princeton University Press.

Haberler, Gottfried, and Thomas D. Willett. 1968. *U.S. Balance-of-Payments Policies and International Reform: A Critical Analysis.* Washington: American Enterprise Institute.

Hamilton, James D. 1992 "Was the Deflation during the Great Depression Anticipated? Evidence from the Commodity Futures Market." *American Economic Review* 82 (March): 157–78.

Harrison, George. Papers. Federal Reserve Bank of New York and Columbia University.

Holtfrerich, Carl-Ludwig. 1986. *The German Inflation 1914–1923: Causes and Effects in International Perspective.* New York: Walter de Gruyter.

Humphrey, Thomas M. 1980. "Eliminating Runaway Inflation: Lessons from the German Hyperinflation." Federal Reserve Bank of Richmond *Economic Review* 66 (July/August): 3–7.

_____. 1998. "Mercantilists and Classicals: Insights from Doctrinal History." Federal Reserve Bank of Richmond *1998 Annual Report.*

James, Harold. 1986. *The German Slump: Politics and Economics 1924–1936.* Oxford: Clarendon Press.

_____. 1999. "The Reichsbank 1876–1945." In *Fifty Years of the Deutsche Mark,* ed. Deutsche Bundesbank. Oxford: Oxford University Press: 3–53.

Johnson, Peter A. 1998. *The Government of Money: Monetarism in Germany and the United States*. Ithaca: Cornell University Press.

Keynes, John Maynard. [1919] 1971. "The Economic Consequences of the Peace." In *The Collected Writings of John Maynard Keynes,* vol. 2. London: Macmillan.

_____. [1923] 1971. "A Tract on Monetary Reform." In *The Collected Writings of John Maynard Keynes,* vol. 4. London: Macmillan.

Laidler, David W., and George W. Stadler. 1998. "Monetary Explanations of the Weimar Republic's Hyperinflation: Some Neglected Contributions in Contemporary German Literature." *Journal of Money, Credit, and Banking* 30 (November): 816–31.

Merkl, Peter H. 1963. *The Origin of the West German Republic.* New York: Oxford University Press.

Nicholls, A. J. 1968. *Weimar and the Rise of Hitler.* New York: St. Martin's Press.

Pontzen, Martin. 1999. "The German Banking Crisis of 1931." *Central Banking* 9 (February): 76–80.

Pringle, Robert. 1998. "Hitler's Central Banker." *Central Banking* 9 (August): 67–72.

Rothbard, Murray N. 1996. "The Gold Exchange Standard in the Interwar Years." In *Money and the Nation State. The Independent Institute,* ed. Kevin Dowd, Richard H. Timberlake, Jr., and Merton Miller. Oakland, CA.

Sargent, Thomas. 1993. "The Ends of Four Big Inflations." In *Rational Expectations and Inflation,* 2nd ed. New York: Harper Collins College: 44-115.

Schacht, Hjalmar. 1927. *The Stabilization of the Mark.* New York: Adelphi Company.

Smyser, W. R. 1999. *From Yalta to Berlin: The Cold War Struggle over Germany.* New York: St. Martin's Press.

Webb, Steven B. 1989. *Hyperinflation and Stabilization in Weimar Germany.* Oxford: Oxford University Press.

Yeager, Leland B. 1976. *International Monetary Relations: Theory, History and Policy.* New York: Harper & Row.

_____. 1981. *Experiences with Stopping Inflation.* Washington: American Enterprise Institute.

Pitfalls in Interpreting Tests of Backward-Looking Pricing in New Keynesian Models

Michael Dotsey

Recently macroeconomists have shown renewed interest in economic models that contain some form of nominal rigidity. These models are referred to generically as New Keynesian models. A particularly important feature of these models is sluggishness in price adjustment. However, there is substantial debate over whether this sluggishness arises from backward-looking adaptive behavior or from forward-looking behavior in the presence of costs in adjusting prices. It is also possible that the economy comprises two types of firms, one type that adjusts the price of its product based on some backward-looking policy and another type that sets its price based on current and anticipated market conditions. Because the nature of price setting is one of the key aspects of New Keynesian models, developing empirical tests that will inform theorists of the correct specification of pricing behavior is essential.

Also, from a policy perspective, understanding how firms set prices is of crucial importance because it determines what the effects of monetary policy will be. For example, as discussed in Ball (1994) and Roberts (1998), credible disinflations are relatively costless in New Keynesian models, but are quite costly from the perspective of traditional backward-looking Keynesian models.

In an attempt to shed empirical light on this question, economists have started investigating the behavior of inflation based on the null hypothesis that

I would like to thank Robert King and Mark Watson for many helpful discussions. I have also benefited from the comments of Thomas Humphrey, Pierre Sarte, John Weinberg, and Alex Wolman. The views expressed herein are the author's and do not necessarily represent those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

firms are indeed forward looking. The goal of this work is to test if forward-looking price behavior is consistent with the actual behavior of prices and inflation. This strategy is attractive as a starting point because it is compatible with firms' optimizing behavior. If inhibitions to perfect price flexibility exist, such as adjustment costs or maintaining long-term customer relationships, then it is optimal for a firm to take account of how a chosen price will affect its future profit stream. That is, the firm's pricing decision will be forward looking in much the same way that current investment decisions are based on expectations of future economic conditions. Seminal work in this area has been carried out by Gali and Gertler (1999) and Sbordone (1998).

Many tests used to assess whether forward-looking pricing behavior adequately captures the behavior of inflation also investigate whether the addition of some backward-looking variables appreciably helps explain inflation. A finding that lags of inflation have marginal predictive content is interpreted to mean that a significant fraction of firms are backward looking. Further, this fraction can be estimated. The empirical debate has largely centered on what relevant variables, such as output gaps or marginal cost, should be included in the specification, how to properly measure the variables in question, and the estimation strategy itself. As of yet, there is no general consensus regarding how important forward-looking behavior is in a firm's pricing decisions.[1]

This article takes a different tack. To believe in forward-looking pricing is one thing; it is an entirely different matter to agree on what form that pricing behavior takes. Is it time or state dependent? If time dependent, which of the leading models best describes pricing behavior? Can it be represented by a Calvo-style or quadratic adjustment-cost model? Or is it more amenable to a staggered contracting model in the spirit of Taylor (1980)? As Kiley (1998) and Wolman (1999) have shown, these various models with forward-looking pricing have different implications for how shocks affect the economy and therefore are likely to give rise to different empirical interpretations of pricing behavior. As is also indicated in Guerrieri (2001), the models lead to very different estimable equations. I show that if data are actually generated by a forward-looking model of the Taylor pricing variety, and one estimates a pricing relationship based on Calvo-type behavior, then the conclusion that a significant fraction of firms are backward looking must follow. Thus, the interpretation of various coefficients in existing tests is open to question. Ascertaining the extent of backward-looking pricing behavior may be a more difficult exercise than is currently acknowledged.

[1] Important articles in this literature include Sbordone (1998, 2001), Gali and Gertler (1999), Gali, Gertler, and Lopez-Salido (2001), Rudd and Whelan (2001), Fuhrer (1997), and Roberts (2001).

1. PRICING MODELS

I will begin by outlining two basic pricing models and their implied empirical tests. The first model is the workhorse model of Calvo (1983), which serves as the basis for an important strand of the empirical literature.[2] The second model is a generalization of the more reasonable specification of staggered pricing behavior as postulated by Taylor (1980). The generalization of this model assumes that a fraction of firms change their price in any given period and that at some point every firm will change its price with probability one. The Taylor model can, therefore, be viewed as a truncated version of the Calvo model. I will also exposit the hybrid model of Gali and Gertler (1999), where a fraction of firms follow a backward-looking rule of thumb convention in setting their price.

Calvo-Style Price Setting

In the Calvo price-setting framework, each firm faces a constant probability, $1-\theta$, that it will be able to adjust its price in the current period and a corresponding probability of θ that it must charge the same price it charged last period. These features imply two equations governing the behavior of prices. One is a backward-looking price level (p_t) equation that is a weighted average of the nominal prices set by firms in prior periods (p^*_{t-j}). Its log-linear approximation takes a particularly simple form,

$$p_t = (1-\theta)\sum_{j=0}^{\infty}\theta^j p^*_{t-j} = \theta p_{t-1} + (1-\theta)p^*_t, \tag{1}$$

where all variables are in logarithms. Equation (1) can also be expressed as a partial adjustment mechanism, $p_t - p_{t-1} = (1-\theta)[p^*_t - p_{t-1}]$. The partial adjustment interpretation indicates that the price level responds only gradually when p^*_t is raised above p_{t-1}, with the extent of price level adjustment equal to the probability of price adjustment. Equation (2) (also a log-linear approximation) describes forward-looking price setting and reflects the notion that firms understand they may not be able to reset their price in future periods. They appropriately set their price to maximize a discounted expected stream of profits. Thus, current price setting depends on future nominal marginal

[2] The quadratic cost-of-adjustment model developed by Rotemberg (1982) gives rise to a similar pricing equation.

cost,

$$p_t^* = (1-\beta\theta)\sum_{j=0}^{\infty}(\beta\theta)^j E_t[(\psi_{t+j}/\psi) + p_t]$$
$$= \theta\beta E_t p_{t+1}^* + (1-\beta\theta)[(\psi_t/\psi) + p_t], \quad (2)$$

where ψ_t is the logarithm of real marginal cost and ψ is the logarithm of the steady state value of real marginal cost and β is the rate at which future utility is discounted. E_t is the conditional expectations operator where expectations are conditioned on all current and past information. Combining equations (1) and (2) yields an equation for inflation of the form

$$\pi_t = \lambda\psi_t + \beta E_t\pi_{t+1}, \quad (3)$$

where $\pi_t = p_t - p_{t-1}$, and $\lambda = (1-\theta)(1-\beta\theta)/\theta$.

Using the Calvo model of price adjustment is attractive because of its tractability and parsimony. It is largely because of these two characteristics that the Calvo model has taken center stage in empirical work regarding forward pricing behavior. The model, however, contains a number of unrealistic features. For example, there exists a measurable fraction of firms that have not changed their price for an arbitrarily long time, and these firms produce a significant portion of total output. Accordingly, one would at least expect all firms to change their price after some finite length of time. It is hard to believe that the costs of adjusting prices are so high that it is not beneficial to change prices frequently. Thus, a useful extension of the model would be to set a finite time limit over which a firm's price remains unchanged. Setting such a time limit makes the pricing formulas much more complex and would not be worthwhile if the implications of the added realism were innocuous. Wolman (1999) and Kiley (1998) indicate that this truncated version of the model yields very different behavior than the original Calvo model. I therefore investigate the pricing implications of the truncated model because it may provide a more realistic version of firm behavior.

Generalized Taylor Staggered Price-Setting

In the Taylor framework, as in the Calvo model, a firm that has not changed its price for j periods faces a probability α_j of changing its price, but at some finite horizon J a firm changes its price with probability one. If $\alpha_j = 0$ for all $j < J$, then the model is the basic staggered price-setting model of Taylor (1980) with $1/J$ of firms changing prices each period. Wolman (1999) argues that a more realistic price-setting model would involve monotonically increasing probabilities, $0 \leq \alpha_j \leq \alpha_{j+1} < 1$ for all $j < J$, and $\alpha_J = 1$. His specification implies that a firm that has not changed its price for a number of periods is more likely to change its price than a firm that recently reset its price.

For ease of comparison with the basic Calvo model, I assume $\alpha_j = \alpha < 1$ for all $j < J$, and for tractability take $J = 3$. As in the Calvo model, price-setting behavior is characterized by two equations (see the appendix), a backward-looking equation describing the price level,

$$p_t = \omega_0 p_t^* + \omega_1 p_{t-1}^* + \omega_2 p_{t-2}^*, \tag{4}$$

and a forward-looking equation depicting optimal price-setting,

$$p_t^* = \rho_0(\psi_t + p_t) + \rho_1 E_t(\psi_{t+1} + p_{t+1}) + \rho_2 E_t(\psi_{t+2} + p_{t+2}). \tag{5}$$

Both of these equations are linearizations around zero inflation of the nonlinear equations that exactly describe model behavior, and the variables in both are expressed as logarithmic deviations from steady state.[3] The parameters ω_j represent the fraction of firms that have not changed their price for j periods and are a function of α. The ρ's arise from the linearization of the optimal price-setting equation and involve the probability α and the time discount factor β that agents use when discounting future utility.

Combining (4) and (5) yields the following difference equation in inflation and marginal cost:

$$\begin{aligned}\{1 + c_1 L + c_2 L^2 + c_3 L^3\} E_{t-2}\pi_{t+2} = \\ -\{1 + a_1 L + a_2 L^2 + a_3 L^3 + a_4 L^4\} E_{t-2}\psi_{t+2},\end{aligned} \tag{6}$$

where E_{t-2} is the expectations operator conditional on information as of $t-2$ and L is the lag operator.[4] As mentioned, the Calvo and Taylor models of price setting result in very different nominal behavior, and these differences carry over to the empirical tests of forward-looking pricing. Equation (6) is the analogue to (3) and contains a number of important differences. First, lagged inflation enters this expression, as does lagged marginal cost. Also, the lead structure in (6) is more complicated, and expectations are conditioned on more distant past information. The different conditioning set will have implications for the admissibility of variables as instruments in the estimation carried out below.[5]

---

[3] The price level equation is given by $P_t = \left[\sum_{j=0}^{2} \omega_j P_{t-j}^{*(1-\epsilon)}\right]^{\frac{1}{1-\epsilon}}$, where variables are in levels and ϵ is the elasticity of demand for the firm's product. The optimal price-setting equation is $p_t^* = \frac{\varepsilon}{\varepsilon-1} \frac{\sum_{j=0}^{2} \beta^j E_t\{(\omega_j/\omega_0)\cdot(\lambda_{t+j}/\lambda_t)\cdot\psi_{t+j}\cdot(P_{t+j}/P_t)^{\varepsilon}\cdot y_{t+j}\}}{\sum_{j=0}^{2} \beta^j E_t\{(\omega_j/\omega_0)\cdot(\lambda_{t+j}/\lambda_t)\cdot(P_{t+j}/P_t)^{\varepsilon-1}\cdot y_{t+j}\}}$, where y is the firm's level of output. For more detail concerning the derivation in the text, see Dotsey, King, and Wolman (1999).

[4] In deriving (6), use was made of the fact that unity is one of the roots of the fourth order polynomial describing the behavior of deviations in the price level around its steady state (see the appendix).

[5] Similar observations are made by Guerrieri (2001).

Further, the different empirical implications generated by (6) apply to more realistic pricing models such as the one used by Wolman (1999), who assumes that the probability a firm will change its price is increasing in the elapsed time since its last price adjustment, and to state-dependent models of the type explored in Dotsey, King, and Wolman (1999). The essential characteristic of these types of models is that they generate higher order difference equations in inflation. They do so as long as firms exist that have not adjusted their price for more than two quarters, a feature needed to match microdata on firm pricing, and as long as all firms adjust their price in some finite time interval.

A Hybrid Calvo Model

To investigate whether backward price-setting behavior is also needed to explain the data, economists have postulated that only a fraction of firms base their price on optimizing behavior and that the remaining firms use a rule of thumb based on past prices and inflation. Within the Calvo framework, Gali and Gertler (1999) describe one such rule that leads to a relatively tractable hybrid Phillips curve. Their pricing rule is depicted by

$$p_t^b = \overline{p}_{t-1}^* + \pi_{t-1}.$$

Backward-looking firms set their price, p_t^b, based on an index reflecting the behavior of all firms who changed their price last period, $\overline{p}_{t-1}^*$, and on a correction term involving lagged inflation, π_{t-1}.

In turn, the current price index reflecting the behavior of all price setters is given by

$$\overline{p}_t^* = (1 - \varpi)\, p_t^* + \varpi\, p_t^b,$$

where ϖ is the fraction of firms that are backward looking. As long as forward-looking price setters compose a significant fraction of firms, the price index of newly set prices will be dominated by forward-looking firms. In the presence of low rates of inflation, the backward-looking price setter's price will not depart far from $\overline{p}_{t-1}^*$. Taken in conjunction, these two assumptions imply that prices set by backward-looking price setters will not depart very far from an optimizing price.

The hybrid model just described implies an equation describing inflation of the form

$$\pi_t = \lambda \psi_t + \gamma_f E_t \pi_{t+1} + \gamma_b \pi_{t-1}, \tag{7}$$

where λ, γ_f, and γ_b are, respectively, nonlinear functions of the discount rate, the probability of price adjustment, and the fraction of firms that are forward

looking. In estimation of equations of the form (7), a significant coefficient on lagged inflation is generally taken to imply some departure from rationality on the part of agents (see Roberts [1998]). In the hybrid model, the departure is represented by a fraction of firms that are backward looking, and, as in Gali and Gertler (1999), that fraction is readily ascertained by uncovering the fundamental parameters of the model.[6] I show below that this interpretation is part of a joint hypothesis, an important component of which is the Calvo model of pricing. If prices are indeed forward looking, but are generated from behavior consistent with the generalized Taylor-style pricing model, then the interpretation may not be correct.

2. INTERPRETING TESTS FOR BACKWARD-LOOKING BEHAVIOR

In this section, data are generated from a generalized Taylor staggered pricing model and then used in tests based on Calvo-style pricing to investigate the estimated presence of backward-looking price setting given the knowledge that all firms in the model are forward looking. Other than the pricing behavior, which is depicted by equations (4), (5), and (6), the particular details of the model are not overly important. What is important is that data on marginal cost and inflation are being generated in a manner that is consistent with the underlying state variables of the model. Such treatment is consistent with the empirical work in this area, where only the pricing behavior is carefully exposited. The full model is that of Dotsey and King (2001) without intermediate inputs, and it is driven by shocks to money growth, technology, money demand, and government spending. Thus, the state variables are the aforementioned shocks, past relative prices, and the capital stock.[7]

Before I test the model, it is worth reiterating an important feature of the pricing equations, namely that real marginal cost is the appropriate variable to be included in the determination of inflation. This point has been strongly emphasized by Gali and Gertler (1999) and Sbordone (1998). Many authors, however, have used the output gap, defined as the deviation of the level of output from its long-run or trend level, as the principal determinant of inflation.[8] If we take the various New Keynesian sticky price models as the null to be tested, we see that this alternative procedure is a mistake.

Output-gap measures produce serious problems of measurement error under the null of a New Keynesian model. Under suitable assumptions about

[6] Another interpretation is that expectations only adapt gradually to their rational value.

[7] The shocks all have a standard deviation of 1 percent and the respective autoregressive parameters are 0.8, 0.6, 0.7, and 0.4 for technology, money supply, money demand, and government spending. Thus, I have made no attempt to accurately calibrate the driving processes.

[8] For example, see Fuhrer (1997) and Roberts (2001).

technology and factor markets, the relationship between marginal cost and potential output, y_t^* (which is output that would occur if prices were counterfactually flexible), is $\psi_t - \overline{\psi} = \kappa(y_t - y_t^*)$, where $\overline{\psi}$ is steady state marginal cost. The right-hand-side term may be rewritten as the sum of two terms, $(y_t - y_t^{trend}) + (y_t^{trend} - y_t^*)$, where y^{trend} is some measure of trend output. The first term corresponds to the output gap and the last term embodies the measurement error associated with using the output gap. The bias induced by this measurement error will depend on the way trend output is measured and other features of the economy, notably the conduct of policy. For example, if policy kept the price level constant, then there would be no variation in marginal cost in response to a technology shock. There would, however, be variation in the output gap, causing its coefficient in a Phillips curve relationship to be biased downward. In response to other shocks and to other policy rules, the effects of the misspecification on the estimated coefficients could become quite complicated.

Testing the Generalized Taylor Price-Setting Model

The test of forward-looking pricing behavior is implemented by using the equation describing inflation. In our example, this test would be based on equation (6) and could be carried out using the Generalized Method of Moments (GMM). The correct orthogonality condition is

$$E_{t-2}\{[\pi_t + (1/c_2)(\pi_{t+2} + c_1\pi_{t+1} + c_3\pi_{t-1} + \psi_{t+2} + a_1\psi_{t+1} + a_2\psi_t + a_3\psi_{t-1} + a_4\psi_{t-2})]s_{t-2}\} = 0, \quad (8)$$

where in this example the instruments should be the twice-lagged states from the economic model, s_{t-2}. Thus, under the null of a generalized Taylor price-setting model, the equation describing inflation should be tested using a fairly complicated orthogonality condition that includes lags of marginal cost and inflation. Again, in performing the test one should use the actual states as instruments. In practice, a Calvo-type model is tested with instruments that are not the true states. The actual set of state variables is not used in the test because the econometrician does not have access to a time series on the past prices set by adjusting firms or the various economic shocks. Thus, under the null of generalized Taylor-style price setting, the tests commonly used to determine whether forward-looking price setting explains the behavior of inflation are misspecified. Relevant variables and restrictions are omitted, and the instrument set is incorrectly specified.

Testing the Calvo Model

To analyze the potential consequences of model misspecification, I investigate the empirical results when tests that assume the underlying model is of the Calvo variety are conducted on data generated by a generalized Taylor price-setting model. I perform two sets of estimates, one based on a sample of 25,000 observations, referred to as the population estimates, and the other based on 500 simulations involving samples of 200 observations, referred to as the finite sample estimates.

Based on equation (3), the orthogonality condition is

$$E_t\{(\pi_t - \lambda\psi_t - \beta E_t\pi_{t+1})z_t\} = 0,$$

where z_t is an instrument vector containing three lags each of inflation, labor share, and output, and, as described above, λ is a combination of the time preference parameter β and the probability that a firm will not be able to reset its price, θ.[9] The population estimates of these two parameters are 0.58 and 0.35, whereas the average finite sample estimates are 0.56(0.24) and 0.36(0.035), with standard errors in parenthesis. The estimate of β is well below its true value of 0.99 and is also substantially less than that estimated by Gali and Gertler (their estimate is 0.926).[10] The estimate of θ implies a mean lag in the Calvo model of roughly 1.5 quarters, which is smaller than the true mean lag of 2.4 quarters. Gali and Gertler's estimate of θ implies a rather long mean lag of 8.6 quarters and indicates that three-period staggering is insufficient for capturing the underlying price stickiness in the U.S. economy. Restricting the coefficient on β to one only slightly affects the estimate of θ. In population the estimate is 0.37 and in sample it is 0.38(0.038).

Estimating a Calvo model when the true model involves three-period Taylor-type contracts implies both a misspecification and that the instruments are correlated with the error term. The correlation arises because the true error term includes two lags of marginal cost, as well as expectational errors of future inflation and marginal cost that are based on information up to two periods ago. This correlation is confirmed by the rejection of the overidentifying restrictions. This rejection of the orthogonality of the instruments also occurs when actual data is used. Although Gali and Gertler indicate that their instruments pass the test for overidentification, that result appears to be due to the choice of a number of poor instruments. When I perform the above estimation on their data, using a set of instruments similar to the one used in

[9] The specification in terms of structural parameters is $E_t\{(\theta\pi_t - (1-\theta)(1-\beta\theta)\psi_t - \theta\beta E_t\pi_{t+1})z_t\}$ and corresponds to specification 1 in Gali and Gertler (1999).

[10] If I use Gali and Gertler's method 2, the population estimates for β and θ are 0.72 and 0.47, respectively.

testing model data, I replicate their point estimates almost exactly.[11] However, the model fails the test for instrument orthogonality at 10 percent significance levels.[12] The analysis presented in this article indicates that the failure may be a result of underlying price behavior that conforms in fact more closely with a staggered price-setting model.

Testing the Hybrid Model

I will now test to see if lagged inflation is statistically significant when added to the Calvo specification. From equation (6), which describes the behavior of inflation in the true model, one would expect lagged inflation to be significant in the estimation. However, because the coefficients c_2 and c_3 are both negative, one might expect the coefficient on lagged inflation to be negative. The orthogonality condition in the GMM estimation is

$$E_t\{(\pi_t - \lambda\psi_t - \gamma_f E_t\pi_{t+1} - \gamma_b\pi_{t-1})z_t\} = 0,$$

where $\lambda = (1-\varpi)(1-\theta)(1-\beta\theta)/\phi$, $\gamma_f = \beta\theta/\phi$, $\gamma_b = \varpi/\phi$, and $\phi = \theta + \varpi(1-\theta(1-\beta))$. The population estimates of β, θ, and ϖ are 0.60, 0.36, and 0.10, respectively, and the finite sample estimates are 0.60(0.24), 0.37(0.038), and 0.13(0.073). The latter estimates imply a value of $\lambda = 0.94$, $\gamma_f = 0.47$, and $\gamma_b = 0.25$. The value of γ_b is exactly the same as that found by Gali and Gertler on U.S. data.

The positive coefficient on the lagged inflation term occurs because the error term in the regression includes not only expectational errors, but also lagged marginal cost terms and a term involving two period leads of inflation and marginal cost. Further, when the instrument set is insufficiently lagged, the expectational errors will also be correlated with the instruments. Thus, the coefficients in the regression will be biased. The bias involves complicated terms arising from the relationships between the instruments and the explanatory variables as well as from the correlations between the omitted variables that appear in the error term and the variables in the regression. Regarding the latter, the correlation between lagged marginal cost and lagged inflation is 0.75 and between twice-lagged marginal cost and lagged inflation is 0.55. If one estimates the linear relationship implied by the above orthogonality condition, ignoring the relationship between $(\lambda, \gamma_b, \gamma_f)$ and (β, θ, ϖ), then it turns out that γ_f is biased downward and the other two coefficients are biased upward. Thus, the misspecification inherent in the Calvo model implies a downward

[11] Using method 2, my estimate of β is 0.965 compared to their estimate of 0.941, and my estimate of θ is 0.895 while theirs is 0.884.

[12] My instrument set is three lags of inflation, labor share, and output growth.

bias in the estimated importance of forward-looking behavior and an upward bias in the importance of backward-looking behavior.[13]

Fundamental Inflation

I now compute what is termed *fundamental inflation* in order to analyze how well inflation predicted by the estimated model matches inflation generated by the theoretical model. Using the estimates from the regression with once-lagged instruments, I can calculate fundamental inflation (inflation that is generated entirely by the pricing equation of the model) as in Gali and Gertler (1999) by solving difference equation (7). One eigenvalue of this difference equation, δ_1, is less than one while the other, δ_2, lies outside the unit circle. The solution for fundamental inflation is

$$\pi_t = \delta_1 \pi_{t-1} + \frac{\lambda}{\delta_2 \gamma_f} \sum_{k=0}^{\infty} \left(\frac{1}{\delta_2}\right)^k E_t \psi_{t+k}.$$

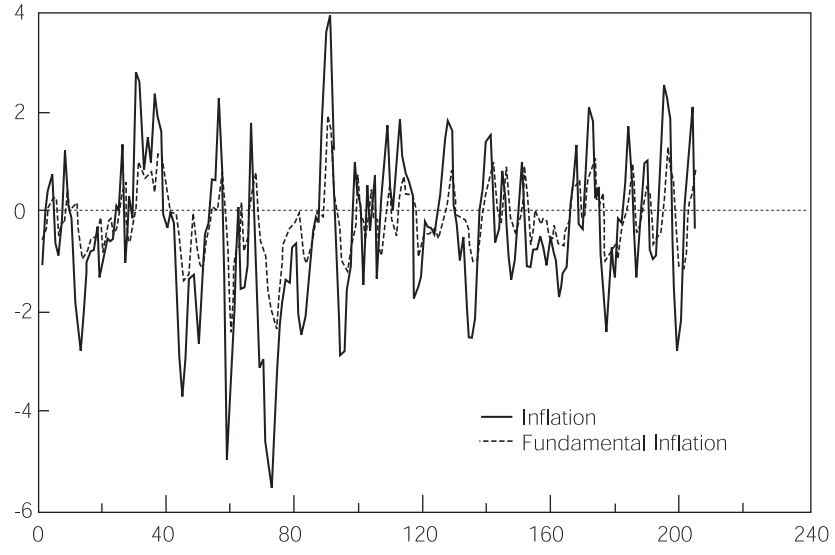
To calculate the summation term, I estimate a Vector Autoregression (VAR) describing inflation and marginal cost on a typical simulation of the model. The simulation in question produced estimates of β, θ, and ϖ of 0.61, 0.36, and 0.095, respectively. The VAR included four lags of each variable, and the forward-looking sum was derived from the estimated equations. Figure 1 depicts the results. Fundamental inflation explains much of the actual movement in inflation, and there is no evidence of systematic bias. The correlation between fundamental and actual inflation is 0.71. In calculating fundamental inflation, it is important to note that even if the coefficient on lagged inflation is small, backward-looking behavior may be important for the dynamics of inflation. The importance arises because the dynamics are governed by the eigenvalues, which are in turn functions of all the underlying parameters.

3. CONCLUSION

This article critically examines the common interpretation of a finding that lagged inflation helps explain the behavior of current inflation. The common interpretation is that some departure from optimality exists in the pricing behavior of firms. A popular explanation of this departure involves the presence of backward rule-of-thumb behavior by some fraction of firms, but irrational forecasting of expected inflation is sometimes also invoked as an explanation. Here, I use a generalized Taylor pricing model as a data-generating mechanism and show that incorrectly basing tests on pricing behavior of the type described

[13] This type of bias may also be present in Fuhrer (1997) because his model fails to account for sufficient lags of the output gap.

Figure 1



by a Calvo model can produce significant coefficients on lagged inflation even though all firms are rational and forward looking. Thus, the interpretation of a significant coefficient on lagged inflation in a pricing equation may be more subtle than is currently realized.

APPENDIX

I derive the underpinnings of equation (6) for an economy that has zero inflation. Let α_j denote the probability that a firm that last changed its price $j \leq 3$ periods ago changes its price in the current period, and let $\alpha_3 = 1$. Defining $\eta_j = 1 - \alpha_j$, then the fraction of firms that change their price in the current period, $\omega_0 = 1/(1 + \eta_1 + \eta_1\eta_2)$, the fraction that last changed their price one period ago, $\omega_1 = \eta_1/\ (1 + \eta_1 + \eta_1\eta_2)$, and the fraction that last changed their price two periods ago, $\omega_2 = \eta_1\eta_2/\ (1 + \eta_1 + \eta_1\eta_2)$. The price level under generalized Taylor pricing is given by

$$P_t = \left[\sum_{j=0}^{2} \omega_j P_{t-j}^{*(1-\epsilon)} \right]^{\frac{1}{1-\epsilon}},$$

where variables are in levels and ϵ is the elasticity of demand for the firm's product. The optimal price-setting equation is

$$p_t^* = \frac{\varepsilon}{\varepsilon - 1} \frac{\sum_{j=0}^{2} \beta^j E_t\{(\omega_j/\omega_0) \cdot (\lambda_{t+j}/\lambda_t) \cdot \psi_{t+j} \cdot (P_{t+j}/P_t)^{\varepsilon} \cdot y_{t+j}\}}{\sum_{j=0}^{2} \beta^j E_t\{(\omega_j/\omega_0) \cdot (\lambda_{t+j}/\lambda_t) \cdot (P_{t+j}/P_t)^{\varepsilon-1} \cdot y_{t+j}\}},$$

where y is the firm's level of output and $\beta^j E_t(\lambda_{t+j}/\lambda_t)$ is the rate at which profits are discounted. For more detail concerning the derivation of these two equations, see Dotsey, King, and Wolman (1999).

Log-linearizing the expression for the price level around zero steady state inflation yields $p_t = \omega_0 p_t^* + \omega_1 p_{t-1}^* + \omega_2 p_{t-2}^*$, which is (4) in the text. Log linearizing the equation for the optimal price yields

$$p_t^* = \rho_0(\psi_t + p_t) + \rho_1 E_t(\psi_{t+1} + p_{t+1}) + \rho_2 E_t(\psi_{t+2} + p_{t+2}),$$

which is (5), where $\rho_0 = 1/\Delta$, $\rho_1 = \beta\eta_1/\Delta$, and $\rho_2 = \beta^2\eta_1\eta_2/\Delta$, and $\Delta = 1 + \beta\eta_1 + \beta^2\eta_1\eta_2$. The linearization turns out to be so compact because at zero inflation many of the terms cancel out (for a general derivation, again see Dotsey, King, and Wolman [1999]).

Combining (4) and (5) for the prices p_t^*, p_{t-1}^*, and p_{t-2}^* yields the following difference equation:

$$\{1 + a_1 L + (a_2 - 1/(\omega_0\rho_2))L^2 + a_3 L^3 + a_4 L^4\}E_{t-2}p_{t+2} =$$
$$\{1 + a_1 L + a_2 L^2 + a_3 L^3 + a_4 L^4\}E_{t-2}\psi_{t+2},$$

where $a_1 = (1 + \beta\eta_1)/(\beta\eta_2)$, $a_2 = (1 + \beta\eta_1^2 + \beta^2\eta_1^2\eta_2^2)/(\beta^2\eta_1\eta_2)$, $a_3 = (1 + \beta\eta_1\eta_2)/(\beta^2\eta_2)$, and $a_4 = 1/\beta^2$. One of the roots of the polynomial on $E_{t-2}p_{t+2}$ is one, and factoring this root yields (6),

$$\{1 + c_1 L + c_2 L^2 + c_3 L^3\}E_{t-2}\pi_{t+2} =$$
$$-\{1 + a_1 L + a_2 L^2 + a_3 L^3 + a_4 L^4\}E_{t-2}\psi_{t+2},$$

where $c_1 = 1 + (1 + \beta\eta_1\eta_2)/(\beta\eta_2)$, $c_2 = -(1 + \eta_2 + \beta\eta_1\eta_2)/(\beta^2\eta_2)$, and $c_3 = -1/\beta^2$.

REFERENCES

Ball, Lawrence J. 1994. "Credible Disinflation with Staggered Price-Setting." *American Economic Review* 84 (March): 282–89.

Calvo, Guillermo A. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics* 12 (September): 383–98.

Dotsey, Michael, and Robert G. King. 2001. "Pricing Production and Persistence." NBER Working Paper 8407 (August).

———————, and Alex L. Wolman. 1999. "State Dependent Pricing and the General Equilibrium Dynamics of Money and Output." *Quarterly Journal of Economics* 114 (May): 655–90.

Fuhrer, Jeffrey C. 1997. "The (Un)Importance of Forward-Looking Behavior in Price Specifications." *Journal of Money, Credit and Banking* 29 (August): 338–51.

Gali, Jordi, and Mark Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44 (October): 195–221.

Guerrieri, Luca. 2001. "Inflation Dynamics." Manuscript, Board of Governors of the Federal Reserve System.

Kiley, Michael T. 1998. "Partial Adjustment and Staggered Price Setting," Manuscript, Board of Governors of the Federal Reserve System.

Roberts, John M. 2001. "How Well Does the New Keynesian Sticky-Price Model Fit the Data?" Manuscript, Board of Governors of the Federal Reserve System.

Rotemberg, Julio J. 1982. "Sticky Prices in the United States." *Journal of Political Economy* 60 (December): 1187–281.

Rudd, Jeremy, and Karl Whalen. 2001. "New Tests of the New-Keynesian Phillips Curve." Manuscript, Board of Governors of the Federal Reserve System.

Sbordone, Argia M. 1998. "Prices and Unit Labor Costs: A New Test of Price Stickiness." Manuscript, Rutgers University.

———————. 2001. "An Optimizing Model of U.S. Wage and Price Dynamics." Manuscript, Rutgers University.

Taylor, John B. 1980. "Aggregate Dynamics and Staggered Contracts." *Journal of Political Economy* 88 (February): 1–24.

Wolman, Alexander M. 1999. "Sticky Prices, Marginal Cost, and the Behavior of Inflation." Federal Reserve Bank of Richmond *Economic Quarterly* 85 (Fall): 29–48.

Imperfect Competition and the Pricing of Interbank Payment Services

John A. Weinberg

In a modern economy, a large fraction of payments for goods and services involve the services of one or more banks. The provision of payment services is, in fact, one of the distinguishing characteristics of banks. A bank-intermediated payment instrument, such as a check, typically communicates instructions to the buyer's bank to make payment to the seller or the seller's bank. Often, then, we think of payment services as being bundled with the deposit services provided by banks, although this is not always the case. Credit cards, for instance, involve payments by the card-issuing bank, at which the cardholder need not hold deposits. Still, many payment services do arise naturally as byproducts of holding deposits with a bank, and some authors have recently begun to focus on this payments function in the theory of banking (McAndrews and Roberds 1999; Prescott and Weinberg 2000). Checks and debit cards are prominent examples of such payment services, but ATM service, which gives people remote access to cash from their deposit accounts, is also an example even though ATM transactions facilitate purchases with cash, not bank liabilities. Accordingly, the industrial organization of the payment services industry, and even the characteristics of the payment services provided, will generally depend on the organization of the banking industry itself.

One area in which the structure of the banking industry matters for the nature of payment services is that of interbank payments—payments in which

This article derives from work completed while I was a visiting scholar at the Bank of Japan's Institute for Monetary and Economic Studies. I thank the Institute staff for their hospitality and assistance and for their willingness to discuss these issues with me. I also thank Mike Dotsey, Huberto Ennis, Ned Prescott, and Tom Humphrey for helpful comments. The views expressed herein are the author's and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

the services of more than one bank are required. In an economy where banking is dominated by a very few institutions, there may be a relatively large number of transactions in which the buyer and the seller have deposits with the same bank. In these cases, interbank payments are not necessary. On the other hand, if there are many banks and people frequently engage in transactions with customers of diverse institutions, then many payments will be interbank payments, requiring the services of both the buyer's and the seller's banks. In these cases each bank provides services to both its own and the other bank's depositors. The interbank payment services that one bank provides to another's depositors resemble the interconnection services that allow customers of one communication network to connect with those of a second network.[1]

In an environment in which banks compete for depositors, the terms on which they make interbank services available can be powerful strategic tools. By making interconnection very costly, a bank could dissuade potential depositors from placing deposits with competing banks. Such surcharges for interbank services create inefficiency by exceeding the resource costs of providing those services. Hence, there is potential tension between a bank's need to compete for depositors and its need to cooperate in interconnection in order to enhance the quality of its service. The implications of this tension depend on the market structure and the nature of competition in the banking industry. In a perfectly competitive, or perfectly "contestable" market, there is strong reason to expect efficient outcomes, even taking the network characteristics of payment services into account (Weinberg 1997). Competition, however, may be limited by regulatory or other features of the economic environment. For instance, in the United States, banking was historically segmented along geographic lines by an array of branching restrictions. In Japan, such segmentation was perhaps even more extensive, with different classes of institutions having specified sets of services or classes of customers they could serve (Ito 1992). In both of these countries, old barriers between market segments have eroded, increasing the opportunities for direct competition among a wider array of banks. Still, in many economies, limits to competition may remain.

In a segmented environment, the terms of interbank payment arrangements can have at most a limited effect on the competition among banks for depositors. The main concern in such an environment, then, would be the provision of the common resource represented by a comprehensive interbank network. Any conflicts of interest among banks would be mainly related to differences in the value that different banks placed on having access to such a network. For instance, in a banking system in which local clearing houses play an important role in payments within a region, a primary role of an interregional network is to connect the various clearing houses. Accordingly, banks that serve limited geographic areas will be most interested in the services of an

[1] Laffont and Tirole (1998a,b) study interconnection pricing among rival telecommunication networks. McAndrews (1998) applies their model to an interbank setting.

interregional interbank network. In Japan, where there have traditionally been both regional and nationwide banks, the value of participating in an interbank, interregional network is likely smaller for the latter than for the former. A large, nationwide bank could use its own internal branch network to make connections among the various clearing houses. A bank with geographically limited operations, however, could benefit from having access to a national network. Indeed, in the 1940s, when the national clearing system was first established, regional banks took the greatest interest in its development, according to Tsurumi (1999). By way of contrast, the United States had no nationwide banks at the time of the founding of the Federal Reserve System, and the banks that lobbied for the Fed to create a national network for clearing interregional checks were primarily large banks in large cities (Lacker, Walker, and Weinberg 1999).

The pricing of interbank payment services typically falls into one of two broad categories: cooperative or independent. Independent pricing simply means that each bank sets the price for its own interbank services. Fees for clearing and settling checks were set independently in the period prior to the Fed's dominance of the check-clearing system. A contemporary example is the surcharge imposed by a bank for an ATM withdrawal by another bank's depositor. Cooperative price-setting usually takes the form of interbank prices set by a group or consortium of banks. Interchange fees in a payment card or ATM network are examples of this sort of cooperation.

It is generally accepted that cooperation among competing firms in the setting of prices can enable sellers to achieve higher prices and profits than they could obtain with independent pricing. This increase in profits comes at the expense of consumer welfare and economic efficiency. This principle is, of course, the basis for antitrust policy. On the other hand, when firms with some market power sell complementary goods, then cooperation can result in lower prices. When it bundles deposit and payment services, a bank is selling products that are both substitutes for and complements to the products of its rivals. This combination complicates the evaluation of cooperative price-setting.

This article explores the differences between cooperative and independent setting of interbank prices in alternative market environments. I focus specifically on the pricing of interbank payment services when deposit markets are segmented, as compared to when banks compete directly for deposits. An important qualification of this discussion is that it takes the structure of the banking market, given by the set of potential market participants, as fixed. That is, I do not consider the effects of free entry or potential competition. In essence, then, the article explores how changes in the degree of imperfect competition affect the comparison of cooperative and independent pricing.

I address the question of how interbank pricing might respond to a change in the competitive environment in which banks operate, first in fairly general

terms and then in the context of a simple model of bank competition and interconnection. The main insight drawn from this discussion can be summarized as follows. Cooperation in the setting of interbank prices typically leads to lower interbank prices and greater consumer welfare and profits when deposit markets are segmented. On the other hand, when banks compete directly for deposits, cooperation in setting interbank prices can have the effect of dampening competition in the deposit market, given a fixed set of competitors. This could result in higher interbank prices and reduced consumer welfare.

1. THE ELEMENTS OF AN INTERBANK PRICING GAME

I begin by describing a model of price competition between two banks facing demands for deposits and interbank payment services. The demand structure specified below can be derived from a more detailed economic environment involving the need for agents to engage sometimes in storage and consumption activities at physically distinct locations.[2] The same general structure would arise in any economic environment in which a diverse set of buyers and sellers of goods and services acquire both deposit and transaction services from potentially competing banks. While models adapted from the telecommunications literature can fit into this framework, the general structure allows for some additional important features. Specifically, and as will be shown by the example in Section 2, this framework can accommodate differences between competing banks. This is a useful feature since many discussions of competition among banks focus on the relative competitive positions of large and small banks.

Demand Functions and Prices

Consider a market in which two banks raise deposits that can be used to make payments in the purchase of goods. To be concrete, focus on the market for household deposits and the payment services provided by banks to households for making purchases from firms. Also, in the interest of simplicity, suppose that firms are exogenously assigned to banks, some with each bank. A consumer selects a bank at which to deposit; the consumer's choice will affect the set of firms to which it can costlessly make payments. If we assume that consumers are randomly matched with firms for the purpose of making purchases, then each consumer faces some chance that he or she will need to make a purchase from a firm that does not use his or her bank. Completion of such a transaction will require an interbank transaction, in which the firm's bank credits the firm's account and collects funds from the consumer's bank.

[2] See Weinberg (2000).

In order to examine interbank pricing and competition, the description of this market structure must specify demand functions for bank services. Labeling the banks 0 and 1, let z_i represent the number of depositors attracted by bank i, and let x_i represent the number of interbank transactions entered into by a customer of bank i.[3] These quantities will respond to the prices set by the two banks. Assume that each bank sets two prices, one price for deposit services that also covers all same-bank payments and another price for each interbank transaction. Let bank i's price for deposit services be denoted by p_i and let its price for an interbank transaction be given by q_i. More precisely, bank i collects p_i from each consumer who places deposits with it and collects q_i from the other bank's depositors for each purchase they make from firms that use bank i. A more general pricing structure would allow a bank to charge transaction fees to its own depositors as well as to its rival's depositors. The simpler structure specified here is sufficient to capture the combination of complementarity and substitutability that banks face when they make pricing decisions.[4]

The actions of the banks and their customers proceed through three stages. First, banks announce their prices. Next, depositors choose with which bank to place their funds. Finally, depositors make purchases. If a depositor wishes to make a purchase from a seller that is associated with a different bank, then the depositor needs the services of the seller's bank to complete the purchase. The depositor values both the consumption of goods purchased and the deposits left over after buying goods and paying bank fees.

In general, one can assume that both the number of depositors a bank attracts and the number of interbank transactions it services will be functions of the full set of prices, (p_0, p_1, q_0, q_1). In choosing a bank, a depositor will weigh the value of depositing at bank 0 against the value of depositing at bank 1 and against the value of not using banking services at all. The value to a depositor of placing deposits with bank i depends on (p_i, q_j). The demand for deposits at bank 1 (z_1) is decreasing in p_1 and q_0 and either independent of or increasing in p_0 and q_1.[5] An increase in q_0 causes this demand to fall because q_0 is the price paid by bank 1's depositors when they must make a purchase from a customer of bank 0. The dependence of z_1 on p_0 and q_1 is determined by the degree to which the two banks' markets for deposits are segmented. Segmentation of the markets could be the result of fundamental demand characteristics, such as the degree to which consumers find the deposit services of the two banks to be good substitutes. Market segmentation could

[3] This specification treats z_i as both the number of depositors and the value of deposits attracted. Hence, each consumer is assumed to have one unit of funds available for deposit.

[4] The interaction between interbank pricing and the pricing of services to one's own depositors under more general pricing structures is qualitatively similar to that presented in this article.

[5] The treatment of the demand facing bank 0 is symmetric to that for bank 1.

also arise from artificial barriers to competition, such as legal rules that limit the set of consumers a particular bank (or type of bank) may serve.

If there are no consumers who could reasonably choose to bank at either bank, then the markets are fully segmented and p_0 and q_1 will have no effect on z_1. If, on the other hand, the two banks compete directly for at least some customers, then z_1 is increasing in p_0 and q_1, which determine the cost of depositing with bank 0.

For a given depositor at bank 1, the demand for interbank transactions depends only on q_0, the price charged for such transactions. This is due to the assumed timing of the depositor's decisions. When the depositor makes a consumption decision (chooses x), deposits have already been placed with a bank. Hence, the depositor's only remaining decision is to weigh the marginal utility of consumption against its price. Here, the price of consumption is either zero (if the depositor is buying from a seller who uses the same bank as the depositor) or q_j for an interbank purchase (from a seller that uses the other bank [bank j]). The total quantity of interbank transactions on which bank 0 collects q_0 is $z_1 x_1$. The banks' profits can be written as $\Pi_1 = z_1 p_1 + z_0 x_0 (q_1 - c)$ and $\Pi_0 = z_0 p_0 + z_1 x_1 (q_0 - c)$, where c is the cost to the bank of processing and collecting on an interbank payment.[6]

Pricing Behavior

The banks set prices for payment and deposit services to maximize their profits, each taking the other's prices as given. Consider, for instance, bank 1's profit maximization problem. Its first order conditions are[7]

$$\begin{aligned}
\frac{\partial \Pi_1}{\partial p_1} &= z_1 + \frac{\partial z_1}{\partial p_1} p_1 + \frac{\partial z_0}{\partial p_1} x_0 (q_1 - c) = 0; \text{ and} \\
\frac{\partial \Pi_1}{\partial q_1} &= z_0 x_0 + \frac{\partial z_1}{\partial q_1} p_1 + \left(\frac{\partial z_0}{\partial q_1} x_0 + \frac{\partial x_0}{\partial q_1} z_0\right)(q_1 - c) = 0.
\end{aligned}$$

These two equations can be rewritten as

$$\begin{aligned}
1 + \eta^1_{p_1} + \eta^0_{p_1} \frac{z_0 x_0 (q_1 - c)}{p_1 z_1} &= 0; \text{ and} \\
1 + \mu_1 (\eta^0_{q_1} + \epsilon^0_{q_1}) + \eta^1_{q_1} \frac{z_1 p_1}{z_0 x_0 q_1} &= 0,
\end{aligned}$$

where η^i_j is the elasticity of z_i (demand for deposits at bank i) with respect to price j; ϵ^i_j is the elasticity x_i (demand for interbank payment services from

[6] The profit functions reflect the assumption (for simplicity) that variable costs of deposit services are zero.

[7] Similar conditions hold for bank 0.

bank i) with respect to price j; and $\mu_i = \frac{q_i - c}{q_i}$ is the percent mark-up of bank i's interbank price over marginal cost.[8]

The conditions above capture the typical result that a profit-maximizing price is inversely related to the relevant (own-price) demand elasticities. The first condition indicates that, in addition to the price elasticity of its own deposit demand, a bank's choice of a price for its deposit services also depends on the cross-price elasticity of the other bank's deposit demand. This dependence arises because the bank earns profits by providing interbank payment services to its rival's depositors. Since deposits at the two banks are substitute services, own-price and cross-price elasticities have opposite signs; raising bank 1's own deposit price increases bank 0's deposits, thereby increasing bank 1's interbank services. The effect is to amplify a bank's desire to raise deposit prices, other things being equal. The magnitude of this effect depends on the relative contributions that payment services and deposit services make to a bank's profits.

A similar interpretation can be given to the second condition above. When setting its price on payment services, a bank considers both the direct effect of the price on its own sale of payment services and the indirect effect on its sale of deposit services. The latter results because bank 1's payment services are complementary to bank 0's deposit services, which are substitutes for bank 1's own deposit services. Again, the strength of the indirect effect depends on the relative contributions the two services make to a bank's overall business.

Segmented Markets

The joint solution of the two banks' problems and the nature of the interaction between prices of interbank payment services and prices of basic deposit services depend on the nature of competition between the banks. In part, the nature of interbank rivalry is determined by the structure of the banks' external environment. In particular, the degree of integration or segmentation of markets determines whether the banks come into face-to-face competition with each other. This characteristic of the markets is captured by the demand functions, and the degree of segmentation is represented by the values of the elasticities $\eta^i_{p_j}$, for $i \neq j$, and $\eta^i_{q_i}$. These elasticities reflect the responsiveness of a bank's deposits to the other bank's deposit price and to its own interbank payment price. Recall that a bank's interbank price is paid by the other bank's depositors. Hence, q_1 will affect z_1 only if banks 1 and 0 compete directly for customers. When the deposit markets are segmented, $\eta^0_{p_1} = \eta^1_{p_0} = \eta^0_{q_0} = \eta^1_{q_1} = 0$.

[8] It is common to express own-price elasticities as absolute values. Here, $\eta^i_{p_i}$ and $\epsilon^i_{q_i}$ are defined as negative numbers. This seems convenient, as own-price elasticities are combined in some expressions with cross-price elasticities, which may be negative or positive.

When markets are segmented, then the first order conditions above reduce to

$$\begin{aligned}\frac{\partial \Pi_1}{\partial p_1} &= z_1 + \frac{\partial z_1}{\partial p_1} p_1 = 0; \text{ and} \\ \frac{\partial \Pi_1}{\partial q_1} &= z_0 x_0 + (\frac{\partial z_0}{\partial q_1} x_0 + \frac{\partial x_0}{\partial q_1} z_0)(q_1 - c) = 0.\end{aligned}$$

Or, in terms of elasticities,

$$\begin{aligned}1 + \eta^1_{p_1} &= 0; \text{ and} \\ 1 + \mu_1(\eta^0_{q_1} + \epsilon^0_{q_1}) &= 0.\end{aligned}$$

Note that even when markets are segmented, one bank's pricing is not entirely independent of the other bank's prices. Each bank's deposit demand depends on its own deposit price and the other bank's payment service (interconnection) price. That is, z_1 depends on p_1 and q_0. Still, under segmented markets, a bank's pricing of its own deposit services does not interact directly with its pricing of interbank payment services.

In the case of segmented markets, one bank's deposit services are complementary to the other bank's interbank payment services. For instance, an increase in q_0, bank 0's payment service price, reduces the value to potential customers of placing deposits at bank 1. The price increase will generally result in lower demand for bank 1 deposits and lower profit-maximizing value of p_1, bank 1's deposit price. At the same time, an increase in p_0 reduces the total value of deposits bank 0 is able to attract and correspondingly reduces the volume of interbank transactions on which bank 1 can extract a fee. This reduction in demand results in a lower optimal choice of q_1.

When two sellers set the prices of complementary goods noncooperatively, the outcome is often characterized as a problem of "double-marginalization." In effect, the two goods can be thought of as a single service with two distinct components. If both components were sold by a single seller with market power, that seller would recognize the effect of each component's price on the sale of both components. This interdependence limits the seller's interest in raising prices. When the components are sold separately by different firms, each seller is interested in only its own profits and ignores the effects of its price on the other seller's sales. As a result, the distortion due to the deviation of price from marginal cost is compounded by the independent profit-maximizing behavior of two sellers with market power. This compound distortion comes at the cost of both combined seller profits and consumer welfare. Hence, the independent pricing of complementary goods resembles decisionmaking in settings with externalities. Each bank ignores the effect of its interbank price on the sales of the other bank, and their noncooperation leads to a loss of efficiency. Unlike losses occurring in the case of a true externality, however,

this loss occurs here only because competition is imperfect and each bank exercises some market power. If for instance there were additional banks whose deposit and payment services were perfect substitutes for those of bank 0, then both p_0 and q_0 would be competed down to marginal cost. The same would be true for bank 1 in the presence of additional competition.

If instead of setting all prices noncooperatively, banks set their prices for interbank services through negotiation, they can raise their combined profits by setting interbank prices (q_0 and q_1) lower than their noncooperative levels. This process is formalized by assuming that q_0 and q_1 are set to maximize joint profits, conditional on the noncooperative determination of p_0 and p_1. The process represents a mixed form of interaction between sellers in which they collude on interbank prices while they compete in the pricing of deposit services. For many specifications of the demand structure, the optimal negotiated choice for interbank prices is to set them equal to marginal cost. This choice eliminates the double marginalization problem, allowing banks to earn their rents from the markup on deposit services. When deposit markets are segmented, cooperation in setting the interbank prices is equivalent to full cooperation in setting all prices, for banks are local monopolists in their deposit market segments.

To see the effects of cooperating in interbank price-setting in segmented markets, consider the first order condition for choosing q_1 to maximize joint profits ($\Pi_0 + \Pi_1$). In terms of elasticities,

$$1 + \eta^0_{q_1} \frac{p_0}{(q_1 - c)x_0} + \mu_1(\eta^0_{q_1} + \epsilon^0_{q_1}) = 0.$$

This cooperative condition has one more term than the corresponding noncooperative condition: $\eta^0_{q_1} \frac{p_0}{(q_1-c)x_0}$. The extra term reflects the effect of bank 1's choice of interbank price q_1 on bank 0's earnings from deposits priced at p_0. The effect of the added term is to reduce the choice of q_1, other things being equal.

In segmented markets, the mechanism for jointly determining interbank prices is not a matter of great importance. Suppose the jointly optimal interbank prices are $q_0 = q_1 = q$. A relatively simple mechanism that will achieve this result is to delegate the choice of a common interbank price to one of the banks. That is, impose symmetry in interbank prices and let the price level be chosen by either of the banks. Suppose this authority is granted to bank 0. Its choice of q_0 does not affect its own profits, but q_1 does. If the demands facing the two banks are symmetric, then bank 0's optimal choice is to set $q_0 = \widehat{q}$. Bank 1 would make the same choice if it were given the authority to set the q's. Hence, with segmented markets and symmetric demands, delegated setting of reciprocal interbank prices achieves the same interbank price as would be set under joint profit maximization, subject to noncooperative choices of deposit prices. This mechanism, then, results in lower interbank prices than would be

chosen independently by the two banks. There are some cases in which this mechanism results in interbank prices that are equal to marginal cost.

When markets are not segmented, the interaction between deposit prices and payment service prices is more complicated. In this case, the interbank prices (q_0, q_1) are a strategic tool in competition for market share. In addition to raising revenue for bank 0, q_0 imposes a cost on bank 1's depositors that, other things being equal, may induce some consumers to deposit at bank 0 instead. To the extent that bank 0 is able to extract price-cost margins from deposit customers that are large relative to markups on payment services, the bank may find it profitable to use a high interbank price to help attract deposits. It is also not the case that cooperation in setting interbank prices will necessarily improve consumer welfare. That cooperation has ambiguous consequences is one of the messages of the literature on interconnection pricing in telecommunications networks. The interbank price could facilitate collusion in deposit pricing by making depositors less likely to switch banks.

It may be reasonable to think of an increase in competition (or more precisely in the potential competitiveness of the market environment) as being captured by a move from segmented markets to a single integrated market. Such a shift could have many causes. Changes in the regulatory or legal environment could bring banks that had previously enjoyed protected market segments into direct competition. Improvements in technology can make it possible for banks to serve expanding sets of customers. For instance, consumer banking may traditionally have been a local business, with people choosing banks based on their proximity to home or place of business. Technological advances allow consumers to make banking choices that are less dependent on location.

If we think of increasing competition as a shift from segmented to integrated markets, then it becomes clear that the role of interbank prices can change in a more competitive environment. With less competition (segmented markets) the interbank price serves mainly as a potential source for double marginalization. Accordingly, cooperation in setting the interbank price is largely beneficial from the point of view of consumer welfare. As markets become more competitive (integrated), the interbank price plays a more complicated strategic role.

Of course, the degree of competition between two banks also depends in part on the behavior of the banks themselves. Is their pricing competitive, in the sense that price determination can be modeled as the Nash equilibrium of a noncooperative game? Or is there some amount of cooperation between the banks in their price-setting behavior? This aspect of the degree of competition is more difficult to tie directly to the demand and cost fundamentals of the market. Rather, the ability of banks to collude depends on such factors as the legal environment. In a setting with strict antitrust enforcement, it will be difficult for sellers of a product or service to engage in explicit or open price

collusion. Even so, tacit collusion may be possible, in the form of cooperation supported by implicit threats to engage in a price war should any seller cheat on the collusive agreement.[9] The feasibility of such collusion depends on factors like sellers' ability to monitor each other's behavior.

The foregoing discussion has assumed that banks behave as Nash price-setters. Under that assumption, the degree of competition is determined by the demand characteristics, as discussed above.

Suppose that banks do collude in the setting of all prices. In that case, prices are set to maximize joint profits, $\Pi_0 + \Pi_1$. In this case, the first order conditions for (for instance) (p_1, q_1) are

$$\frac{\partial(\Pi_0 + \Pi_1)}{\partial p_1} = z_1 + \frac{\partial z_1}{\partial p_1}[p_1 + x_1(q_0 - c)] + \frac{\partial z_0}{\partial p_1}[p_0 + x_0(q_1 - c)] = 0;$$

and

$$\begin{aligned}\frac{\partial(\Pi_0 + \Pi_1)}{\partial q_1} &= \frac{\partial z_1}{\partial q_1}[p_1 + x_1(q_0 - c)] + \frac{\partial z_0}{\partial q_1}[p_0 + x_0(q_1 - c)] \\ &\quad + z_0[\frac{\partial x_0}{\partial q_1}(q_1 - c) + x_0] \\ &= 0.\end{aligned}$$

As with other conditions stated above, these last two can be expressed in terms of demand elasticities as

$$\begin{aligned}1 + \eta^1_{p_1}[1 + \frac{x_1(q_0 - c)}{p_1}] + \eta^0_{p_1}\frac{p_0 z_0 + z_0 x_0(q_1 - c)}{p_1 z_1} &= 0; \text{ and} \\ 1 + \mu_1(\eta^0_{q_1} + \epsilon^0_{q_1}) + \eta^0_{q_1}\frac{p_0}{x_0 q_1} + \eta^1_{q_1}\frac{z_1 p_1 + z_1 x_1(q_0 - c)}{z_0 x_0 q_1} &= 0.\end{aligned}$$

For any given configuration of demand, cooperative price-setting tends to result in higher deposit prices (p's) and lower payment services prices (q's) than does noncooperative pricing. Payment services provide interconnection between banks, allowing one bank's customers to use another bank's facilities. The prices charged for these services, then, are prices charged to another bank's depositors. When prices are set noncooperatively, a bank ignores the effect that raising the payment services price has on its rival's demand and profits. Taking this effect into account initiates cooperation, which results in a moderation of the desire to raise this price. Hence, when banks collude in the setting of deposit prices, either explicitly or implicitly, the role of the interbank price resembles its role in segmented markets.

[9] Green and Porter (1984).

One additional issue regarding tacit (or implicit) collusion involves the role that interbank prices might play in coordinating collusive pricing. Banks must monitor implicit agreements not to engage in aggressive competition in deposit prices, and the monitoring of a rival bank's deposit arrangements with its customers may be difficult compared to monitoring prices of interbank payment services. If, for instance, bank 1 charges a fee to bank 0's depositor, that fee is typically collected through bank 0 (that is, through the interbank clearing and settlement system). Hence, bank 0 will directly observe the fees its customers face from bank 1. The ease of monitoring interbank prices could give them a role to play in the enforcement of broader agreements among banks.

2. AN EXAMPLE

The strategic interaction among banks (or firms in general) in setting interconnection prices can be illustrated by an example in which consumers are assumed to have "home" locations on the "Hotelling" line (Hotelling 1929). That is, each consumer's location is given by a point in the unit interval, $z \in [0, 1]$. There are two banks, located at either end point of the interval. The cost to a consumer located at z of depositing funds at bank 0 is τz, and the cost of depositing at bank 1 is $\tau(1 - z)$. A consumer receives utility W from deposit services and U from payment services. One could interpret W as the balances deposited with the bank. If the consumer is able to use his or her deposit balances to make a purchase of goods from a store, then U will represent the net benefit that the consumer receives from such a transaction. Hence, a "payment service" here might be a transfer of funds from the consumer's account to the store's account. Alternatively, a payment service might be the withdrawal of cash at a cash dispensing terminal close to the place where the consumer will make a purchase. In either case, the net value received by the consumer will be $W + U$ minus fees paid to banks. A consumer also has the option of not depositing funds in a bank. For simplicity, assume that by not using bank services the consumer limits his or her ability to make certain purchases. Specifying the value to the consumer of not depositing funds with a bank as W captures this assumption.

Consumers face uncertainty about where they will want to consume final goods. This uncertainty translates into uncertainty regarding the bank from which the consumer will need deposit services. With probability ϕ, a consumer needs the services of bank 0. This might be interpreted as a consumer's desire to transfer funds to a merchant who banks with bank 0 or as a consumer's need to withdraw funds from a machine owned by bank 0. With probability $(1 - \phi)$, the consumer needs the payment services of bank 1.

Bank i bundles deposit services and payment services to its own depositors under a single price p_i and charges q_i for payment services provided to

the other bank's depositors. The net benefits that a consumer derives from depositing with either bank are given by

$$\begin{aligned} V_0 &= W + U - p_0 - (1-\phi)q_1 - \tau z; \text{ and} \\ V_1 &= W + U - p_1 - \phi q_0 - \tau(1-z). \end{aligned}$$

If, for a given z, the greater of V_0 and V_1 is greater than W, then the consumer deposits with whichever offers the greater value. Let z_i denote the consumer for whom $V_i = W$. Then, the case of segmented markets, as discussed above, is the case in which $z_0 < z_1$. In this case, there is a set of consumers (those between z_0 and z_1) who do not use banking services. Consumers between 0 and z_0 deposit at bank 0, while those between z_1 and 1 deposit at bank 1. Given this specification of demand, banks' profit functions (when markets are segmented) can be written as[10]

$$\begin{aligned} \Pi_0 &= z_0 p_0 + \phi(1-z_1)q_0; \text{ and} \\ \Pi_1 &= (1-z_1)p_1 + (1-\phi)z_0 q_1. \end{aligned}$$

This specification of segmented markets involves a "gap" in the market for banking services that represents consumers who choose not to deposit their funds in the banking system. While there are, in fact, such "unbanked" consumers in many economies (close to 10 percent of all households in the United States), one need not take this specification literally. The choice of interbank prices would be similar in any setting in which a bank's choice of q had no effect on its own deposits. This would be true, for instance, if deposit market segmentation were established by legal or regulatory rules.

Noncooperative price-setting by banks in this example leads to the following Nash equilibrium prices: $p_0 = p_1 = \frac{U}{3}$; $q_0 = \min[\frac{U}{3\phi}, U]$; $q_1 = \min[\frac{U}{3(1-\phi)}, U]$. The reason interbank prices must be less than U is that consumers can always choose not to use interbank services, forgoing the utility U. With these prices, the market division is given by $z_0 = (1 - z_1) = \frac{U}{3\tau}$, so that the two banks have equal market shares.[11]

When the noncooperative equilibrium has this segmented markets characteristic, cooperation in the setting of interconnection prices is equivalent to full cooperation in all prices. This is true because with segmented markets, each bank is a local monopolist in its segment of the deposit services market. Still, cooperation results in a preferred outcome for both banks and consumers.

[10] For simplicity, this example assumes that the marginal costs of both deposit and payment services are zero. Assuming positive marginal costs would not alter the nature of the strategic interaction among banks. However, assuming a higher marginal cost for interbank payment services than for same bank services would add an important dimension to the efficiency properties of equilibrium allocations.

[11] This characterization of the equilibrium assumes that $\tau > \frac{2}{3}U$.

Under this pricing scenario, interbank prices (q_0, q_1) are set equal to marginal cost ($q_0 = q_1 = 0$), and deposit prices are $p_0 = p_1 = \frac{U}{2}$. Hence, deposit prices go up while interbank charges go down. The net effect on consumer welfare is positive, as is demonstrated by the fact that more consumers choose to use bank services under this pricing scenario than under noncooperative pricing. With the cooperative prices, market shares are $z_0 = z_1 = \frac{U}{2\tau}$.

Whether the equilibrium features segmented or integrated markets depends, of course, on the parameters of the model. In particular, U gives the value of having access to payment services, and τ gives the consumer's marginal cost of using bank services. As τ gets smaller or U gets bigger, more consumers will seek to use bank services, and eventually the marginal consumer's decision will be between banks rather than whether to deposit at all. When the market becomes integrated in this way, banks' shares of the market are determined by the point (z), at which a consumer is indifferent between the two banks ($V_0 = V_1 > W$). Denoting this point by $\hat{z}$, we have

$$\hat{z} = \frac{1}{2} + \frac{1}{\tau}[(p_1 + \phi q_0) - (p_0 + (1 - \phi)q_1)],$$

and banks' profit functions are

$$\begin{aligned} \Pi_0 &= \hat{z}p_0 + \phi(1 - \hat{z})q_0, \\ \Pi_1 &= (1 - \hat{z})p_1 + (1 - \phi)\hat{z}q_0. \end{aligned}$$

Under these conditions, banks have a heightened incentive to raise the interconnection price compared to the case of segmented markets. With segmented markets, q_0 has no effect on bank 0's sale of deposit services to its own customers. Here, raising q_0 raises the cost to consumers of depositing with bank 1. When the market is integrated, any loss of depositors by bank 1 is matched by a gain at bank 0. Indeed, in this example the profit-maximizing choice for q_0 and q_1 is $q_0 = q_1 = U$. Deposit prices are then $p_0 = 2\tau + \phi U$ and $p_0 = 2\tau + (1 - \phi)U$.

With an integrated market, it is no longer true that banks can raise their combined profits by agreeing to lower interconnection prices. In particular, each bank's profits are lower if interconnection prices are set at marginal cost. That is, cooperation on interbank prices alone does not tend to drive those prices down to marginal cost. On the other hand, if banks collude on both interbank and deposit prices, then joint profits are maximized by setting interbank prices equal to zero.[12]

[12] Actually, in this example, where consumers end up using either zero or one unit of interbank services, the joint profit maximizing solution determines only the sums $p_0 + (1 - \phi)q_1$ and $p_1 + \phi q_0$. In an extended example, with downward sloping demand for interbank services, joint maximization would drive the interbank prices to marginal cost.

3. CONCLUSION

In many economies, the business of banking is undergoing profound changes. Boundaries between markets, both geographically and in terms of product lines, are being removed by regulatory changes and technological advances. These changes present challenges to traditional ways of handling interbank clearing and settlement arrangements. If the terms for interbank transactions are established by industry-based, collaborative organizations, how will such arrangements respond to the entry of new market participants? This article has suggested that increasing (though still imperfect) competition creates a complicated set of incentives for banks with regard to the terms for interbank payment services. Neither competition nor cooperation in setting these prices is guaranteed to always yield desirable results from the point of view of consumer welfare. This does not necessarily imply the need for a regulatory mechanism in determining interbank prices. The development of such a mechanism, managed by a governmental authority, is subject to its own drawbacks—including, for instance, the difficulty faced by a regulator in obtaining the information necessary to set optimal interconnection prices. Rather than direct regulation, however, there may be call to carefully monitor of industry practices in interconnection pricing. Such monitoring was perhaps less important in an environment with less direct competition among banks. It is somewhat ironic, then, that increasing competition may actually increase concerns for the competitive impact of interbank payment services pricing.

REFERENCES

Green, Edward J., and Robert H. Porter. 1984. "Noncooperative Collusion under Imperfect Price Information." *Econometrica* 52: 87–100.

Hotelling, Harold. 1929. "Stability in Competition." *The Economic Journal* 39 (March): 41–57.

Ito, Takatoshi. 1992. *The Japanese Economy*. Cambridge, Mass.: The MIT Press.

Lacker, Jeffrey M., Jeffrey Walker, and John A. Weinberg. 1999. "The Fed's Entry into Check Clearing Reconsidered." Federal Reserve Bank of Richmond *Economic Quarterly* 85 (Spring): 1–31.

Laffont, Jean J., Patrick Rey, and Jean Tirole. 1998a. "Network Competition I: Overview and Nondiscriminatory Pricing." *Rand Journal of Economics* 29 (Spring): 1–37.

———. 1998b. "Network Competition II: Price Discrimination." *Rand Journal of Economics* 29 (Spring): 38-56.

McAndrews, James. 1998. "Direct Presentment Regulation in Payments." Federal Reserve Bank of Chicago Conference on Bank Structure and Performance, Proceedings.

———, and William Roberds. 1999. "Payment Intermediation and the Origins of Banking." Federal Reserve Bank of Atlanta Working Paper 99-11.

Prescott, Edward S., and John A. Weinberg. 2000. "Incentives, Communication and Payment Instruments." Federal Reserve Bank of Richmond Working Paper 00-11.

Sidak, J. Gregory, and Daniel F. Spulber. 1997. *Deregulatory Takings and the Regulatory Contract.* Cambridge: Cambridge University Press.

Tsurumi, Masayoshi. 1999. "The History of the Payment System in Japan." Manuscript, Faculty of Economics, Hosei University, Tokyo, Japan.

Weinberg, John A. 1997. "The Organization of Private Payment Networks." Federal Reserve Bank of Richmond *Economic Quarterly* 82 (Spring): 25-43.

———. 2000. "Interconnection and Rivalry Between Banks." Wharton Financial Institutions Center Working Paper 00-15.

Recent Developments in Monetary Policy Analysis: The Roles of Theory and Evidence

Bennett T. McCallum

1. INTRODUCTION

Academic thinking about monetary economics—as well as macroeconomics more generally—has altered drastically since 1971–1973 and so has the practice of monetary policy. The former has passed through the rational expectations and real-business-cycle revolutions into today's "new neoclassical synthesis" whereas policymaking has rebounded, after a bad decade following the breakdown of the Bretton Woods system, into an era of low inflation that emphasizes the concepts of central bank independence, transparency, and accountability while exhibiting substantial interest in the consideration of alternative rules for the conduct of monetary policy.[1]

My assignment in this paper is to consider the roles of economic theory and empirical evidence in bringing about these changes—in particular, changes in policy formulation. Have they been driven primarily by theoretical reasoning

H. J. Heinz Professor of Economics, Carnegie Mellon University, and Visiting Scholar, Federal Reserve Bank of Richmond. This article previously appeared in the *Journal of Economic Methodology* 6:2, 1999: 171–98 (http://www.tandf.co.uk), and in *Macroeconomics and the Real World,* vol. 1 (Oxford University Press 2000) and is reprinted here with permission. It was prepared for the conference "Theory and Evidence in Macroeconomics," held 15–17 October 1998, in Bergamo, Italy. For helpful suggestions and criticism, I am indebted to Marvin Goodfriend, Jim Hartley, Bob Hetzel, Bob King, Mary Morgan, Edward Nelson, and Harald Uhlig. The views expressed in this article do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] A different (but not incompatible) account of post-Bretton Woods developments in monetary analysis is provided by Goodfriend and King (1997), who coined the term "new neoclassical synthesis."

or by accumulated evidence? As a related matter, has the evolution reflected health or sickness in the macro-monetary branch of economic science?

In discussing actual monetary policymaking, there is a difficulty stemming from the possibility that in practice policy choices are dominated by responses to current political pressures, with economic reasoning of any form playing a strictly subordinate role in the thought processes of voting members of policymaking bodies such as the United States Federal Open Market Committee. There is reason to believe, however, that economic analysis has been playing an increasing role in monetary policy considerations and, in any event, there would be little for economists to discuss if we were to conclude that actual policy is independent of such analysis. Consequently, most of the discussion below will take writings of central bank economists, together with official publications such as inflation reports, as providing some indication of actual monetary policy practices.

Also, it should be admitted at the outset that evaluation of the relative contributions of theory and evidence is extremely difficult. In fact, a proper quantitative evaluation is probably impossible, since economic science evolves by way of a complicated back-and-forth interaction of theoretical and empirical considerations. Moreover, these considerations are often combined in the work of a single analyst; for example, most of the researchers listed below in Table 1 rely on such a combination in their own work. Consequently, some of this back-and-forth takes place within the minds of individual researchers and thus may not show up at all in the exposition of papers written to report results. Under such circumstances, it is clear that measurement of the relative contributions of theory and evidence must be highly problematic, at best. Accordingly, what is presented in this paper might be regarded more as a number of observations relevant to the issue, rather than as an actual evaluation. My hope is that these observations will shed some light on the evolution of monetary analysis while establishing that both theory and evidence have played important roles.

The outline of the paper is as follows. In Section 2, general analytic trends in macroeconomics will be briefly outlined as a background. Then Section 3 takes up the evolution of monetary policymaking in practice and Section 4 does the same for the formal analysis of monetary policy. Section 5 treats a special topic and Section 6 concludes.

2. TRENDS IN MACROECONOMICS, 1973–1998

The years 1971–1973 make a good starting point for our discussion because they mark sharp breaks in both macroeconomic thinking and in institutional arrangements relevant to the conduct of monetary policy. In terms of institutions, of course I have in mind the breakdown of the Bretton Woods exchange-rate system, which was catalyzed by the United States's decision of August

1971 not to supply gold to other nations' central banks at $35 per ounce. This abandonment of the system's nominal anchor naturally led other nations to be unwilling to continue to peg their currency values to the (overvalued) U.S. dollar, so the par value exchange-rate agreements disintegrated. New par values were painfully established in the December 1971 meeting at the Smithsonian Institution, but after a new crisis in February 1973 the par-value system crumbled in March 1973 and has not been reassembled as of 1998.[2]

In terms of macroeconomics, the years 1971–1973 featured the publication of six papers that initiated the rational expectations revolution. The most celebrated of these, certainly, is Lucas's (1972a) "Expectations and the Neutrality of Money," but his (1972b) and (1973) were also extremely influential as were Sargent's (1971) and (1973). Curiously, however, the first publication to use rational expectations in a macro-monetary analysis was none of these but rather Walters (1971), which has apparently had almost no influence.[3]

At first there was much resistance to the hypothesis of rational expectations, partly because in macroeconomics it was initially associated rather strongly with the policy-ineffectiveness proposition.[4] There were also several other misconceptions, one of which continues today in the argument that it is implausible that all of an economy's agents would believe in the particular model of the economy being used by the analyst.[5] Actually, that is not the assumption required for rational expectations. The latter presumes instead that agents form expectations so as to avoid systematic expectational errors in actuality, which implies that they behave as if they knew the structure of the actual economy. Then expectations will agree with the analyst's model of the economy, but the reason is that this model is by construction the analyst's best attempt to depict the true structure of the economy (otherwise, he/she would use a different model).

Be that as it may, the hypothesis of rational expectations (RE) gradually swept the field in both macroeconomics and microeconomics, a major reason being that it is almost certainly unwise for policy to be conducted under the presumption that any particular pattern of expectational errors will prevail in the future—and ruling out all such patterns implies that expectational errors are orthogonal to information sets (i.e., implies rational expectations). During the late 1970s there was much interest in alternative specifications of price

---

[2] A very brief analysis of the source of the collapse is given below in Section 3.

[3] One reason, perhaps, is that Walters used a different term, namely, "consistent expectations." His paper's first footnote states in part: "What I call consistent expectations is formally similar to Richard Muth's rational expectations." Actually, of course, Richard F. Muth—at the time a leading scholar in the field of housing economics—is the brother of John F. Muth.

[4] On the latter, see McCallum (1980).

[5] A variant is the claim that it is implausible that all agents would believe in the same model of the economy. But, first, this is an objection to macroeconomics, not rational expectations, and second, there are some rational expectations models in which agents' expectations are not all alike.

adjustment behavior, since with RE some but not all forms of price adjustment behavior will lead to policy ineffectiveness. Around 1980, however, such research virtually ceased (which is not to say that work with models including slow price adjustments—e.g., Taylor (1989)—ceased). Other topics involving consumption/saving and labor supply behavior became popular for a while, notable contributions including Hall (1978), Hansen and Singleton (1982), and Mankiw, Rotemberg, and Summers (1985).

Then shortly following the appearance of Kydland and Prescott (1982), the era of real-business-cycle (RBC) analysis began.[6] For the next dozen years, a large fraction of all research by leading macroeconomic analysts involved RBC reasoning or issues in one way or another, pro or con.[7] In standard RBC analysis it is assumed that price adjustments take place very quickly so that, for practical purposes, there is continuous market clearing for all commodities—including labor—in which case monetary policy actions will in most models have little or no effect on real macroeconomic variables at cyclical frequencies. Typically, moreover, the RBC models imply that cyclical fluctuations that are observed in real variables are the consequence of technology shocks, not real shocks to preferences or government fiscal variables. Now, of course this has been a highly controversial hypothesis and I am on record as finding it quite dubious (McCallum, 1986, 1989). But it would be wrong to be altogether negative about RBC analysis because much of it has been devoted to the development of new tools of theoretical and empirical analysis, tools that can be employed without any necessary adherence to the RBC hypothesis about the source of cyclical fluctuations.

In recent years, moreover, these tools have been applied in precisely this fashion. Thus a major movement has been underway to construct, estimate, and simulate models in which agents are depicted as solving dynamic optimization problems and interacting on competitive—or, more often, monopolistically competitive—markets, but with some elements of nominal price or wage "stickiness" built into the structure. The match between these models and actual data is then investigated, often by quasi-RBC procedures, for both real and nominal variables and their interaction. Thus the objective of this line of work is to combine the theoretical discipline of RBC analysis with the greater empirical veracity made possible by the assumption that nominal

[6] This statement oversimplifies greatly, in several respects. First, Kydland and Prescott (1982) was clearly previewed by Kydland and Prescott (1980). Second, there were important early contributions by other RBC analysts, including King, Long, Plosser, and Rebelo—and, as stated below, RBC analysis developed out of earlier work by Lucas, Barro, Sargent, and others. Third, the rise of RBC analysis was somewhat more gradual than the exposition in the text indicates.

[7] A partial exception was work involving unit-root or cointegration analysis, which was quite popular. But this work lay more in the domain of econometrics than macroeconomics, and besides there were prominent issues concerning this topic's relation to RBC analysis—see, e.g., Nelson and Plosser (1982).

prices do not adjust instantaneously.[8] Basically, the attempt is to develop a model that is truly structural, and therefore immune to the Lucas (1976) critique of econometric policy analysis.

The mere description of these developments in macroeconomics makes it apparent that they have been driven by a combination of theoretical and empirical impulses. The rational expectations onslaught was primarily theoretical in origin, building upon recognition of the fact that all other expectational hypotheses permit systematic (hence correctable) expectational errors. But the logical basis for the upsurge of the RBC movement can be viewed as principally empirical.[9] Here the point is that RBC models are in essence equilibrium business cycle models of the type promoted by Lucas (1972a, 1975) but with the monetary shocks eliminated and technology shocks emphasized. And this change in emphasis came about, it can be argued, largely because empirical analysis of various types suggested that the cyclical real effects of monetary policy shocks were in fact very small in relation to the overall variability of output and employment. Some crucial studies providing such evidence were Sims (1980), Litterman and Weiss (1989), Eichenbaum and Singleton (1986), and Nelson and Plosser (1982).[10]

Then there came the more recent movement to incorporate gradual price adjustment—"sticky prices"— into optimizing macro models. By its very nature, the impetus for this movement must have been mainly empirical. For there is no body of theory that tells us that price behavior is sticky; to the contrary it is rather difficult to incorporate sticky prices in a model that stresses optimizing general equilibrium theory.[11]

Thus it has to be the force of evidence that has brought about this important change. Moreover, I think it is only fair to recognize that the RBC movement has itself been strongly concerned with empirical veracity, even though in that regard its preferred measures have been quite different from those used in orthodox time series econometrics.

---

[8] The first of these papers of which I am aware is King (1990). Other notable efforts with publication dates prior to 1997 include Benassy (1995), Cho (1993), Cho and Cooley (1995), Cooley and Hansen (1995), Hairault and Portier (1993), Kimball (1995), King and Watson (1996), Rotemberg (1996), Ohanian, Stockman, and Kilian (1995), and Yun (1996). For references and a useful review, see Nelson (1998). Some more recent studies will be mentioned below.

[9] This statement has been disputed by several readers, and I must confess that the following argument is not entirely straightforward. But I continue to believe that the leaders of the RBC movement were not Keynesians who were won over by arguments for theoretical purity, but instead were adherents of the Lucas-Barro equilibrium approach who discovered that it was very difficult empirically to assign much importance to monetary shocks.

[10] This is not to imply that these studies are immune to criticism; in fact I have quarreled with some of them myself. But the point is that, rightly or wrongly, they were influential.

[11] One reader has suggested that it is illogical for me to cite "evidence" as providing stimulus for the rise and also the decline of RBC analysis. But I contend that this is not illogical, for different types of evidence were predominant during the two phases of intellectual development.

With respect to these measures a few brief words may be appropriate before we move on. The "standard" set of RBC measures was established in the famous Kydland-Prescott (1982) paper, which focused on three sets of second moments for variables that had been "detrended." These were: (i) variances of important real variables including output, labor input, average labor productivity, consumption, investment, and capital; (ii) correlations with output of the other variables listed in (i); and (iii) autocorrelations and, to a lesser extent, lead and lag correlations with output.[12] Thus the RBC empirical verification program has been to assess the conformity of these measures as generated by RBC models with actual values pertaining to quarterly data for the U.S. and other economies.

It has been argued by many analysts that these measures provide an inadequate basis for judging the veracity of a macroeconomic model. One problem is that a model may match the data nicely according to the second-moment measurers (i), (ii), and (iii) and yet fail dramatically to fit the data in other respects, as illustrated by Altug (1989) and Watson (1993). In this regard there now exists a sizeable literature on the topic of "calibration vs. estimation."[13] A different type of concern is that the production function residuals (e.g., Prescott 1986), on which the RBC analysis relies, may not be measures of technology shocks at all, but may instead reflect primarily phenomena of an entirely different origin. Some evidence pointing rather strongly in that direction has been presented by Evans (1992), Basu (1996), and Gali (1997). Also, Cogley and Nason (1995) and others have shown that the dynamic properties of typical RBC models come almost entirely from the properties of the stochastic process assumed to generate the technology shocks, rather than from the modeled behavior of agents.

3. DEVELOPMENTS IN MONETARY POLICY

The 1971–1973 collapse of the Bretton Woods system created, for the first time in history, a situation in which the world's leading central banks were responsible for conducting monetary policy without an externally-imposed monetary standard (often termed a "nominal anchor"). Previously, central banks had normally operated under the constraint of some metallic standard (e.g., a gold or silver standard), with wartime departures being understood to be temporary, i.e., of limited duration. Some readers might not think of the Bretton Woods system as one incorporating a metallic standard, but by

[12] Actually, the lead and lag correlations appeared somewhat later, in Kydland and Prescott (1986).

[13] See, for example, Hoover (1995) and the symposium, with articles by Kydland and Prescott, Hansen and Heckman, and Sims, in the Winter 1996 issue of the *Journal of Economic Perspectives.*

design it certainly was, since the values of all other currencies were pegged to the U.S. dollar and the latter was pegged to gold at $35 per ounce.[14] In practice, United States officials—Treasury and Federal Reserve—did not treat the $35/oz standard as if it were a constraint. This was possible initially because the large devaluation of the dollar relative to gold in 1933–1934 had left the dollar undervalued, so several years of postwar inflation could therefore take place before the dollar became overvalued relative to gold—i.e., until the free market dollar price of gold began to significantly exceed $35/oz. But the effects of these years of mild inflation did gradually accumulate and by 1961 the market price of gold had risen (the value of the dollar had fallen) to about $35/oz. Various patch-up attempts were made to permit the U.S. to continue to conduct policy without conforming to the requirements of the official standard, but another 10 years of slow but steady U.S. inflation generated an unsustainable position—so the system collapsed.

Faced with the responsibility of establishing a monetary standard of their own design, the world's central banks did not perform well at first and inflation reached levels that were unprecedented for a sustained period without any widespread war. Germany and Japan began to get inflation under control by the middle 1970s but it remained high in the other G-7 nations. In the U.S. and the U.K. there was a tendency for central banks to deny that their own behavior was an essential ingredient to the inflation process[15] and considerable importance was attached by central banks to employment, output, and other real macroeconomic objectives. The exact nature of central bank thinking during these years is a matter of dispute,[16] but I am myself inclined to share the judgment of Taylor (1996), who depicts central bankers as acting under the influence of 1960s academic ideas that posited the existence of a long-run and exploitable Phillips-type tradeoff between inflation and unemployment rates.[17]

During the 1970s, there was considerable discussion of policy regimes featuring money growth targets. In Germany, the Bundesbank adopted a monetary targeting strategy that has, with some modifications, been officially employed ever since. The other large-nation central bank that was most successful in avoiding inflation in the late 1970s and 1980s, the Bank of Japan, also apparently gave some emphasis to monetary targets (although in this case the extent

---

[14] The other nations at Bretton Woods would never have agreed to a system based on a paper dollar standard.

[15] An interesting document in this regard is Burns (1979), a speech given in 1975.

[16] One account that is more detailed but basically consistent with the one given here is Goodfriend (1997).

[17] In my opinion it is entirely clear that the above-optimal inflation of the 1970s cannot plausibly be attributed to the time-inconsistency motivation depicted in the famous analysis of Barro and Gordon (1983); for this model requires that central bankers believe that the public forms its expectations rationally. In fact, central bank policymakers and economists both exhibited considerable hostility to the hypothesis of rational expectations until the middle 1980s.

of dedication to this strategy was apparently smaller). In the United States, monetary growth targets were given official status by the Humphrey-Hawkins Act of 1975, but evidently played a rather small role in actual policymaking until October 1979.[18] Then on October 6 the Fed began its so-called "monetarist experiment," i.e., the period (ending in July 1982) during which M1 targets were actively pursued by means of a new operating procedure that featured a nonborrowed reserves instrument. Interest rates quickly rose dramatically, but the effort foundered during 1980 as a result of the selective credit controls that were imposed and then quickly removed. Finally, a period of genuine monetary stringency was begun at the end of 1980 and maintained until the middle of 1982. In response, inflation fell quite rapidly—as did output and employment.[19]

From 1983 until 1990, U.S. inflation fluctuated gently around a midpoint of about 4 or 4.5 percent per year. A monetary tightening during 1989 interacted with the Persian Gulf oil crisis of 1990 to begin another recession that was mild but lengthy. By late 1992, U.S. inflation had declined further to the 2–3 percent range that has persisted since. Whether the Fed was deliberately seeking a reduction in the trend inflation rate during 1989–1990 is a matter of some dispute.

In terms of operating procedures, the Fed gradually reverted after August 1982 to a scheme that centers on the Federal funds rate as its instrument (or "operating target"). In addition, interest rates have come to receive more attention—via the term structure but also long-term rates in an unaugmented state—as indicators of monetary conditions. Thus monetary aggregate growth rates have been downgraded in policymaking significance to the point that the biannual congressional hearings, because they legislatively require reference to these figures, always include a few minutes of distinct awkwardness in the Fed's testimony.

Outside the U.S., a major development has been the emergence of the European Monetary Union. Partly because of the so-called convergence criteria needed to qualify for participation in the single-currency Euro scheme, to be guided by the ECB (European Central Bank), inflation rates across Europe have fallen remarkably, averaging close to 1.0 percent for the most recent years (1996–1998).

Also highly noteworthy has been the arrival of inflation targeting as a new framework for the conduct of monetary policy. Actually, most central banks, among those that are not constrained by formal exchange rate commitments, do not adhere to any clear-cut and announced procedures in conducting monetary

---

[18] Targets were announced for several money stock measures, which often gave conflicting signals, and target misses were treated as irrelevant bygones during the 1970s.

[19] For additional discussion of the 1979–1982 period, including a tabulation of the adjusted M1 growth rates that the FOMC was using at the time (which reveal the tightening during 1981 more clearly than unadjusted values), see Broaddus and Goodfriend (1984).

policy. But of those that have adopted explicit policy frameworks, virtually all have opted for targets expressed in terms of inflation rates, not money stock or nominal income growth rates.[20] Most notable, probably, is the arrangement in New Zealand, which came first and which stipulates that the central bank governor can be removed if the agreed-upon 0–2% inflation target band is not met.[21]

Overall, the most fundamental change since the 1970s has been the assumption of responsibility by central banks for performance in terms of inflation rates. In 1998, it would be extremely surprising to run across a central bank statement that discussed medium-term inflation prospects in a manner suggesting that these are unaffected by monetary policy behavior. So, even though we are here discussing practice and not analysis, one could ask whether theory or evidence has been more responsible for the change in opinion. In this regard there is a "multicollinearity problem" because, as it happens, both theory and evidence have pointed strongly in the same direction, i.e., toward the proposition that there is no permanent stimulus to real variables from monetary leniency so that sustained easy conditions will produce just inflation, without any lasting boost to output or employment. There is of course some formal econometric evidence in this regard,[22] but even more influential to policymakers, probably, was the informal perception of the 1970s as a decade of experience with high inflation accompanied by no enhancement in terms of output and employment. Thus we have theory, formal evidence, and informal "experimental" evidence all pointing in the same direction—toward the idea that from a long-term perspective monetary policy's main influence is on growth of the price level with little or no lasting effect on real output's level or growth rate. From this conception it is a natural step to view inflation prevention as the main macroeconomic duty of a modern central bank, with a secondary objective of dampening cyclical fluctuations, and today's general policy climate falls into place.

4. MONETARY POLICY ANALYSIS

We now turn to the topic of central concern in this paper, analysis of monetary policy arrangements by economists—i.e., by monetary economics specialists in universities, central banks, and other analytical organizations.[23] In that regard it is quite gratifying to report that in recent years there has been a large amount of interaction between central bank and academic analysts, so that

[20] Among the reviews of inflation targeting are Haldane (1995), Leiderman and Svensson (1995), and Bernanke and Mishkin (1997).

[21] Since the election of 1996, the target band has been widened to 0–3 percent.

[22] See, for example, King and Watson (1994).

[23] Such as the IMF or the economic policy institutes of (e.g.) Germany.

today (August 1998) one would be hard-pressed to tell, for many research papers, whether a particular one had been written by members of one group or the other.[24] To illustrate that point, as well as others to be made below, it will be useful to refer to two major conferences held in the first half of 1998. The first of these is an NBER conference on "Monetary Policy Rules" held January 15–17 in Islamorada, Florida, and the second is a Riksbank-IIES conference "Monetary Policy Rules" held June 12–13 in Stockholm. Since the conference titles are the same, they will be referred to below as the NBER and Riksbank conferences. The former was organized by John B. Taylor (Stanford University), the latter by Claes Berg (Sveriges Riksbank) and Lars E.O. Svensson (Institute for International Economic Studies, Stockholm University).

Paper authors, discussants, and panelists at these two conferences are listed in Table 1. It will be noted that there is some overlap in the lists. More importantly, it will be noted that there is substantial participation by both academic and central bank economists in both conferences, especially the Riksbank's.[25] To verify the similarity in concerns and techniques exhibited by central bank and academic authors, the reader is invited to sample the papers themselves. They are forthcoming in an NBER conference volume and an issue of the *Journal of Monetary Economics*; as of August 1998 virtually all the papers could be downloaded from NBER or IIES home pages on the world wide web.

The situation just described is vastly different from that obtaining as recently as the middle 1980s, when academic and central bank economists had much less interaction and much less similarity of viewpoint.[26] If one introspects about reasons for the change, one can easily think of several contenders, among which are some that involve adjustments on the part of both groups. One fact is that several (regional) Federal Reserve Banks have, since the late 1970s, employed academic economists as consultants, a practice that makes each group more familiar with research assumptions held to be essential by the other—e.g., academics have become more knowledgeable about realistic operating procedures while central bank economists have become more comfortable with analysis utilizing rational expectations. Conferences held by Federal Reserve Banks and some by academics (e.g., NBER and Carnegie-Rochester conferences) have brought central bank and academic researchers together more often. Ph.D. graduates of leading universities have taken positions at the Federal Reserve Board and regional Feds and have played crucial

[24] This is a slight exaggeration, since reference to simulation results obtained with Fed models will signal a Federal Reserve author, etc. But the methods and the general characteristics of the models used are extremely similar.

[25] It might also be noted that Alan Blinder and Frederic Mishkin are listed as professors but would recently have been categorized as central bank officials.

[26] In this case one bit of evidence is provided by the interchange between Brunner and Meltzer (1983) and Axilrod (1983).

Table 1 Programs for NBER and Riksbank-IIES Conferences

A	**NBER Conference, 15–17 January 1998**
1	Bennett McCallum and Edward Nelson, Carnegie Mellon Univ., "Performance of Operational Policy Rules in an Estimated Semi-Classical Structural Model." *Discussant*: Mark Gertler, New York Univ.
2	Julio Rotemberg, Harvard University, and Michael Woodford, Princeton Univ., "Interest Rate Rules in an Estimated Sticky-Price Model." *Discussant*: Martin Feldstein, Harvard Univ.
3	Laurence Ball, Johns Hopkins Univ., "Policy Rules for Open Economies." *Discussant*: Thomas Sargent, Stanford Univ.
4	Andrew Haldane and Nicoletta Batini, Bank of England,"Forward Looking Rules for Monetary Policy." *Discussant*: Donald Kohn, Federal Reserve Board
5	Glenn Rudebusch, FRB of San Francisco, and Lars Svensson, Institute for International Economic Studies, "Policy Rules for Inflation Targeting." *Discussant*: James Stock, Harvard Univ.
6	Andrew Levin, Volcker Wieland, and John Williams, Federal Reserve Board, "Are Simple Monetary Rules Robust to Model Uncertainty?" *Discussant*: Lawrence Christiano, Northwestern Univ.
7	John Taylor, Stanford Univ., "An Historical Analysis of Monetary Policy Rules," *Discussant*: Richard Clarida, Columbia Univ.
8	Robert King, University of Virginia, and Alexander Wolman, FRB of Richmond, "What Should Monetary Policy Do When Prices are Sticky?" *Discussant*: Benjamin Friedman, Harvard Univ.
9	Arturo Estrella, FRB of New York, and Frederic Mishkin, Columbia Univ., "The Role of NAIRU in Monetary Policy: Implications of Uncertainty and Model Selection." *Discussant*: Robert Hall, Stanford Univ.
B	**Riksbank-IIES Conference, 12–13 June 1998**
1	Frederic Mishkin, Columbia Univ., "International Experiences with Different Monetary Policy Regimes." *Discussant*: Charles Goodhart, London School of Economics and Bank of England
2	John Taylor, Stanford Univ., "The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank." *Discussant*: Leonardo Leiderman, Bank of Israel
3	Jürgen von Hagen, Mannheim Univ., "Money Growth Targeting." *Discussant*: Stephen Cecchetti, FRB of New York
4	Bennett McCallum and Edward Nelson, Carnegie Mellon Univ., "Nominal Income Targeting in an Open-Economy Optimizing Model." *Discussant*: Glenn Rudebusch, FRB of San Francisco
5	Dale Henderson with Christopher Erceg and Andrew Levin, Federal Reserve Board,"Output-Gap and Price Inflation Volatilities: Reaffirming Tradeoffs in an Optimizing Model." *Discussant*: Stefan Gerlach, Bank for International Settlements
6	Lars Svensson, IIES, "Inflation Targeting as a Monetary Policy Rule." *Discussant*: Alan Blinder, Princeton Univ.
7	Claes Berg, Sveriges Riksbank, and Lars Jonung, Stockholm School of Economics, "Pioneering Price Level Targeting: the Swedish Experience 1931–37." *Discussant*: Mervyn King, Bank of England
8	Panel Discussion: Alan Blinder, Princeton Univ.; Donald Brash, Reserve Bank of New Zealand; Otmar Issing, European Central Bank; Mervyn King, Bank of England; and Guido Tabellini, Bocconi Univ.

roles in the development of Fed models and procedures.[27] The "*Economic Review*" publications of Federal Reserve Banks have become more open to articles of a nearly academic style, which has fostered increased understanding in both directions, and more Federal Reserve Banks have encouraged their research staff members to publish in academic publications. And there are still more channels of communication, not as open or as regularly used during the 1960s and 1970s, that could in principle be listed. Also, Taylor's paper (1993) succeeded brilliantly in interesting central bankers in the consideration of rule-guided policymaking.

The typical method for conducting monetary policy analysis in the NBER and Riksbank conferences can be summarized as follows. An analytical macroeconomic model is developed that includes three major components: (i) a monetary policy rule that specifies quarterly settings for an interest rate instrument, (ii) an IS-type relation or set of relations that specifies how interest rate changes affect aggregate demand and output, and (iii) a price-adjustment equation or set of equations that specifies how inflation behaves in response to output (measured relative to capacity) and expectations regarding the future. Typically, these models feature rational expectations. They may be estimated by various strategies including the estimation procedure termed "calibration" but, whatever the strategy, an attempt is made by the researcher to develop a quantitative model in which parameter values (including disturbance term variances, covariances, and autocovariances) are consistent with actual time series data. Frequently, some effort is taken to make the policy rule operational, i.e., one that is based on a feasible specification of the instrument variable and plausibly available information. Furthermore, in many (but not all) cases the model utilized is obtained by consideration of optimal choices by individual agents in a dynamic and stochastic environment. Then stochastic simulations are conducted using the specified model and alternative policy rules, with summary statistics calculated to represent performance in terms of average values[28] of various macroeconomic measures such as the mean or variability of inflation, the output gap,[29] interest rates, etc.[30] Some models are constructed so that each simulation implies a related utility level for a representative individual agent; in such cases, utility-based performance measures can be calculated.

---

[27] Special mention might be made of former students of John Taylor, including Joseph Gagnon, Andrew Levin, Volker Wieland, and John Williams.

[28] Averaged over numerous replications for each model plus rule specification.

[29] By the output gap I mean the percentage (or fractional) difference between output and its market-clearing, natural-rate, or capacity value. These concepts are subtly but significantly different from model to model.

[30] The King-Wolman and Henderson-Erceg-Levin papers have models compact enough that analytical solutions are used rather than stochastic simulations.

Having outlined the dominant manner in which monetary policy analysis is currently being conducted, our task now is to discuss changes from the research style or styles prevalent in 1971–1973 and then to attempt to attribute these changes to the influence of theory or evidence. Perhaps the most fundamental difference between the procedure outlined above and standard practice as of 1971–1973 is the incorporation of rational expectations. Expectations are important in any dynamic analysis, of course, but if these are rational rather than conforming to some fixed distributed-lag structure then they must be treated quite differently in the study of policy rules, as was emphasized in Lucas's famous critique paper (1976).[31] In particular, the model's equations must not muddle together lagged values from forecasting (expectational) relations and lags in variables due to other causes, such as adjustment costs. This distinction is necessary since, with rational expectations, the coefficients in the forecasting rules will be different with different policy rules—and so cannot be held fixed in comparisons of alternative policy rules. The same is not true of adjustment-cost parameters.

Now clearly the switch from the fixed-lag to the rational expectations hypothesis was the consequence primarily of theoretical, rather than empirical, analysis. At the time it seemed a rather drastic step, but after the fact it has come to be recognized as an entirely natural extension of the usual approach of neoclassical economic analysis to an area of economic activity (expectation formation) that had previously been treated in a non-standard manner. Today, many economists trained after 1980 appear, empirically, to have difficulty in even contemplating any other expectational hypothesis. Also, it should be remembered that Lucas's critique itself was not new, but merely a (brilliantly persuasive) application of Marschak's (1953) fundamental insight that policy analysis requires a structural (as opposed to reduced-form) model.

There have recently been a few attempts to argue that, whatever the theoretical attractions of rational expectations, evidence suggests that the Lucas critique is of little or no consequence empirically. The most extensive and prominent such argument is perhaps that of Hendry and Ericsson (1991) and Ericsson and Irons (1995), who document that an estimated model of money demand shows no symptoms of parameter change (due to coefficient changes in forecasting equations) across periods with different monetary policy rules in effect. A detailed analysis of these studies is beyond the scope of the present paper, but a basic objection to the Hendry-Ericsson-Irons argument can be presented very briefly. It is simply that money demand relations provide an inappropriate laboratory for the study of Lucas-critique effects. The reason is that standard theoretical analysis of money demand behavior, as represented by, e.g., McCallum and Goodfriend (1987), Lucas (1988), Woodford (1995),

[31] The point had been made earlier—e.g., in Lucas (1972a) and Sargent (1971)—but was brought out forcefully and at length in Lucas (1976).

Walsh (1998, Ch. 3), and many others, indicates that forecasting (i.e., expectational) relations are not involved in the optimality conditions (Euler equations) that are typically termed "money demand functions."[32] In other words, these relations are ones that are not predicted to shift with policy changes, under Lucas critique reasoning. Thus a failure to shift with policy changes is irrelevant to the issue. A much better laboratory for consideration of this issue would be Phillips-curve relationships, in which expectational variables are prominent.[33]

A related but somewhat different empirical criticism of rational expectations analysis has recently been put forth by Fuhrer (1997). In an analysis based upon a price adjustment (Phillips curve) relation that is formulated so as to nest expectational (forward looking), inertial (backward looking), and mixed specifications, Fuhrer finds that the expectational terms provide statistically insignificant explanatory power: "I find that expectations of future prices are empirically unimportant in explaining price and inflation behavior" (Fuhrer 1997, 349). This would appear to strike a significant blow to the hypothesis of rational expectations, suggesting that expectations are instead formed as fixed-weight distributed lags of past values. My own response to this argument may not be widely accepted, but it has been held for many years (see McCallum 1980, 718).[34] It is that the incorporation of the rational expectations hypothesis is much more important for policy evaluation than at the estimation stage of the research project. It is fairly plausible that systematic expectational errors can be found in data for past years, distant or recent. But it would be unwise—as mentioned above—to expect any given pattern of expectational errors to prevail in the future, especially if policy is designed to exploit this error pattern. But to conduct policy analysis without assuming rational expectations is to design policy in a manner that attempts to do precisely that, i.e., to exploit a particular pattern. Thus it is desirable to design policy under the assumption of rational expectations even if one has utilized some other expectational hypothesis in estimating the model utilized. Interestingly, Fuhrer himself often uses rational expectations models in his own policy-analysis studies (e.g., Fuhrer 1995).

Another apparent change in monetary policy analysis since 1971–1973 is that such analysis is now typically conducted in terms of a choice among alternative policy rules, as contrasted with the choice of policy actions to

---

[32] More precisely, these conditions that relate real money balances to a transaction quantity variable and an opportunity cost variable are obtained by combining first-order optimality conditions with respect to current consumption and money holdings. There is another construct that could more properly be termed a money demand function, but it would include an infinite sequence of expected future values of all variables taken parametrically by the household, so has a very different specification. On all this, see McCallum and Goodfriend (1987).

[33] Using the Phillips curve as its laboratory, Alogoskoufis and Smith (1991) find dramatic confirmation of Lucas critique effects.

[34] Another response is that model mispecifications are likely to yield results spuriously suggesting the importance of lagged variables.

be taken in a particular episode. But this change is basically a necessary concomitant of the rational expectations assumption and therefore needs no separate discussion.

Rational expectations does not itself imply the absence of a long-run tradeoff between inflation and unemployment. But analysts, like the policymakers mentioned above, moved during the 1970s to near-unanimous acceptance of the Friedman-Phelps-Lucas view that there is no exploitable long-run tradeoff between inflation and output or employment (measured relative to capacity). Undoubtedly, this move was influenced by the same brute experiences as those seen by policymakers, but for analysts there was also some formal econometric work that probably played a role. Thus it was the case that Solow (1969), Tobin (1969), Gordon (1970), and others began quickly to conduct standard tests based on time series regression estimates very promptly after receiving the challenge of Friedman (1966, 1968) and Phelps (1967). These first studies suggested, as veterans of the period will recall, that long-run tradeoffs, did exist—that the long-run Phillips curve was not vertical. But after Sargent (1971) and Lucas (1972b) pointed out the logical flaw that invalidates these studies if expectations are rational, other tests conducted in more appropriate ways by Sargent (1973), McCallum (1976), and Barro (1977) indicated that long-run tradeoffs were not present—a position subscribed to in subsequent studies by Gordon (1975). Thus empirical evidence (of various types) was probably dominant in bringing about a crucial change in analytical views.

The foregoing should not be taken to imply that there are no remaining disagreements concerning long-run relationships between real and monetary variables. Indeed, there are major differences implied by various types of price adjustment models that are currently in use. For example, the type of price adjustment scheme most frequently discussed in practical policymaking circles is that of NAIRU models, where the name is an acronym for non-accelerating-inflation rate of unemployment. In nontechnical publications—even including a symposium in *Economic Perspectives* (Winter 1997)—models of the NAIRU type are often discussed as if they reflected the property known as the natural rate hypothesis (NRH). But the latter, as formulated by Lucas (1972b), asserts that there is no time path of the price level (or the money stock) that would (if maintained) keep output permanently away from its market-clearing natural-rate path. Thus if y_t denotes the log of output and $\bar{y}_t$ is its market-clearing or natural-rate value, the NRH asserts that the unconditional expectation $E(y_t - \bar{y}_t)$ will be unaffected by the selection among monetary policy regimes. Not only will a high inflation rate fail to keep $E(y_t - \bar{y}_t)$ above zero, but so will an increasing (often termed "accelerating") inflation rate or one with an increasing second (or *n*th!) difference in p_t, the log of the price level. By contrast, models of the NAIRU type typically possess the implication that a maintained increase in the inflation rate, such as $\Delta p_t = \Delta p_{t-1} + \delta$ for $\delta > 0$, *will* keep $E(y_t - \bar{y}_t) > 0$. Indeed, the very name NAIRU suggests this

property, for it suggests a stable relationship between the increase in inflation and $y_t - \bar{y}_t$. But that implies that a properly chosen Δp_t pattern can keep $y_t - \bar{y}_t$ above zero permanently, in contradiction to the NRH.[35]

Another prominent class of price adjustment model is the staggered contracts class typified by Calvo (1983), Rotemberg (1982), and Taylor (1980). These also fail to possess the NRH property, but in the opposite direction: they imply that an ever-increasing inflation rate will tend to keep output permanently low! While I personally consider this violation to be a mark against these models, one that suggests the presence of some dynamic misspecification, the implications are not nearly so dangerous from a policy perspective as those of the NAIRU class. One price adjustment model that does satisfy the NRH is the "P-bar model" used by McCallum and Nelson (1998). Its main weakness is that it fails to produce strong positive serial correlation in inflation rates—i.e., sticky inflation—which seems to be a feature of quarterly data in the U.S. and elsewhere. However, the only compact model known to me that does tend to generate inflation persistence is that of Fuhrer and Moore (1995), which fails to satisfy the NRH (although it fails by less than the others mentioned above).

A striking feature of the typical models in the NBER and Riksbank conferences is that they include no money demand equations or sectors. That none is necessary can be understood by reference to the following simple three-equation system.

$$y_t = \alpha_0 + \alpha_1 E_t y_{t+1} + \alpha_2(R_t - E_t \Delta p_{t+1}) + \alpha_3(g_t - E_t g_{t+1}) + v_t \qquad (1)$$

$$\Delta p_t = E_t p_{t+1} + \alpha_4(y_t - \bar{y}_t) + u_t \qquad (2)$$

$$R_t = \mu_0 + \mu_1(\Delta p_t - \Delta p^*) + \mu_2(y_t - \bar{y}_t) + e_t \qquad (3)$$

Here equations (1)–(3) represent an expectational IS equation, a price-adjustment relationship, and a Taylor-style monetary policy rule, respectively. The basic variables are y_t = log of output, p_t = log of price level, and R_t = nominal one-period interest rate, so Δp_t represents inflation, $R_t - E_t \Delta p_{t+1}$ is the real interest rate, and $y_t - \bar{y}_t \equiv \tilde{y}_t$ is the fractional output gap (output relative to its capacity or natural rate value, whose log is $\bar{y}_t$). Also, g_t represents the log of government purchases, which for present purposes we take to be exogenous. In this system, E_t denotes the expectations operator conditional on information available at time t, so $E_t \Delta p_{t+1}$ is the rational expectation formed at t of Δp_{t+1}, the inflation rate one period in the future.

[35] It should be mentioned, perhaps, that a failure to satisfy the NRH is not the same thing as the absence of monetary superneutrality.

The basic point at hand is that with g_t and $\bar{y}_t$ exogenous, and expectations formed rationally, the three equations (1)–(3) are sufficient in number to fully determine the time paths of the model's three endogenous variables, namely, y_t (or $\tilde{y}_t$), p_t and R_t. Thus there is no need for a money-demand equation.[36] If nevertheless one such as

$$m_t - p_t = \gamma_0 + \gamma_1 \gamma_t + \gamma_2 R_t + \varepsilon_t$$

were appended to the system, it would not be inconsistent with (1)–(3) but would be irrelevant in the sense that it would play no role in determining the behavior of y_t, Δp_t or R_t. Its role would be merely to determine the amount of money (m_t, in log terms) that would be demanded and which would therefore necessarily be supplied by the central bank in the process of setting interest rates in conformity with the policy rule (3).

It can be seen that the absence of any money demand function—or any money stock variable!—in the prototype system (1)–(3) reflects two properties of the latter. These are that no "real balance" term $m_t - p_t$ appears in the IS relation (1) and that the interest rate R_t is used as the policy instrument. So we ask, what are the methodological precepts that lead to those two aspects of (1)–(3)?

In the case of the second aspect, that R_t is specified as the instrument variable, the rationale is almost entirely empirical. The fact is that actual central banks in industrial countries conduct monetary policy in a manner that is much more accurately depicted by writing R_t rather than m_t (even if interpreted as the monetary base) as the instrument or operating-target variable. Thus, such policy rules are studied even by economists who might be regarded as possessing "monetarist" tendencies and possibly even believing that policy might be improved if central banks used m_t as their instrument (e.g., McCallum and Nelson, 1998).

The first aspect of the system (1)–(3), that no "money" term appears in the IS function (1), is by contrast of an *a priori* origin. Traditionally, of course, it has been usually presumed in analysis of an IS-LM style that IS functions do not include real balance terms.[37] But in recent work, the IS relationship has often been of the expectational variety that includes $E_t y_{t+1}$ as in Kerr and King (1996), McCallum and Nelson (1997), and Rotemberg and Woodford (1997). Such relations are obtained from explicit optimization analysis of the dynamic choice problems faced by individual agents, so they have arguably greater claim to theoretical validity than traditional specifications. But in such analyses, the absence of any monetary real-balance variable depends upon the common assumption that separability obtains in the indirect utility function

[36] Indeed, the Fed's major new quarterly econometric model was constructed without any money demand function or any reference to any monetary aggregate—see Brayton et al. (1997).

[37] More accurately, it has been assumed that real balance terms should be included in principle but are of negligible importance practically; the classic reference on this is Patinkin (1965).

that reflects the transaction-facilitating properties of the medium of exchange (i.e., money). There is, however, no compelling theoretical basis for that assumption, which is presumably made for analytical convenience. Indeed, it could be argued that separability is not very plausible. Accordingly, the absence of any real balance term in (1), and the omission of monetary variables from model (1)–(3), hinges on the presumption that nonseparabilities of the relevant type are quantitatively unimportant—i.e., that the marginal utility of consumption is (for a given rate of consumption) virtually independent of the level of real money balances. The justification for that presumption has not been explicitly discussed in the studies cited.

It has been mentioned that the NBER and Riksbank conferences featured considerable agreement among participants concerning research strategy. Furthermore, Taylor (1998, 4–5) argues that there exists a fair amount of substantive agreement; specifically that model simulations at the NBER conference "show that simple policy rules work remarkably well in a variety of situations; they seem to be surprisingly good approximations to fully optimal policies" and "simple policy rules are more robust than complex rules across a variety of models." Also, "introducing information lags as long as a quarter does not affect the performance of the policy rules by very much." In addition, Taylor mentions other issues about which these studies do not reflect agreement, including: the value of interest-rate "smoothing" terms in the rules; whether responses should be geared to expected feature values rather than currently-observed values; and about measurement of potential or natural-rate output.

One area of disagreement among researchers concerns the distinction between "optimal control" and "robustness" approaches for the design of monetary policy rules. For some analysts, the task of policy rule design is to develop an appropriate macroeconomic model of the economy and then conduct an optimal control exercise to determine what the best policy rule would be for the economy in question. Neither step is trivial, but both represent rather straightforward scientific problems. There would remain the task of convincing actual policy makers to implement this rule, of course, but that is a matter of persuasion rather than scientific investigation. Notable examples of this approach are Feldstein and Stock (1994) and Svensson (1997). To other economists, by contrast, a crucial feature of the policymaking process is the lack of professional agreement concerning the appropriate specification of a model suitable for monetary policy issues. Various members of this group would emphasize different portions of a macro model,[38] but to all in

[38] My own candidate for the weakest component in a macroeconomic model is the price-adjustment (Phillips curve) sector. In McCallum (1999, fn. 14) the argument is stated as follows. "It is not just that the economics profession does not have a well-tested quantitative model of the quarter-to-quarter dynamics, the situation is much worse than that: we do not even have any basic agreement about the qualitative nature of the mechanism. This point can be made by mentioning some of the leading theoretical categories, which include: real business cycle models; monetary

this group it seems hard to avoid the conclusion that agreement upon model specification is predominately absent—and that different models give rise to different alleged implications for policy. Thus these latter economists believe that in practice the optimal control strategy collapses in response to the question, "What is the appropriate model specification?" As a consequence, the approach favored by these analysts is to search for a policy rule that possesses robustness in the sense of yielding reasonably desirable outcomes in policy simulation experiments conducted with a wide variety of models.[39] It is not necessary that the collection of models all be designed and simulated by a single researcher or research team; the work of Bryant et al. (1993) and Taylor (1998) represent studies by over a half-dozen research teams each.

In evaluating candidate policy rules, it would clearly be desirable to have at hand an established specification of the appropriate ultimate *goals* of monetary policy. In that regard there are several important issues, including whether a CB should keep actual inflation or expected inflation close to some normative value; what that normative value should be (or should it change over time?); and how heavily the variability of output—or is it output relative to capacity (measured how?) or consumption?—should be weighted in relation to the inflation criterion. Of course in optimizing models that are specified at the level of individuals' utility and production functions, the answers to such questions are implicit to the solution to the optimal control problem. But again the fundamental difficulty mentioned above intrudes in a crucial manner, for these answers must depend significantly upon the model's specification. Thus the absence of agreement regarding model specification implies that there can be at present no consensus as to the precise goals that are appropriate. In practice, nevertheless, there seems currently to be a substantial amount of agreement about *actual* CB objectives; namely that most CBs desire to keep realized inflation close to zero (allowing for measurement error) and to keep output (or employment) close to a capacity or natural-rate value that grows with the capital stock, the labor force, and technical progress. As a matter of logic it cannot be rigorously established that these objectives are optimal (from the perspective of individuals' preferences), but it seems a reasonable judgment

misperception models; semi-classical price adjustment models; models with overlapping nominal contracts of the Taylor variety or the Fischer variety or the Calvo-Rotemberg type; models with nominal contracts set as in the recent work of Fuhrer and Moore; NAIRU models; Lucas supply function models; MPS-style markup pricing models; and so on. Not only do we have all of these basic modeling approaches, but to be made operational each of them has to be combined with some measure of capacity output—a step that itself involves competing approaches—and with several critical assumptions regarding the nature of different types of unobservable shocks and the time series processes generating them. Thus there are dozens or perhaps hundreds of competing specifications regarding the precise nature of the connection between monetary policy actions and their real short-term consequences. And there is little empirical basis for much narrowing of the range of contenders."

[39] Some representatives are Bryant et al. (1993), McCallum (1999), and Taylor (1998).

that they probably provide an appropriate specification of CB macroeconomic goals.

A related research-design issue that has attracted some attention involves the distinction between "instrument rules" and "target rules," in the language of Svensson (1997). Rules of the former type, exemplified by Taylor (1993) and presumed in my foregoing comments, specify period-by-period settings of a controllable instrument variable as in equation (3) above. The second type of rules, by contrast, specify target values for some variable or combination of variables that the CB can influence but not directly control, with these values obtained (at least in the work of Svensson) by optimal control exercises on the basis of a designated model and objective function. Thus the specification of a target rule amounts logically to the selection by the analyst of a model and an objective function, whereas an instrument rule reflects the analyst's hypothesis that the CB would (whatever its model and objectives) achieve satisfactory results if it were to implement the rule.[40]

It has been emphasized that models represented by the system (1)–(3) specify slowly adjusting price levels. Thus central-bank policy actions, represented by changes in R_t, typically have effects on real output, y_t. It is of some importance to ask, then, what is the scientific justification for models of this type in preference to ones (including RBC models) in which monetary policy actions have no systematic effects on real variables. Is it theory or empirical evidence that indicates that prices are sticky and monetary policy able to influence output?[41]

Here the argument is much the same as that of Section 2 above—the answer must be empirical evidence since neoclassical theory certainly does not entail price stickiness. But then the question becomes, what *type* of evidence has indicated strongly and clearly that sticky prices and monetary effects on real variables are a feature of actual economies? For it is unclear how to look for the former, empirically, and there is no shortage of empirical studies that fail to find major effects of the latter type (e.g., Eichenbaum and Singleton 1986). Regarding sticky prices *per se* there is some survey evidence provided by Blinder (1994) and studies of particular commodities by others but it is my impression that most analysts have judged these to be non-compelling. More influential, I believe, has been the perception that sharp major changes in monetary policy conditions (e.g., in the United States during 1981) have in fact had major real effects in the same direction, together with the belief that price stickiness provides the most satisfactory means of rationalizing that fact.

[40] There is also a major controversy as to whether a CB can implement a rule of the "committed" or "non-discertionary" type. Since my own affirmative position on this issue (e.g., McCallum 1999) is somewhat unorthodox, I propose not to discuss it in the present paper.

[41] There is small but significant school of thought that attributes real effects of policy to financing constraints in flexible-price models, e.g., Christiano and Eichenbaum (1995). Discussion of this position is beyond the scope of the present paper.

Ironically, the empirical study that has probably attracted the most support for this viewpoint is Romer and Romer (1989), a study that, like Hoover and Perez (1994), I find rather unsatisfactory.[42]

One source of difficulty in formal empirical studies of monetary policy effects on real variables has been the common practice of focusing attention on real responses to policy innovations—i.e., unexpected components —in vector autoregression (VAR) studies. Although Christiano, Eichenbaum, and Evans (1998) have reported effects of this type in an extensive study, they are not quantitatively large in many VAR studies. But the VAR approach seems inherently to miss the major effects, because the measured innovation component of policy-variable fluctuations is extremely small relative to the systematic component in terms of variability. Thus, for example, the systematic component's variability is about 16 times as large as for the innovation component in an estimated interest rate policy rule for the U.S., 1955–1996 (quarterly data). One way of making the relevant point is by consideration of an extreme case. Suppose that a central bank's policy rule is activist but entirely systematic, i.e., is devoid of random components. Then a well-designed VAR study would attribute no importance to monetary policy in affecting output—or inflation!—although it could be that the systematic component of policy was in fact very important.

A related point concerns the way in which empirical evidence works in persuading specialists in monetary economics (and, probably, other areas). In that regard it is almost never the case that an analyst's view on some important hypothesis is crucially dependent upon the results of a formal econometric study—no single study is decisive. This is in part because conclusions about crucial properties of macroeconomic systems almost always require identification of structural parameters in a well-specified model, yet both identification and "correct specification" are exceedingly difficult to achieve in the macroeconomic context, for there is usually some highly dubious feature of any manageable model. Instead, it is the cumulative effect of several econometric studies and/or various bits of evidence obtained in more informal ways, all taken together, that is usually persuasive to an analyst.

This section has argued that monetary policy analysis, like macroeconomics more generally and the practice of monetary policy, has been significantly influenced over the last 25 years by both theoretical and empirical developments. That such is the case is certainly desirable, I would think.

---

[42] It is obvious that the Romer and Romer (R&R) dummy variable is not exogenous, for it reflects actions taken in response to recently-prevailing macroeconomic conditions. Thus my own summary statement (McCallum 1994, 334) is that their study differs from previous attempts to measure monetary policy effects primarily as follows: "the R&R dummy reflects changes in only one direction, does not reflect the intensity of policy actions, and is based on statements rather than actions. Thus one is led to wonder how use of this dummy, instead of a traditional measure, constitutes an improvement over prior practice."

Indeed, there is an important sense in which significant scientific progress inevitably requires both theoretical and empirical inputs. Evidence is necessary, obviously, because theories with content may by construction be untrue, i.e., grossly inconsistent with relevant facts. But theory too is necessary, for one can only make sense of facts and measurements within the disciplining context of a theoretical structure that gives these facts some coherence and delineates what is and is not relevant, etc.

5. COINTEGRATION AND MONETARY ANALYSIS

In this section, I would like to take up an issue that is highly pertinent to this paper's theme although it failed to find a place in the above evolutionary discussion. This issue concerns a claim, which appears occasionally in the literature, to the effect that a failure of real money balances, real income, and nominal interest rates to be *cointegrated* implies the absence of any long-run relationship of the type that is necessary for the validity of traditional monetary economics. Cuthbertson and Taylor (1990, 295), for example, have expressed the claim as follows: "If the concept of a stable, long-run money demand function is to have any empirical content whatsoever, then m_t [i.e., log money]. . . must be cointegrated with log prices, log income, and interest rates." Engle and Granger (1987) presented evidence contrary to the cointegration hypothesis; several other researchers have reached the opposite conclusion but only after accepting the presumption that cointegration is necessary for standard monetary analysis.

My objective here is to argue that this presumption is basically mistaken. Of course there is a technical sense in which it is correct: if $m_t - p_t$, y_t and R_t are all integrated (difference-stationary) of order one[43] but not cointegrated, then the disturbance entering any linear relation among them must by definition be nonstationary, so $m_t - p_t$ and any linear combination of y_t and R_t can drift apart as time passes. But it is highly misleading to conclude that in any practical sense a long-run relationship is therefore nonexistent. The following argument is entirely interpretive; it includes no suggestion of *technical* error in the literature discussed. But that does not diminish its importance.

To develop the argument, let us consider again the example of a traditional money demand function of the form (4). Suppose that $m_t - p_t$, y_t and R_t are all I(1) variables and that each has been processed by the removal of a deterministic trend. Then the cointegration status of relationship (4) depends

[43] A time-series variable is integrated of order one, written I(1), if it must be differenced once to obtain a variable that is covariance stationary. This will be the case for an ARMA variable if its autoregressive parameter has a unit root.

upon the properties of the disturbance term ε_t: if its process is of the difference-stationary type that includes a unit AR root, then the variables in (4) will not be cointegrated.

But the traditional view of money demand theory, as represented by the studies cited above, provides no reason for believing that ε_t would instead be trend stationary (i.e., would possess no AR unit root component). Indeed, it would seem almost to suggest the opposite—for the theoretical rationale for (4) is built upon the transaction-facilitating function of money, but the technology for effecting transactions is constantly evolving. And since technical progress cannot be directly measured by available variables, the effects of technical change (not captured by a deterministic trend) show up in the disturbance term, ε_t. But the nature of technological progress is such that changes (shocks) are typically not reversed. Thus one would expect *a priori* there to be an important permanent component to the ε_t process, making it one of the integrated type—and thereby making $m_t - p_t$ not cointegrated with y_t and R_t.

In such a case, however, the "long-run" messages of traditional monetary analysis could easily continue to apply. Provided that the magnitude of the variance of the innovation in ε_t is not large in relation to potential magnitudes of Δm_t values, it will still be true that inflation rates will be principally determined, over long periods of time, by money growth rates. And even without that proviso, long-run monetary neutrality may still prevail, superneutrality may be approximately but not precisely valid, etc. That the disturbance ε_t is of the difference-stationary class is simply not a source of embarrassment or special concern for supporters of the traditional theory of money demand, some of whom have estimated money demand relations like (4) after *assuming* that ε_t is a pure random walk![44]

6. CONCLUSION

The picture painted in the preceding discussion is one that attributes major changes in the analysis and practice of monetary policy over the years 1973–1998 to a combination of theoretical and empirical influences. This is not a very dramatic conclusion; indeed, one might say that it is almost empty. It would be possible to add a bit of debatable content, by asserting that the mixture of influences has been reasonably appropriate—about the right amount of theory and empirics—but I would not feel comfortable in doing so. Partly that is because I would have preferred that models with complete price flexibility

[44] More generally, I would *a priori* expect cointegration among basic variables—ones that enter utility or production functions, not their differences—to be quite rare. for behavioral relations typically include disturbance terms that represent unobservable (and thus omitted) variables, which include shocks to preferences and technology. But these shocks would seem likely to include significant random-walk components, as argued above. Thus disturbance terms in behavioral relations should, according to this argument, typically possess unit-root components.

had not been quite so dominant during the years (say) 1982–1992, although the surge of work with sticky price models in recent years may have largely made up for their previous neglect. But an equally important reason stems from the question: what type of evidence could be presented in support of a "reasonably appropriate" contention? Unfortunately, I know of no satisfactory way of making such a determination, especially since most influential studies involve a blend of theory and evidence. For example, the rather abstract theoretical analysis of Lucas (1972a) was not actually devoid of empirical content in the sense that its theorizing was specifically designed to rationalize a set of broad facts that were (and are) of genuine, fundamental importance. Consequently, I am left with the rather limp conclusion with which this paragraph began.

In conclusion, then, it may be appropriate to add the opinion that the current state of monetary economics is not as highly unsatisfactory as has been claimed over the past decade or so by various commentators at conferences and seminars. The type of claim that I have in mind does not often make it into print, so I cannot provide citations, but I am confident that many readers can supply examples from their own experiences. In any event, the state of monetary economics seems to me to be about as healthy as that of economic analysis in general. The contrary opinion is rather widely held for three reasons, I would suggest, none of which is sound. First, much of the negative opinion has been put forth by economists who are themselves proponents of an entirely unsatisfactory theory of money demand, one involving overlapping generations models in which the asset termed "money" plays no role as a (transaction-facilitating) medium of exchange. Since this role provides the defining characteristic of money, as distinct from other assets, it is not surprising that proponents of such a theory would find it unsatisfactory. Second, rather inconsequential differences among proponents of the transaction-facilitating approach to money demand theory—e.g., cash-in-advance, money-in-utility function, shopping time, or transaction-cost-in-budget constraint models—have tended to obscure the fundamental similarity of principles and implications among the variants of this approach. Third, the profession's poor level of understanding of the precise nature of the dynamic connection between monetary and real variables—i.e., of price adjustment relations—has tended to reflect discredit upon monetary economics, although this relation belongs to the realm of macroeconomics more broadly. In that regard, moreover, it is an illusion to believe that macroeconomics is itself in poor condition in relation to microeconomics, an illusion generated by the fact that applied macro features a much more ambitious agenda than applied micro—the understanding of quarter-to-quarter dynamic movements in variables rather than just steady-state values. Correcting for that difference, the extent of disagreement seems about the same in the two sub-disciplines.

REFERENCES

Alogoskoufis, G. L., and Smith, R. (1991) "The Phillips curve, the persistence of inflation, and the Lucas critique," *American Economic Review* 81: 1254–75.

Altug, S. (1989) "Time-to-build and aggregate fluctuations," *International Economic Review* 30: 889–920.

Axilrod, S. H. (1983) "A comment on Brunner and Meltzer," *Carnegie-Rochester Conference Series* 18: 105–12.

Barro, R. J. (1977) "Unanticipated money growth and unemployment in the United States," *American Economic Review* 67: 101–15.

Barro, R. J., and Gordon, D. B. (1983) "A positive theory of monetary policy in a natural rate model," *Journal of Political Economy* 91: 589–610.

Basu, S. (1996) "Procyclical productivity: increasing returns or cyclical utilization," *Quarterly Journal of Economics* 111: 719–51.

Bernanke, B. S., and Mishkin, F. S. (1997) "Inflation targeting: A new framework for monetary policy?" *Journal of Economic Perspectives* 11: 97–116.

Blinder, A. S. (1994) "On sticky prices: academic theories meet the real world," *Monetary Policy,* ed. N.G. Mankiw. Chicago: University of Chicago Press.

Brayton, F., Levin, A., Tryon, R., and Williams, J. C. (1997) "The evolution of macro models at the Federal Reserve Board," *Carnegie-Rochester Conference Series on Public Policy* 47: 43–81.

Broaddus, A., and Goodfriend, M. (1984) "Base drift and the longer-run growth of M1: Experience from a decade of monetary targeting," Federal Reserve Bank of Richmond *Economic Review* 70 (Nov./Dec.): 3–14.

Brunner, K., and Meltzer, A. H. (1983) "Strategies and tactics for monetary control," *Carnegie-Rochester Conference Series* 18: 59–104.

Bryant, R. C., Hooper, P., and Mann, C. L. (1993) *Evaluating Policy Regimes: New Research in Empirical Macroeconomics.* Washington: Brookings Institution.

Burns, A. (1979) "The real issues of inflation and unemployment," *Federal Reserve Readings on Inflation.* Federal Reserve Bank of New York.

Calvo, G. A. (1983) "Staggered prices in a utility prices in a utility maximizing framework," *Journal of Monetary Economics* 12: 383–98.

Christiano, L. J., and Eichenbaum, M. (1995) "Liquidity effects, monetary policy, and the business cycle," *Journal of Money, Credit, and Banking* 27: 1113–36.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1998) "Monetary policy shocks: What have we learned and to what end?" NBER Working paper 6400. Forthcoming in *Handbook of Macroeconomics,* ed. J. B. Taylor and M. Woodford. Amsterdam: North Holland.

Cogley, T., and Nason, J. M. (1995) "Output dynamics in real-business cycle models," *American Economic Review* 85: 492–511.

Cuthbertson, K., and Taylor, M. P. (1990) "Money demand, expectations, and the forward-looking model," *Journal of Policy Modeling* 12: 289–315.

Eichenbaum, M., and Singleton, K. J. (1986) "Do equilibrium real business cycle theories explain postwar U.S. business cycles?" *NBER Macroeconomic Annual 1986.* Cambridge, MA: MIT Press.

Engle, R. F., and Granger, C. W. J. (1987) "Co-integration and error correction: Representation, estimation, and testing," *Econometrica* 55: 251–76.

Ericsson, N., and Irons, J. (1995) "The Lucas critique in practice: Theory without measurement," in *Macroeconometrics: Developments, Tensions, and Prospects,* ed. by K. D. Hoover. Boston: Kluwer Academic Press, 263–312.

Evans, C. L. (1992) "Productivity shocks and real business cycles," *Journal of Monetary Economics* 29: 191–208.

Feldstein, M., and Stock, J. H. (1994) "The use of a monetary aggregate to target nominal GDP," in *Monetary Policy,* ed. N. G. Mankiw. Chicago: University of Chicago Press.

Friedman, M. (1966) "Comments," in *Guidelines, Informal Controls, and the Marketplace,* ed. G. P. Schultz and R. Z. Aliber. Chicago: University of Chicago Press.

_____. (1968) "The role of monetary policy," *American Economic Review* 58: 1–17.

Fuhrer, J. C. (1997) "The (un)importance of forward-looking behavior in price specifications," *Journal of Money, Credit, and Banking* 29: 338–50.

_____. (1995) "The persistence of inflation and the cost of disinflation," *New England Economic Review* Jan/Feb: 3–16.

_____, and Moore, G. R. (1995) "Inflation persistence," *Quarterly Journal of Economics* 109: 127–59.

Gali, J. (1997) "Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations?" NBER Working Paper 5721.

Goodfriend, M. (1997) "Monetary policy comes of age: A 20th century odyssey," Federal Reserve Bank of Richmond *Economic Quarterly* 83 (Winter): 1–22.

Goodfriend, M., and King, R. E. (1997) "The new neoclassical synthesis and the role of monetary policy," *NBER Macroeconomics Annual 1997*: 231–82.

Gordon, R. J. (1970) "The recent acceleration of inflation and its lessons for the future," *Brookings Papers on Economic Activity,* No. 1: 8–41.

———. (1975) "The impact of aggregate demand on prices," *Brookings Papers on Economic Activity,* No. 3: 613–62.

Haldane, A. G., ed. (1995) *Targeting Inflation.* Bank of England.

Hall, R. E. (1978) "Stochastic implications of the life-cycle-permanent income hypothesis: theory and evidence," *Journal of Political Economy* 86: 971–87.

Hansen, L. P., and Singleton, K. J. (1982) "Generalized instrumental variables estimation of nonlinear rational expectations models," *Econometrica* 50: 1269–86.

Hendry, D. F. and Ericsson, N. R. (1991) "Modeling the demand for narrow money in the United Kingdom and the United States," *European Economic Review* 35: 833–86.

Hoover, K. D. (1995) "Facts and artifacts: calibration and the empirical assessment of real business cycle models," *Oxford Economic Papers* 47: 24–44.

Hoover, K. D., and Perez, S. J. (1994) "Money may matter, but how could you know?" *Journal of Monetary Economics* 34: 89–100.

Kerr, W. and King, R. G. (1996) "Limits on interest rate rules in the IS-LM model," Federal Reserve Bank of Richmond *Economic Quarterly* 82 (Spring): 47–75.

King, R. E. (1990) "Money and business cycles," Working Paper. Forthcoming in *Journal of Monetary Economics.*

King, R. E., and Watson, M. W. (1994) "The post-war U.S. Phillips curve: a revisionist econometric history," *Carnegie-Rochester Conference Series* 41: 157–219.

Kydland, F. E., and Prescott, E. C. (1980) "A competitive theory of fluctuations and the feasibility of stabilization policy," in *Rational*

Expectations and Economic Policy, ed. S. Fischer. Chicago: University of Chicago Press.

_____. (1982) "Time to build and aggregate fluctuations," *Econometrica* 50: 1345–70.

_____. "The workweek of capital and its cyclical implications," Working paper.

Leiderman, L., and Svensson, L. E. O., eds. (1995) *Inflation Targets.* London: Centre for Economic Policy Research.

Litterman, R. B., and Weiss, L. (1989) "Money, real interest rates, and output: a reinterpretation of postwar U.S. data," *Econometrica* 53: 129–56.

Lucas, R. E., Jr. (1972a) "Expectations and the neutrality of money," *Journal of Economic Theory* 4: 103–24.

_____. (1972b) "Econometric testing of the natural-rate hypothesis," in *The Econometrics of Price Determination,* ed. O. Eckstein. Washington, D.C.: Board of Governors of the Federal Reserve System.

_____. (1973) "Some international evidence on output-inflation tradeoffs," *American Economic Review* 63: 326–34.

_____. (1975) "An equilibrium model of the business cycle," *Journal of Political Economy* 83: 1113–44.

_____. (1976) "Econometric policy evaluation: a critique," *Carnegie-Rochester Conference Series* 1: 19–46.

_____. (1988) "Money demand in the United States: a quantitative review," *Carnegie-Rochester Conference Series* 29: 137–67.

Mankiw, N. G., Rotemberg, J. J., and Summers, L. (1985) "Intertemporal substitution in macroeconomics," *Quarterly Journal of Economics* 100: 225–51.

Marschak, J. (1953) "Economic measurements for policy and prediction," in *Studies in Econometric Method,* ed. W. C. Hood and T. J. Koopmans. New York: John Wiley.

McCallum, B. T. (1976) "Rational expectations and the natural rate hypothesis: some consistent estimates," *Econometrica* 44: 43–52.

_____. (1980) "Rational expectations and macroeconomic stabilization policy," *Journal of Money, Credit, and Banking* 12: 716–46.

_____. (1986) "On 'real' and 'sticky-price' theories of the business cycle," *Journal of Money, Credit, and Banking* 18: 397–414.

———. (1989) "Real business cycle models," in *Modern Business Cycle Theory,* ed. R. J. Barro. Cambridge, MA: Harvard University Press.

———. (1994) "Comment on 'Federal Reserve Policy: Cause and Effect' by M. D. Shapiro," in *Monetary Policy,* ed. N. G. Mankiw. Chicago: University of Chicago Press for NBER.

———. (1999) "Issues in the design of monetary policy rules," in *Handbook of Macroeconomics,* ed. J. B. Taylor and M. Woodford. Amsterdam: North-Holland Pub. Co.

McCallum, B. T., and Goodfriend, M. S. (1987) "Demand for money: theoretical studies," in *The New Palgrave,* ed. J. Eatwell, M. Milgate, and P. Newman. New York: Stockton Press.

McCallum, B. T., and Nelson, E. (1997) "An optimizing IS-LM specification for monetary policy and business cycle analysis," NBER Working paper 5875. Forthcoming in *Journal of Money, Credit, and Banking 31*.

———. (1998) "Nominal income targeting in an open-economy maximizing model," NBER Working paper 6675.

Nelson, C. R., and Plosser, C. I. (1982) "Trends and random walks in macroeconomic time series," *Journal of Monetary Economics* 10: 139–62.

Nelson, E. (1998) "Sluggish inflation and optimizing models of the business cycle," *Journal of Monetary Economics* 42: 303–22.

Patinkin, D. (1965) *Money, Interest, and Prices,* second ed. New York: Harper & Row.

Phelps, E. S. (1967) "Phillips curves, expectations of inflation, and optimal unemployment overtime," *Economica* 34: 254–81.

Prescott, E. S. (1986) "Theory ahead of business cycle measurement," *Carnegie-Rochester Conference Series* 25: 11–44.

Romer, C. D., and Romer, D. (1989) "Does monetary policy matter? A new test in the spirit of Friedman and Schwartz," *NBER Macroeconomics Annual 1989*: 121–70.

Rotemberg, J. J. (1982) "Monopolistic price adjustment and aggregate output," *Review of Economic Studies* 44: 517–31.

Rotemberg, J. J., and Woodford, M. (1997) "An optimization based econometric framework for the evaluation of monetary policy," *NBER Macroeconomics Annual 1997*: 297–345.

Sargent, T. J. (1971) "A note on the accelerationist controversy," *Journal of Money, Credit, and Banking* 3: 50–60.

———. (1973) "Rational expectations, the real rate of interest, and the natural rate of unemployment," *Brookings Papers on Economic Activity,* No. 2: 429–72.

Sargent, T. J., and Wallace, N. (1975) "'Rational' expectations, the optimal monetary instrument, and the optimal money supply role," *Journal of Political Economy* 83: 241–54.

Sims, C. A. (1980) "A comparison of interwar and postwar business cycles: monetarism reconsidered," *American Economic Review Papers and Proceedings* 70: 250–57.

Solow, R. M. (1969) *Price Expectations and the Behavior of the Price Level*. Manchester: Manchester University Press.

Svensson, L. E. O. (1997) "Inflation forecast targeting: implementing and monitoring inflation targets," *European Economic Review* 41: 1111–46.

Taylor, J. B. (1980) "Aggregate dynamics and staggered contracts," *Journal of Political Economy* 88: 1–23.

———. (1989) "Monetary policy and the stability of macroeconomic relationships," *Journal of Applied Econometrics* 4: S161–78.

———. (1993) "Discretion versus policy rules in practice," *Carnegie-Rochester Conference Series* 39, 195–214.

———. (1996) "How should monetary policy respond to shocks while maintaining long-run price stability—conceptual issues," in *Achieving Price Stability,* Federal Reserve Bank of Kansas City.

———. (1998) "The robustness and efficiency of monetary policy rules as guidelines for interest rate setting by the European Central Bank," Working paper.

Tobin, J. (1969) "Discussion," *Inflation: Its Causes, Consequences, and Control,* ed. S. W. Rousseas. New York: New York University.

Walsh, C. E. (1998) *Monetary Theory and Policy.* Cambridge, MA: MIT Press.

Walters, A. A. (1971) "Consistent expectations, distributed lags, and the quantity theory," *Economic Journal* 81: 273–81.

Watson, M. W. (1993) "Measures of fit for calibrated models," *Journal of Political Economy* 101: 1011–41.

Woodford, M. (1995) "Price-level determinacy without control of a monetary aggregate," *Carnegie-Rochester Conference Series* 43: 1–46.