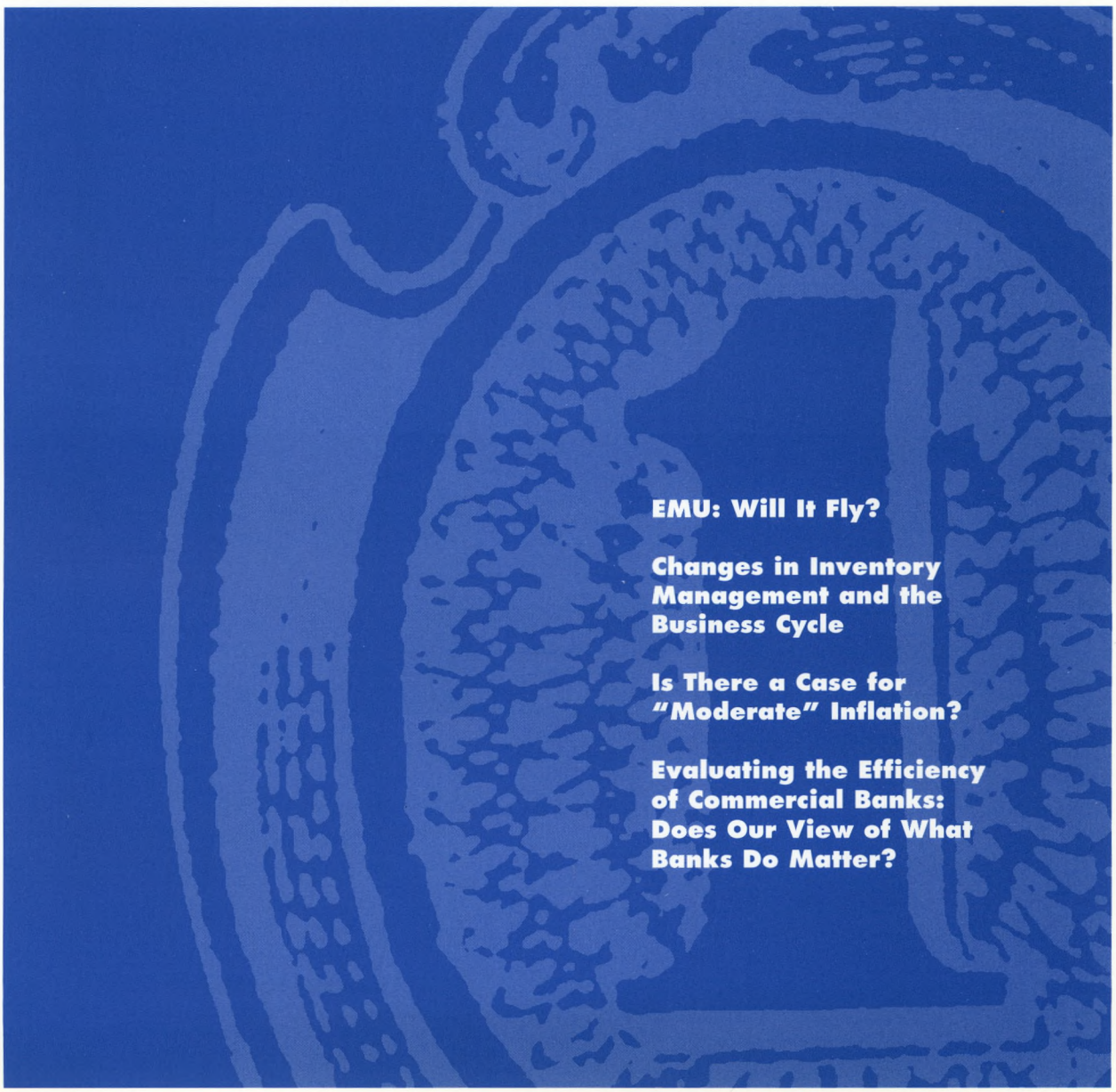
Federal Reserve Bank of St. Louis

REVIEW

JULY/AUGUST 1995 • VOLUME 77, NUMBER 4



EMU: Will It Fly?

Changes in Inventory Management and the Business Cycle

Is There a Case for "Moderate" Inflation?

Evaluating the Efficiency of Commercial Banks: Does Our View of What Banks Do Matter?

REVIEW

Review is published six times per year by the Research Department of the Federal Reserve Bank of St. Louis. Single-copy subscriptions are available free of charge. Send requests to: Federal Reserve Bank of St. Louis, Public Affairs Department, P.O. Box 442, St. Louis, Missouri 63166-0442.

REVIEW
JULY/AUGUST 1995

CONTENTS

Volume 77, Number 4

3 **EMU: Will It Fly?**
Patricia S. Pollard

Will a monetary union be established in Europe before the end of this century? Prior to the establishment of a monetary union within the European Union, countries are expected to meet designated criteria regarding their interest rates, inflation rates, government finances and exchange rates. These criteria were designed to ensure a commitment to price stability and economic convergence among potential entrants. Patricia S. Pollard examines the progress that has been made in fulfilling these criteria and the countries' prospects for meeting the entry criteria. Furthermore, she examines the options available to the European Union if there aren't enough countries to fulfill the criteria before the target date for monetary union.

17 **Changes in Inventory Management and the Business Cycle**
Donald S. Allen

The change in business inventory has always played a major role during downturns in the business cycle. Innovations in inventory management, such as "just-in-time" methods and computerized stock management using bar code scanning, have become more popular over the last decade. Donald S. Allen looks at whether these changes are expected to have a major impact on the amplitude or duration of the business cycle.

27 **Is There a Case for "Moderate" Inflation?**
Alvin L. Marty and Daniel L. Thornton

Two objections to making price stability the primary objective of monetary policy are: (1) Moderate inflation is good for the economy and (2) It is less costly to live with moderate inflation than to eliminate it. Reviewing the first objection, Alvin L. Marty and Daniel L. Thornton consider four arguments, namely, that moderate inflation enhances economic stability, increases output per person, increases the efficiency of inter-industry wage adjustments, and enhances the efficacy of countercyclical monetary policy. Considering the second point, they argue that concern for transitional unemployment is a frail foundation for a policy of moderate inflation.

39 **Evaluating the Efficiency of Commercial Banks: Does Our View of What Banks Do Matter?**
David C. Wheelock and Paul W. Wilson

An inefficient business wastes resources, either by producing less than the feasible level of output from a given amount of input or by using excessive input to produce a given amount of output. Researchers often find that banks are quite inefficient, but don't agree on how best to measure that inefficiency, or even how to measure bank production. David C. Wheelock and Paul W. Wilson show that the average estimated inefficiency, and even the ranking of banks by their inefficiency, is sensitive to whether bank loans and deposits are measured in dollar amounts, or the number of loans and accounts. Their research indicates that in the absence of a standard view of how to measure bank production, the extent to which banks are inefficient will remain an open question.

Patricia S. Pollard is an economist at the Federal Reserve Bank of St. Louis. Jerram C. Betts provided research assistance.

EMU: Will It Fly?

Patricia S. Pollard

In December 1991, the leaders of the member states of the European Union met in Maastricht, the Netherlands, to conclude the negotiations on a *Treaty on European Union*. The Maastricht Treaty, as it is commonly known, encompasses a wide range of issues, from foreign affairs and security policy to citizenship, health and tourism. Primarily, however, the Maastricht Treaty is known for formalizing the intentions of the member states of the European Union to create an economic and monetary union (EMU) by the end of this century. The main features of EMU are the creation of a single monetary policymaking body and a single currency for the European Union.

While EMU seemed certain in December 1991, within a year the outlook had turned much bleaker. In a referendum in June 1992, Danish voters rejected the treaty. This was followed by a series of exchange rate crises affecting the European Union in 1992 and 1993. Despite these setbacks, the Maastricht Treaty was ultimately approved by all member states (a second referendum passed in Denmark in 1993) and the treaty entered into force on November 1, 1993. In accordance with the treaty, the European Union is laying the groundwork for monetary union: creating the institutions and studying the technical details necessary to meld as many as 14 independent monetary policymaking bodies into one cohesive system.[1] Furthermore, to make themselves eligible for entry into EMU, countries are undertaking policies aimed at achieving economic convergence across the European Union.

This economic conversion is seen as an integral part of the process toward monetary union. Indeed, the Maastricht Treaty is based on the idea that economic convergence is a prerequisite for monetary union. The treaty creates a series of criteria which countries must meet to join the monetary union. These criteria are designed to ensure that potential entrants share a commitment to that union.

Much has been written critiquing the usefulness of economic convergence prior to monetary union.[2] Some papers, such as De Grauwe (1994), focus on whether the convergence indicators detailed in the treaty are the proper indicators to ensure a well-functioning monetary union. This article does not enter this discussion; rather, given the criteria established by the Maastricht Treaty, it assesses the progress of the members of the European Union in meeting these criteria. After illustrating the lack of progress of the EU in meeting them, I consider the two main alternatives available to the member states that hope to achieve monetary union in the near future. One is to allow latitude in the application of the convergence criteria and the other is to view the starting date for monetary union as flexible.

BACKGROUND

Serious discussion in Europe of a move to monetary union began in 1988 with the decision of the European Council to create a Committee for the Study of Economic and Monetary Union. This committee was chaired by Jacques Delors, the president of the European Commission.[3] The Delors Committee, as it was commonly known, was given a mandate to examine the issue of EMU and to develop a program aimed at its implementation. In 1989, the committee issued a report stating:

> "Economic and monetary union in Europe would imply complete freedom of movement for persons, goods, services and capital, as well as irrevocably fixed exchange rates between national currencies and

[1] Belgium and Luxembourg already operate in a monetary union.

[2] See, for example, De Grauwe (1992) and Portes (1993).

[3] See the shaded insert, "Institutions of the European Union" on page 2 for an explanation of the institutional structure of the European Union.

INSTITUTIONS OF THE EUROPEAN UNION

The *European Commission* is the executive branch of the European Union government. The president of the commission, who serves a two-year renewable term, is chosen by the European Council. The other 19 commissioners are appointed by their national governments for four-year renewable terms. France, Germany, Italy and the United Kingdom each appoint two commissioners and the remaining 11 EU countries each appoint one commissioner. Although the president of the commission has no control over the selection of commissioners, he does control the selection of the portfolios assigned to each commissioner. During their term in office, the commissioners are expected to represent the interests of the European Union, not those of their home countries.

The *Council of Ministers* consists of the representatives of the national governments. The composition of the Council of Ministers depends on the issue being considered. For example, issues related to the Common Agricultural Policy are addressed by the agricultural ministers of the member states, whereas finance matters are addressed by the finance ministers. Within the Council of Ministers, each country is allocated a number of votes based loosely on the size of its population. France, Germany, Italy and the United Kingdom have 10 votes each. Spain has eight. Belgium, Greece, the Netherlands, Portugal and Sweden have five votes each. The remaining countries, Austria, Denmark, Finland and Ireland, have three votes each. In sum, there are 85 votes. To pass by qualified majority, a measure must receive at least 61 votes. Thus, two large states and two small states can form a blocking coalition.

The *European Council* consists of the heads of state or government of the member countries. The president of the European Commission is a non-voting member of the European Council. The presidency of the European Council rotates among the member states on a six-month basis. The European Council holds a meeting at the end of the six-month period (in December and June).

The *European Parliament* is the legislative branch of the European Union. The 626 members of Parliament are elected in national elections and serve renewable five-year terms. In the Parliament, members are grouped according to their party affiliation, not their nationality. The European Parliament is the weakest institution within the European Union, having mainly consultative powers. The exception to this weakness is in budgetary issues, over which it has considerable control. The European Parliament may dismiss the European Commission en masse, but cannot dismiss individual members of the Commission.

> finally, a single currency. This, in turn, would imply a common monetary policy and require a high degree of compatibility of economic policies and consistency in a number of other policy areas, particularly the fiscal field. These policies should be geared to price stability, balanced growth, converging standards of living, high employment and external equilibrium" (Committee for the Study of Economic and Monetary Union, 1989, p. 17).

The recommendations of the Delors Committee formed the basis for the negotiations on EMU in the Maastricht Treaty.

In the plan suggested by the Delors Report, and incorporated in the Maastricht Treaty, EMU was to be achieved in three stages. Broadly speaking, stage one would emphasize economic convergence and stage two would emphasize institutional convergence. The final steps to full EMU would occur during stage three.

During stage one, which began in July 1990, the member countries of the European Union were to achieve greater convergence in economic performance through increased policy coordination. Stage one was also to be characterized by the completion of the single internal market and removal of all

capital controls.[4] In addition, all currencies would be linked in the Exchange Rate Mechanism (ERM), and procedures would be established for budgetary policy coordination.[5] The goals for the completion of stage one have yet to be met because the currencies of five countries do not participate in the ERM.

In accordance with the Maastricht Treaty, stage two began on January 1, 1994. During this stage, the member states are to make their central banks independent. As part of the steps toward independence, central banks are prohibited from providing overdraft facilities to their governments and from directly financing the government debt. The European Monetary Institute (EMI) began operations at the start of stage two. It is charged with ensuring cooperation between national central banks and strengthening the coordination of national monetary policies. The EMI is also to begin preparations for a single currency and the conduct of a single monetary policy. Perhaps most importantly in this regard, it is to create the instruments and procedures necessary for the operation of a single European monetary policy. Also, during stage two, countries are to achieve further economic convergence, as detailed by the criteria in the Maastricht Treaty.

The most important role of the EMI is to ensure that the technical barriers to EMU are removed prior to the start of stage three. These barriers include cross-country differences in the conduct of monetary policy, financial regulations, payments systems and currencies. The EMI is studying issues related to the conduct of monetary policy. For example, should the future European Central Bank target the money supply as the German Bundesbank does, or should it target inflation, as the Bank of England does? Another issue being studied by the EMI is the design and implementation of the single currency system. This is a politically volatile issue because each country has an interest in having the new currency resemble its own.

Stage three will mark the final transition to a full-fledged monetary union. At the start of stage three, exchange rates between member countries will be permanently fixed. The governments of the member countries of the monetary union, acting in consultation with the European Commission and the European Central Bank, will determine the exchange rates at which currencies are to be fixed. The determination of these fixed exchange rates requires the unanimous consent of the member states. As the final step to EMU, individual currencies will be replaced with a common currency. Monetary policy decisions will be made by the independent, supranational European Central Bank. According to the Maastricht Treaty, stage three must start by January 1, 1999.

The exact starting date will be determined as follows. By December 1996, an inter-governmental conference comprised of the leaders of the European Union countries must meet to determine if EMU is ready to commence. Prior to this meeting, the European Commission and the EMI are to issue reports detailing the progress made by each country in meeting the convergence criteria. These reports will be sent to the Council of Ministers. The Council of Ministers will use these reports to determine:

- whether each member state fulfills the necessary conditions for the adoption of a single currency; and

- whether a majority of the member states fulfill the necessary conditions for the adoption of a single currency (*Treaty on European Union*, Article 109j.2).

The decisions of the Council of Ministers will be made on the basis of a "qualified" majority vote. The determinations of the Council of Ministers will be forwarded to the European Parliament, which will make its own recommendation on the readiness of the member states to move to the final stage of monetary union.

Taking into account the decisions of the Council of Ministers and the European Parliament, the European Council at the inter-governmental conference must then decide, again by qualified majority:

[4] In accordance with the Maastricht Treaty, Greece was allowed to maintain capital controls until the end of June 1994.

[5] The Exchange Rate Mechanism, created in 1979, set narrow margins for exchange rate fluctuations between member countries. Normally, each currency was allowed to fluctuate by ± 2.25 percentage points against any other member currency. Some currencies, however, were given wider margins of fluctuation (± 6 percentage points) to smooth their transition upon entering the ERM.

Table 1

Progress in Meeting Convergence Criteria

	Number of Criteria Met				
	1990	1991	1992	1993	1994
Belgium	2	3	3	3	3
Denmark	5	4	4	3	3
France	5	5	4	4	4
Germany	5	4	4	3	5
Greece	0	0	0	0	0
Ireland	4	4	4	3	3
Italy	0	0	0	0	0
Luxembourg	5	5	5	4	5
Netherlands	3	4	3	3	3
Portugal	0	0	0	0	0
Spain	1	1	1	1	0
United Kingdom	3	3	2	2	3
Austria	4	4	3	2	3
Finland	2	2	1	1	2
Sweden	2	3	2	1	1
Number meeting criteria	4	2	1	0	2

- whether a majority of the member states meet the necessary conditions for monetary union;
- whether it is appropriate ... to enter the third stage; and if so,
- set the date for the beginning of the third stage (*Treaty on European Union*, Article 109j.3).

If no date for the start of monetary union has been set by the end of 1997, the treaty obligates the leaders of the European Union countries to meet by July 1, 1998, to determine, based on the same procedure outlined above, which member states fulfill the conditions for monetary union. These states are then to enter the third stage on January 1, 1999. For monetary union to begin prior to 1999, a majority of countries must meet the criteria established by the Maastricht Treaty. However, in 1999, according to the treaty, EMU will commence for those countries (however few) that meet the entry conditions.

The countries that do not meet the entry conditions and are excluded from EMU will, according to the Treaty, be referred to as "member states with a derogation" (*Treaty on European Union*, Article 109k.2). This exclusion, however, need not be permanent. At least once every two years, following the guidelines outlined above, the European Council will decide by qualified majority which member states with a derogation have fulfilled the entry criteria and admit them to the monetary union.

CONVERGENCE CRITERIA

As noted above, entry into EMU is dependent upon the fulfillment of what the Maastricht Treaty calls "necessary conditions." What are these conditions? First, to facilitate the common monetary policy, each member must guarantee the independence of its central bank and pass national legislation in accordance with the protocol establishing the European Central Bank. Second, in making their reports on the progress of countries in meeting the necessary conditions, the European Commission and the EMI are to consider the progress made in developing a common currency, "the results of the integration of markets, the situation and development of the balances of payments on account and an examination of the development of unit labour costs and other price indices" (*Treaty on European Union*, Article 109j.1).

Most attention, however, has been focused on the conditions that the Maastricht Treaty says are designed to ensure "the achievement of a high degree of sustainable convergence" (*Treaty on European Union*, Article 109j.1). Convergence must be achieved in exchange rates, inflation rates, long-term interest rates and government finances. The treaty and two separate protocols detail these convergence criteria as follows:[6]

- The currency of each member state must have remained within the normal fluctuation margins of the ERM for a least two years prior

[6] See *Protocol on the Convergence Criteria referred to in Article 109j of the Treaty Establishing the European Community* (1992) and *Protocol on the Excessive Deficit Procedure* (1992).

to the examination. Specifically, a member state may not have devalued its currency against any other currency within the ERM on its own initiative.

- The average inflation rate for any member state during the year prior to the examination by the European Commission must have been no more than 1.5 percentage points above the average rate of inflation in the three best-performing countries during this same period.

- The long-term interest rate (on government bonds or comparable securities) of any member state during the year prior to the examination by the European Commission must have been no more than 2 percentage points above the average long-term interest rate of the three countries with the lowest inflation rates during this same period.

- The government budget deficit of any member state may not exceed 3 percent of that country's GDP at the time of the examination.

- The government debt of any member state may not exceed 60 percent of the country's GDP at the time of the examination.[7]

Table 1 summarizes the performance of each current EU member state in fulfilling the convergence criteria during the years 1990-94. As this table shows, the path toward convergence has not been smooth. On the basis of these five criteria, more countries met the eligibility requirement in 1990, the year before the treaty was concluded, than in any subsequent year. Denmark, France, Germany and Luxembourg met all five convergence criteria in 1990.[8] The number of countries fulfilling the criteria declined in each following year, reaching a low of zero in 1993. In 1994, the performance of the members of the European Union improved slightly, with Germany and Luxembourg meeting all five criteria.

As the performance of the countries in 1990 and 1994 is compared, only Belgium improved its overall performance on the criteria. In contrast, six countries met fewer criteria in 1994 than they met in 1990. This worsening performance reflects the crises in the ERM and a deterioration in the public finances of many countries.

Exchange Rate Criterion

Although the ERM had functioned smoothly since 1987, it was beset by a series of crises during 1992 and 1993. These crises resulted in the September 1992 withdrawal of the British pound and the Italian lira from the ERM, and the February 1993 devaluation of the Irish pound. The Portuguese escudo and the Spanish peseta were devalued several times throughout 1992 and 1993. As a result of these crises, fewer countries met the exchange rate convergence criterion in 1994 than in 1990 (see Table 2).

The exchange rate crises ended in August 1993 with the expansion of the bilateral bands from ±2.25 percent to ±15 percent for all pairs of currencies with the exception of the Dutch krona/Deutsche mark. The consensus within the European Union is that these wider bands have reduced currency speculation and thus have lessened the prospects for exchange rate crises within the ERM. Thus, no return to the narrow margins is likely. The maintenance of the expanded margins presents no problem for the fulfillment of the convergence criteria as long as the European Commission and the European Council agree that the treaty's reference to "normal fluctuating margins" means margins of ±15 percent.

In March 1995, the currencies within the ERM again experienced sharp fluctuations. The movement in the exchange markets away from dollars and into Deutsche marks caused problems for weaker currencies within the ERM. As a result of this turbulence, the escudo and the peseta were both devalued. In the absence of any further devaluations, only eight of the 15 member countries of the European Union would meet the exchange

[7] As discussed in the *Protocol on the Excessive Deficit Procedure*, the deficit and debt ratios are based on general government budgets, that is, the central government, regional or local governments and social security funds. Commercial operations of the public sector are excluded. The deficit is defined as net borrowing by the government. Net borrowing excludes any portion of the deficit that is used for "the acquisition of loans or other financial assets" by the government. Thus, for example, the funds borrowed by the German government that were in turn lent to agencies in eastern Germany do not show up in these deficit figures (Collignon and others, 1994). Privatization proceeds cannot be used to reduce the deficit, although some countries are trying to change this provision. Whereas the deficit ratio is based on net borrowing, the debt ratio is based on gross debt.

[8] If Austria had been a member of the European Union, it too would have met all five convergence criteria in 1990. Although it was not a member of the ERM, its currency has shadowed the Deutsche mark.

Table 2

Convergence Indicators: Exchange Rate

	1990	1991	1992	1993	1994
Belgium	yes	yes	yes	yes	yes
Denmark	yes	yes	yes	yes	yes
France	yes	yes	yes	yes	yes
Germany	yes	yes	yes	yes	yes
Greece	n.m.	n.m.	n.m.	n.m.	n.m
Ireland	yes	yes	yes	no	no
Italy	no	no	no	n.m.	n.m.
Luxembourg	yes	yes	yes	yes	yes
Netherlands	yes	yes	yes	yes	yes
Portugal	n.m.	n.m.	no	no	no
Spain	no	no	no	no	no
United Kingdom	no	no	no	n.m.	n.m.
Austria	n.m.	n.m.	n.m.	n.m.	n.m.
Finland	n.m.	n.m.	n.m.	n.m.	n.m.
Sweden	n.m.	n.m.	n.m.	n.m.	n.m.
Number meeting criterion	7	7	7	6	6

Notes: n.m. indicates that the country was not a member of the ERM during any part of the relevant year.
The Irish pound was devalued by 10 percent in February 1993.
The Italian lira was devalued by 3.7 percent in January 1990 when it was incorporated into the narrow (2.25 percent) bands. The lira left the ERM in September 1992.
The Portuguese escudo was devalued by 6 percent in November 1992 and by 6.5 percent in May 1993.
The Spanish peseta was devalued by 5 percent in September 1992, by 6 percent in November 1992 and by 8 percent in May 1993.

rate criterion at the end of 1996. The currencies of five countries — Finland, Greece, Italy, Sweden and the United Kingdom — are not participating in the ERM and therefore will not meet the two-year rule by the end of 1996. As a result of the recent devaluations of their currencies, Portugal and Spain will not meet the criterion by the end of 1996.

Inflation Criterion

Comparing 1990 to 1994, the performance of the EU countries with regard to the inflation criterion has improved. As shown in Table 3, seven of the present 15 EU countries met the inflation criterion in 1990. This number fell to five in 1993, but rebounded strongly with 11 countries meeting the criterion in 1994. Greece, Italy, Portugal and Spain were the countries with inflation rates exceeding the criterion in 1994. Although these four countries have not met the criterion in any year, each country has made progress in lowering its inflation rate over the period in question.

The economic recovery currently under way in Europe is expected to lead to a slight increase in inflation in most member countries by 1996. Because the criterion is based on the performance of the three countries with the lowest inflation, a general increase in the rate of inflation will not affect the overall performance of countries. As shown in Table 3, the increase in the inflation forecast for 1996 is not expected to reduce the number of countries satisfying the inflation criterion. Moreover, the inflation performance of the countries not currently meeting the criterion is expected to improve over the next two years.

Interest Rate Criterion

The interest rate criterion has been the one that countries have usually found easiest to meet. Furthermore, the member countries showed steady improvement over the period 1990-94. In 1990, as shown in Table 4, nine countries had long-term interest rates within the limit set forth in the Maastricht Treaty. This number rose to 10 in 1991 and increased to 11 in 1993. In 1994, however, the number of countries meeting the interest rate criterion slipped back to 10. In 1994, Greece, Italy, Portugal, Spain and Sweden did not meet this criterion. The former four have never met the criterion.

Public Finance Criteria

The two public finance criteria have caused the biggest problems for countries in their quest to join the EMU. In 1990, nine of the current 15 EU countries met the deficit criterion while only three met it in 1994. Similarly, nine countries met the government debt criterion in 1990 but only four

Table 3

Convergence Indicators: Inflation

Percent

	1990	1991	1992	1993	1994	1995	1996
Belgium	3.7	2.5	2.1	2.6	2.4	1.9	2.4
Denmark	2.7	2.4	1.8	1.0	1.7	2.3	2.7
France	2.8	3.2	2.4	2.2	1.8	1.9	2.1
Germany	2.8	4.0	4.7	3.8	2.7	2.3	2.5
Greece	19.2	18.8	15.1	13.6	10.9	9.6	8.9
Ireland	1.4	2.5	2.8	1.6	3.0	2.9	2.7
Italy	5.9	6.9	5.2	5.1	4.7	5.2	4.5
Luxembourg	3.6	2.9	2.8	3.6	2.2	2.3	2.5
Netherlands	2.2	3.2	3.0	2.1	2.2	1.8	2.2
Portugal	11.7	12.5	10.0	7.9	5.1	4.5	4.5
Spain	6.5	6.4	6.4	5.6	5.1	4.9	4.5
United Kingdom	5.5	7.4	4.7	3.4	2.5	3.0	3.0
Austria	3.1	3.4	3.9	3.5	3.3	2.8	2.9
Finland	6.0	5.6	4.1	3.9	1.6	1.7	3.3
Sweden	9.6	10.2	2.2	5.8	3.0	3.2	3.2
Convergence criterion	3.6	4.0	3.6	3.1	3.4	3.3	3.7
Number meeting criterion	7	8	7	5	11	11	11

Notes: Prior to 1992, data for Germany is for western Germany only.
Data for 1995 and 1996 are forecasts
Convergence criterion is based on data for the 12 member states prior to 1995 and the 15 states thereafter.
SOURCE: *European Economy* (April/May 1995, Supplement A, Table 10)

did in 1994. Much of this decline can be attributed to the expansionary nature of fiscal policies in reaction to the recession of the early 1990s, from which Europe is just beginning to recover. The effect of the recession on public finances can be seen by considering the example of Finland. Output growth in Finland fell from 5.7 percent in 1989 to -7.1 percent in 1991. Consequently, Finland's government budget balance declined from 5.4 percent of GDP in 1990 to a low of -7.8 percent in 1993. The government budget deficit shrank in 1994 as its economy moved out of recession.

The economic recovery currently under way in Europe is expected to lead to a gradual improvement in the budget balances of the EU countries. Nevertheless, only six of the 15 countries are expected to meet the budget deficit criterion in 1996. The recovery is expected to have less of an effect on countries' performance with respect to the debt criterion. The ratio of debt to GDP is expected to increase through 1996 in most countries.

The criterion limiting the government debt to 60 percent of GDP has been the most difficult for countries to meet. Only Luxembourg has a debt ratio well below that level. The other three countries that met this criterion in 1994 (France, Germany and the United Kingdom) all have debt-to-GDP ratios close to 50 percent. Among those countries not meeting the criterion, some have debt ratios so high that they would have to run substantial budget surpluses for a number of years to meet it. For example, Buiter, Cosetti and Roubini (1993) calculated that based on the 1991 debt levels and assuming a 5 percent nominal GDP growth

Table 4

Convergence Indicators: Long-Term Interest Rates

Percent

	1990	1991	1992	1993	1994
Belgium	10.1	9.3	8.6	7.2	7.8
Denmark	11.0	10.1	10.1	8.8	8.5
France	10.4	9.5	9.0	7.0	7.5
Germany	8.9	8.6	8.0	6.3	6.7
Greece	18.5	18.8	17.7	18.2	n.a.
Ireland	10.1	9.2	9.1	7.7	8.1
Italy	13.4	13.0	13.7	11.3	10.6
Luxembourg	8.6	8.2	7.9	6.9	6.4
Netherlands	9.0	8.7	8.1	6.7	7.2
Portugal	16.8	18.3	15.4	12.5	10.0
Spain	14.7	12.4	12.2	10.2	9.7
United Kingdom	11.8	9.9	9.1	7.8	8.2
Austria	8.7	8.6	8.3	6.6	6.7
Finland	13.2	11.9	12.1	8.2	8.4
Sweden	13.6	10.9	10.4	8.5	9.5
Convergence criterion	12.0	11.5	11.2	9.7	9.4
Number meeting criterion	9	10	10	11	10

SOURCES: *European Economy* (1995, Number 59, Table 54) and *OECD Economic Outlook* (June 1995, Number 57, Annex Table 36)

rate, Belgium needs a government surplus of more than 9 percent of GDP a year for each year through 1996 to meet the convergence criteria. To meet the criteria by the end of 1998, Belgium would need an annual government surplus greater than 5 percent of GDP.

Summary on Convergence

To summarize, the data indicate that inflation and interest rate convergence are taking place in the European Union. The outlook for the next two years anticipates further convergence with respect to these two criteria. In contrast, the public finances of the EU members have worsened since the establishment of the convergence criteria. Although the government budget balances of most member states are expected to improve through 1996, the debt ratios are unlikely to show significant improvement. Turning to the exchange rate criterion, five countries are not members of the ERM and thus do not meet the convergence criterion. For the remaining 10 counties, although the wider bands eliminated tensions within the ERM between August 1993 and March 1995, there is now evidence that even these bands cannot prevent pressure from accumulating on weak currencies.

PROSPECTS FOR EMU

For the 1996 inter-governmental conference to set a date for monetary union, eight countries must fulfill all of the convergence criteria. If there are no further devaluations within the ERM, eight countries—Austria, Belgium, Denmark, France, Germany, Ireland, Luxembourg and the Netherlands—will fulfill the exchange rate criterion in 1996. Thus, if EMU is to get off the ground prior to 1999, all eight of these countries must meet the other four convergence criteria. However, the debt/GDP ratios of four of these countries—Belgium, Denmark, Ireland and the Netherlands—are not expected to be close to the 60 percent reference value by the end of 1996.

Thus, based on the five convergence criteria, it is almost certain that a majority of the EU countries will not be ready for monetary union when the inter-governmental conference is held in 1996. If EMU is postponed, the next issue is: How many countries will be eligible at the start of 1999, the last possible date for monetary union in accordance with the treaty? Barring unforeseen economic shocks, Germany and Luxembourg should both be eligible for monetary union. The eligibility of the remaining 13 countries is less certain, even leaving aside the uncertain future of the ERM. Austria and France are the most likely additional candidates. Both, however, could run into problems meeting the government budget requirement, and Austria is not expected to meet the debt criterion.

Belgium and Italy have public debts totaling more than 100 percent of their

respective GDPs. It will be many years before these debt ratios come close to meeting the 60 percent limit. Denmark, Finland, Ireland, the Netherlands and Sweden also have high debt ratios unlikely to fall within the target range by the end of the century. The Dutch central bank last year calculated that if the Netherlands limited its annual public sector deficit to 1 percent of GDP, and achieved an average nominal GDP growth of 4 percent a year, it would still take 10 years to reach the 60 percent public debt target (*Financial Times*, January 17, 1995). While 4 percent was the average nominal GDP growth for the Netherlands during 1985-94, its average yearly budget deficit has been more than double 1 percent of GDP over the last 10 years.[9]

Portugal and Spain are likely to have difficulty meeting several of the criteria. Although they both have substantially lowered their inflation rates in recent years, the 5.1 percent Portuguese and Spanish inflation rates remain outside the ceiling. The debt ratios of both countries also have grown recently and that of Spain is likely to remain a problem as long as it maintains its high unemployment rate (estimated at more than 22 percent in 1994). No one expects that Greece will be a candidate for monetary union for many years to come. It alone among the EU countries still has double-digit inflation.

The remaining country, the United Kingdom, is a good candidate for meeting all of the eligibility requirements for monetary union, except the exchange rate criterion. The United Kingdom is unlikely to rejoin the ERM in the next few years. Even ignoring this problem, opposition to EMU is strong within the British government and Britain is one of two European Union countries that have the right to refuse entry into the monetary union.[10] A change in the government from the ruling Conservative party to the opposition Labour party is likely to increase the prospects for Britain joining EMU simply because the latter is much more amenable to the idea of monetary union than the former.

Responses to the Lack of Progress in Meeting the Convergence Criteria

The reality that a majority of countries will not meet the convergence criteria in 1996, and that most, including some key countries, are unlikely to meet the criteria in 1998, has generated three responses within the European Union. One reaction has been to label the idea of monetary union impractical. A second suggests that the public finance criteria for monetary union can be and should be interpreted with some leeway. A third reply suggests that the timetable for monetary union should be interpreted with some flexibility.

Abandoning EMU

Those who have reacted to the difficulty in meeting the convergence criteria by labeling EMU impractical are basically opposed to the idea of monetary union. They see the lack of progress in meeting the criteria as a means to gain support for the idea of abandoning the treaty. Proponents of this view, most notably some members of the British Parliament, have reacted to each crisis within the ERM with predictions of the demise of monetary union. For example, British Prime Minister John Major responded to the August 1993 widening of the bands of the ERM with the statement that the Maastricht timetable for monetary union was now "totally unrealistic." The reaction of Norman Lamont, the former chancellor of the exchequer in Britain, was even more pointed. He claimed that the crisis in the ERM meant "the end of monetary union in Europe" (*Financial Times*, August 3, 1993). In practice, this group supports strict adherence to the convergence criteria, since this will delay the starting date for monetary union.

Flexibility in Interpreting the Convergence Criteria

In opposition to this group are those who not only support EMU but believe that the earlier the starting date the better. This latter group favors a liberal interpretation of the convergence criteria. One reason for

[9] A reduction in public debt can occur through several means besides a government surplus. Both nominal GDP growth and a reduction in interest rates on government debt will reduce the debt/GDP ratio. Nominal GDP growth may be achieved through growth in output or inflation. This might lead one to think that inflating away the debt would be a compelling option. Such a strategy, however, will only work in the short run. An increase in inflation raises the interest rate at which the government must borrow to finance its debt. The shorter the maturity of the outstanding debt, the shorter the period of time before which the engineered inflation will affect the interest rate on the debt. Furthermore, any such attempt by the government to meet the debt convergence criterion through inflation is likely to have long-term repercussions for the interest rate at which the government borrows by reducing the government's credibility.

[10] In Maastricht, the United Kingdom refused to conclude negotiations on the treaty unless it was given the right to opt-out of EMU. Denmark is the other country with the right to opt-out of monetary union. It negotiated this right following the rejection of a referendum on the treaty. After securing the opt-out provision, a new referendum approved the treaty.

supporting a quick move to monetary union is the belief that a long transition period may itself be the source of instability. A proponent of this view is Portes (1993). In addition to arguing that a long transition period creates instability, Portes contends that the convergence criteria are unnecessary because "monetary union will deliver convergence—at least the extent required to maintain it." De Grauwe (1994) takes this argument one step further by claiming that the convergence criteria cannot be met prior to EMU.

Although support for a quick move to monetary union is generally tied to the belief that convergence is not a necessary prerequisite for EMU, support for a flexible approach to the criteria is based on additional reasons. One is to provide a wide participation in EMU. Another is the fear among countries that have little chance of meeting the requirements that non-participation in EMU will be costly both politically and economically. In the political sphere, countries are afraid that remaining outside EMU will reduce their political power within the EU, particularly as the inner core of countries (the members of EMU) become more interdependent. In economic terms, countries are concerned that exclusion from EMU may be viewed as a mark against them, and result in a higher interest rate premium and a weakness in their currencies.

Supporters of a flexible approach to the convergence criteria make reference to the Maastricht Treaty to bolster their case. The treaty provides an opening for a relaxation of both the deficit and the debt criteria. The 3 percent deficit/GDP ratio and the 60 percent debt/GDP ratio are referred to in the treaty as reference values, not fixed limits as are the criteria for inflation and interest rates. The treaty says that these reference values must be met unless, in the case of the deficit:

- either the ratio has declined substantially and continuously and reached a level that comes close to the reference value; or

- alternatively, the excess over the reference value is only exceptional and temporary and the ratio remains close to the reference value (*Treaty on European Union,* Article 104c.2.a).

In addition, in preparing its report on whether an excessive deficit exists, the Commission is to take into account:

- whether the government deficit exceeds government investment expenditure (gross fixed capital formation); and

- all other relevant factors, including the medium-term economic and budgetary position of the Member State (*Treaty on European Union,* Article 104c.3).

These clauses provide the commission a means by which to relax the deficit requirement. As noted by Collignon and others (1994), the treaty could be interpreted as applying the deficit criterion to only the part of the deficit not accounted for by government investment, and only requiring the 3 percent ratio to be met "when the economy was near full capacity." Looking at the data in Table 5, one could argue that Austria, Denmark and the Netherlands all meet the deficit criterion since their budget deficits remain close to the reference level, and that the elevated levels are merely temporary — caused by the recession.[11]

With respect to the debt criterion, the Maastricht Treaty states that the reference level (60 percent debt/GDP) is binding "unless the ratio is sufficiently diminishing and approaching the reference value at a satisfactory pace" (*Treaty on European Union,* Article 104c.2.b).

The debt levels of all the countries, with the exception of Ireland and the Netherlands, have increased between 1990 and 1994, as shown in Table 6. In Ireland's case, substantial progress has been made in reducing its debt ratio. Ireland has met the deficit convergence criterion in every year and has reduced its debt ratio from 97 percent of GDP in 1990 to 90 percent in 1994. In the fall of 1994, the European Council, assessing the progress of countries toward the

[11] During the Maastricht negotiations, several countries proposed adopting a concept of cyclically adjusted deficits. The proposal was rejected because of measurement problems (Bini-Smaghi and others, 1994).

Table 5

Convergence Indicators: Government Budget Balance

Percent of GDP

	1990	1991	1992	1993	1994	1995	1996
Belgium	-5.4	-6.5	-6.7	-6.6	-5.3	-4.2	-3.9
Denmark	-1.5	-2.1	-2.9	-4.5	-4.0	-1.9	-1.2
France	-1.6	-2.2	-3.9	-6.1	-6.0	-4.9	-3.9
Germany	-2.1	-3.3	-2.9	-3.3	-2.5	-2.1	-2.4
Greece	-14.0	-11.6	-12.3	-13.2	-12.5	-11.3	-10.2
Ireland	-2.2	-2.1	-2.4	-2.4	-2.3	-2.8	-2.6
Italy	-10.9	-10.2	-9.5	-9.6	-9.0	-7.9	-8.1
Luxembourg	5.9	2.3	0.8	2.1	2.3	1.4	1.5
Netherlands	-5.1	-2.9	-3.9	-3.3	-3.1	-3.2	-2.5
Portugal	-5.5	-6.5	-3.3	-7.0	-5.8	-5.6	-4.7
Spain	-3.9	-4.9	-4.2	-7.5	-6.6	-6.0	-4.8
United Kingdom	-1.5	-2.6	-6.1	-7.8	-6.9	-4.8	-2.9
Austria	-2.2	-2.4	-2.0	-4.1	-4.0	-4.6	-3.9
Finland	5.4	-1.5	-5.9	-7.8	-5.6	-5.0	-1.1
Sweden	4.2	-1.1	-7.8	-13.4	-10.4	-9.1	-5.8
Convergence criterion	-3.0	-3.0	-3.0	-3.0	-3.0	-3.0	-3.0
Number meeting criterion	9	9	5	2	3	4	6

Notes: Prior to 1991 the data for Germany are for western Germany only.
Data for 1995 and 1996 are forecasts.
SOURCE: *European Economy* (April/May 1995, Supplement A, Table 21).

Maastricht criteria, accepted the recommendation of the commission and determined that Ireland met the debt criterion. This decision indicates some willingness on the part of the European Union to reward countries that are making efforts to control public deficits yet remain outside the numerical targets. However, it does not mean that such a policy will be followed at the inter-governmental conference in 1996. The decision that Ireland met the debt convergence requirement was not without controversy. Germany, in particular, had severe reservations about the exemption. Furthermore, while the previous European Commission, the term of which ended in December 1994, supported a flexible interpretation of the convergence criteria, it is not clear that the present commission also supports this view. In May 1994, the then-commissioner for economic and monetary affairs stated that it had "always been understood that the judgement on whether a member state fulfills the conditions for participation in stage 3 would be based on an assessment, and not on a mechanical application of the convergence criteria" (*Financial Times*, May 16, 1994). In contrast, the president of the current commission, Jacques Santer, in his first speech before the European Parliament, pledged that the commission would insist on strict application of the criteria (*Financial Times*, January 18, 1995).

Flexibility in the Starting Date for Monetary Union

In contrast to those who have responded to the lack of progress in meeting the convergence criteria by suggesting that the

Table 6

Convergence Indicators: Government Debt

Percent of GDP

	1990	1991	1992	1993	1994	1995	1996
Belgium	130.8	130.1	131.1	137.2	136.1	134.3	132.3
Denmark	59.6	64.6	69.0	80.3	75.6	76.1	75.4
France	35.4	35.7	39.6	45.8	48.5	51.2	52.8
Germany	43.8	41.5	44.1	48.2	50.1	58.2	58.1
Greece	82.6	86.1	92.3	115.2	114.1	115.3	116.2
Ireland	96.8	96.9	94.2	97.0	89.8	84.6	86.8
Italy	97.9	101.3	108.4	119.4	125.4	124.9	124.4
Luxembourg	5.4	4.9	5.7	6.9	7.2	7.6	7.8
Netherlands	78.8	78.9	79.9	81.4	78.1	78.1	77.1
Portugal	68.6	69.3	61.7	66.6	69.2	70.5	70.7
Spain	45.1	45.8	48.3	59.9	62.3	64.6	65.2
United Kingdom	35.0	35.7	41.9	48.5	50.1	51.5	51.5
Austria	58.3	58.7	58.4	62.8	64.5	66.2	67.4
Finland	14.5	23.0	41.5	57.1	60.1	64.4	64.6
Sweden	43.5	53.0	67.1	76.2	79.1	84.6	85.7
Convergence criterion	60.0	60.0	60.0	60.0	60.0	60.0	60.0
Number meeting criterion	9	8	7	6	4	4	4

Notes: Data for the United Kingdom in 1990 are based on OECD calculations of general government gross financial liabilities. All other data are based on the Maastricht Treaty's definition of public debt.
Prior to 1991 the data for Germany are for western Germany only.
Data for 1995 and 1996 are forecasts.

SOURCES: *European Economy* (April/May 1995, Supplement A, Table 22) and *OECD Economic Outlook* (June 1995, Number 57, Table 34).

convergence criteria be treated with flexibility are those who believe that the 1999 deadline should be viewed as flexible. The proponents of a flexible timetable believe that strict adherence to the convergence criteria is a necessary condition for a well-functioning monetary union. Thus, rather than relaxing the criteria to guarantee that an optimal number of countries will participate in EMU, they suggest that the date for monetary union be delayed if the criteria are not met by a sufficient number of countries. German Chancellor Helmut Kohl was the first leader to publicly address this issue. In 1993, he stated that strict adherence to the convergence criteria might delay monetary union beyond 1999.

The October 1993 ruling of the German Constitutional Court supported those who argue that the timetable for monetary union is more flexible than the criteria. The court, in ruling on the constitutionality of the Maastricht Treaty, wrote that strict adherence to the convergence criteria was essential to Germany's participation in EMU. In other words, the criteria could not be weakened without the consent of the German parliament.

The German central bank, the Bundesbank, has been perhaps the most vocal advocate of a strict application of the convergence criteria. Both Hans Tietmeyer, the current president of the bank, and his predecessor, Helmut Schlesinger, have made statements on several occasions favoring a strict interpretation of the Maastricht criteria while claiming that the criteria are themselves not strict enough. For example, the

Bundesbank has favored an absolute limit on inflation rather than a relative one, the latter based on the behavior of other countries. The reason for this is to ensure not simply convergence in inflation rates, but also a commitment to price stability. The Bundesbank has also attacked the deficit criterion as setting too high a ceiling. Specifically, Mr. Tietmeyer has stated that the ceiling for the deficit ratio is at least double what it should be. He also has emphasized that the deficit criterion should be met throughout the business cycle (*Financial Times,* November 5, 1994).[12] This statement contrasts with a study prepared for the European Parliament that suggests that "It would be keeping with the spirit of the Treaty, if 3 percent were taken as the 'full employment' deficit during periods of economic expansion" (Collignon and others, 1994, p. 76).

As noted above, the emphasis on a strict interpretation of the convergence criteria is based on the belief that adherence to them is necessary for a well-functioning monetary union. The proponents of strict criteria argue that for EMU to succeed, the member states must show a prior commitment to price stability and follow sound government budgetary policies. Specifically, the emphasis on a strict interpretation of the deficit criterion is based on the idea that "a sound budget position is an indispensable precondition for a successful anti-inflationary monetary policy."[13] There is a concern that within a monetary union, expansionary national fiscal policies (as evidenced by budget deficits in excess of 3 percent of GDP) could conflict with the monetary policy of the supranational central bank. Such a conflict would not only create difficulties for the central bank in its effort to maintain price stability, but also could cause tension among the participants in the monetary union. Would the participants of a monetary union be willing to accept a recession brought about by the anti-inflationary polices of the central bank in an effort to combat the fiscal laxity of other members? Furthermore, although the Maastricht Treaty prohibits the central bank from extending credit to, or directly purchasing the debt of, member states (*Protocol on the Statue of the European System of Central Banks and the European Central Bank,* Article 21), and declares that neither the central bank nor other countries shall be liable for or assume the financial commitments of any member states (*Treaty on European Union,* Article 104b.1), there are those who believe there would be pressure on the central bank to bail out countries experiencing financial difficulties.[14]

CONCLUSION

Despite the many setbacks that have occurred since the December 1991 conclusion of the Maastricht Treaty, most of the countries of the European Union remain committed to monetary union. This commitment, however, has not been enough to produce the economic convergence prescribed by the treaty. Many countries have made progress in reducing their inflation rates, and the divergence in long-term nominal interest rates is declining. On the fiscal side, however, the number of countries meeting the convergence criteria has declined. The recent recession in Europe resulted in a deterioration in the fiscal balances of most countries. In addition, the 1992-93 exchange rate crises resulted in a reduction in the membership of the ERM. Thus, the European Union is further away from a fulfillment of the convergence criteria today than it was in the year prior to the negotiation of the Maastricht Treaty.

By the end of 1996, the member states of the European Union must decide if a majority of countries are ready to proceed to EMU in 1997. As detailed above, it is implausible that a majority of countries will have fulfilled the convergence criteria by the end of 1996. EMU will most certainly be delayed beyond its earliest possible starting date. The Maastricht Treaty states that the final stage of EMU must begin by January 1, 1999, with the membership decided by July 1998. Even by this date, few countries are likely to satisfy the convergence criteria.

Given the lack of progress in meeting the convergence criteria, the European Union faces two options if it is to continue to pursue EMU: Relax the criteria or relax the

[12] Others have claimed that even a government that has a balanced budget during upturns could have a budget deficit exceeding the Maastricht limit during a recession. See, for example, Eichengreen (1992) and Kenen (1992). Eichengreen argues that it may even be optimal for disciplined governments to occasionally have deficits exceeding 3 percent of GDP (p. 50).

[13] Tietmeyer (September 9, 1994).

[14] Support for this view is given by Fratianni, von Hagen and Waller (1992) and Craig (1994).

timetable for monetary union. Which option it chooses will likely not be decided until the July 1998 deadline for determining the membership of EMU. The choice taken by the EU will undoubtedly be influenced by the two countries without whose participation EMU will not occur: France and Germany.

Germany has strongly opposed a relaxation of the convergence criteria.[15] If it maintains this position, few countries are likely to meet the membership requirements for EMU by the end of the decade. More importantly, two countries considered among the core group of EU countries — Belgium and the Netherlands — are not expected to meet the criteria.[16] Without the participation of the core group, monetary union may not be feasible. Thus, it is likely that EMU, like its avian namesake, will remain grounded.

REFERENCES

Barber, Lionel. "Santer steers middle course for EU," *Financial Times* (January 18, 1995).

Bini-Smaghi, Lorenzo, Tommaso Padoa-Schioppa and Francesco Papadia. "The Transition to EMU in the Maastricht Treaty," *Essays in International Finance* (Number 194, November 1994), Princeton University.

Buiter, Willem, Giancarlo Corsetti and Nouriel Roubini. "Excessive Deficits: Sense and Nonsense in the Treaty of Maastricht," *Economic Policy* (April 1993), pp. 57-100.

Collignon, Stefan with Peter Bofinger, Christopher Johnson and Bertrand de Maigret. *Europe's Monetary Future.* Printer Publishers, 1994.

Committee for the Study of Economic and Monetary Union. *Report on Economic and Monetary Union in the European Community.* Office for Official Publications of the European Communities, 1989.

Council of the European Communities, Commission of the European Communities. *Protocol on the Convergence Criteria Referred to in Article 109j of the Treaty Establishing the European Community.* Office for Official Publications of the European Communities, 1992.

_____. *Protocol on the Excessive Deficit Procedure.* Office for Official Publications of the European Communities, 1992.

_____. *Protocol on the Statue of the European System of Central Banks and of the European Central Bank.* Office for Official Publications of the European Communities, 1992.

_____. *Treaty on European Union.* Office for Official Publications of the European Communities, 1992.

Craig, R. Sean. "Who Will Join EMU? Impact of the Maastricht Convergence Criteria on Economic Policy Choice and Performance," Board of Governors of the Federal Reserve System *International Finance Discussion Papers No. 480* (September 1994).

De Grauwe, Paul. "Towards European Monetary Union Without the EMS," *Economic Policy* (April 1994), pp. 147-85.

_____. *The Economics of Monetary Integration.* Oxford University Press, 1992.

The Economist. "The Unpopularity of Two-Speed Europe" (September 14, 1991).

Eichengreen, Barry. "Should the Maastricht Treaty Be Saved?" *Princeton Studies in International Finance* (December 1992).

Fratianni, Michele, Jurgen von Hagen and Christopher Waller. "The Maastricht Way to EMU," *Essays in International Finance* (June 1992).

Gawith, Philip. "Tietmeyer Sets Out Tough Line on EMU Convergence Criteria," *Financial Times* (November 5, 1994).

Marsh, David. "Emu Strain Begins to Show." *Financial Times* (January 17, 1995).

Owen, David. "Major Says Timetable for Monetary Union Unrealistic," *Financial Times* (August 3, 1993).

Portes, Richard. 1993. "EMS and EMU After the Fall," *The World Economy* (No, 1, 1993) pp. 1-16.

Tietmeyer, Hans. Speech at the Association for the Monetary Union of Europe in Frankfurt/Main, September 9, 1994. Printed in Bank for International Settlements, *BIS Review* (October 14, 1994).

[15] It is interesting to note that in the negotiations on the Maastricht Treaty, Germany resisted setting a fixed date for the commencement of monetary union. It believed that fixing a date would result in a loose application of the convergence criteria (*The Economist*, September 14, 1991).

[16] The other members of this core group are France, Germany and Luxembourg.

Donald S. Allen is an economist at the Federal Reserve Bank of St. Louis. Thomas A. Pollmann provided research assistance.

Changes in Inventory Management and the Business Cycle

Donald S. Allen

"I remember one day in the summer of 1975 when a CBO [Congressional Budget Office] staffer returned from a congressional hearing with some amazing news. Alan Greenspan, then President Gerald Ford's chief economic adviser, had just testified that the recession was mostly an inventory correction. We all snickered at the idea that what was, up to then, the deepest recession since the Great Depression could have been 'only' an inventory cycle. When I subsequently studied the data more carefully, however, I learned that Greenspan had been right. Like most of the recessions before and since, the 1973-5 contraction was dominated by changes in inventory investment."

Alan S. Blinder, introduction to *Inventory Theory and Consumer Behavior* (1990)

The change in business inventories is usually less than 1 percent of total Gross Domestic Product (GDP), yet during cyclical contractions this component contributes disproportionately to the change in GDP. As a result, most cyclical contractions have been referred to as inventory cycles. These inventory cycles are characterized by an unanticipated drop in demand resulting in unplanned increases in inventories. Firms respond by cutting production to reduce inventory. This cut in production can exacerbate the downturn by reducing demand further.

Since the 1970s, many firms have made notable changes in their inventory management methods. In particular, large movements in interest rates in the early 1980s and increased global trade have combined to motivate firms to reduce inventory levels relative to sales as part of larger downsizing efforts. More efficient inventory management has been realized by implementing "just-in-time" (JIT) management techniques and the use of bar codes. Will these innovations in inventory management decrease the effect of inventory movements on the business cycle? This article investigates the extent of the changes in inventory management and makes some observations regarding inventory movement and the business cycle. There is evidence to suggest that the use of these innovative inventory control methods is on the rise, but the net effect on the business cycle remains ambiguous.

In the first two sections, I review the role of inventory investment in postwar recessions and the motivations for holding inventory. Next, I document some of the innovations in inventory management that firms have adopted over the last 10-15 years. Finally, I discuss the potential impact of these changes on the business cycle.

THE ROLE OF INVENTORY IN POSTWAR RECESSIONS

The stocks of materials and supplies, partially completed goods and finished goods in the possession of a firm are income-producing assets. These stocks are held temporarily before being sold. As inventories are increased or decreased between the beginning and the end of a period, they add to or subtract from the investment component of GDP. Unlike fixed investment, which is assumed to be the result of specific plans by firms, inventory stocks fluctuate as a result of both active decisions by firms and errors in forecasted demand. This dual effect tends to make inventory investment especially volatile around contractions, usually going

Table 1

U.S. Inventory Investment Movements in Postwar Recessions

Recession Period Peak to Trough[1]	Change in Real GDP[2]	Change in Inventory Investment[2]	Change in Inventory Investment as a Percentage of Change in Real GDP
1948:4 - 1949:4	–14.5	–28.3	195.2
1953:2 - 1954:2	–36.9	–20.0	54.2
1957:3 - 1958:1	–61.1	–21.1	34.5
1960:1 - 1960:4	–15.8	–45.5	288.0
1969:3 - 1970:4	–11.2	–32.2	287.5
1973:4 - 1975:1	–135.1	–84.7	62.7
1980:1 - 1980:2	–98.2	–10.7	10.9
1981:3 - 1982:3	–110.1	–35.0	31.8
1990:2 - 1991:1	–75.1	–44.5	59.3
Mean			113.8

[1] Peaks and troughs correspond to peaks and troughs of real GDP and do not always coincide with official NBER recession dates. .
[2] Billions of 1987 dollars.

from positive at the beginning as a result of unintended accumulation, to negative due to deliberate reduction.

Inventory investment averages less than 1 percent of GDP, but changes in inventory investment can account for a substantial portion of the *change*. In 1994, for example, inventory investment was $47.8 billion (in 1987 dollars) or 0.9 percent of GDP. This level of inventory investment reflected an increase of $32.5 billion over 1993 or 15.5 percent of the $209.5 billion increase in GDP in 1994.

The typical inventory cycle begins with an unexpected reduction in demand which leaves firms with inventory above their desired levels. Production is reduced to lower inventory levels, which can result in layoffs and further reduction in demand. As inventory falls back to desired levels and demand resumes, production may be insufficient to meet demand and maintain inventory levels. The result is that inventory can fall below the desired level, causing increased production to replenish inventory. The degree of undesired accumulation and decumulation is a function of the accuracy of firms' demand projections. This inventory cycle phenomenon has been of varying interest to economists, and research in this area has ebbed and flowed like the business cycle. Metzler (1941) showed analytically how inventory cycles could be generated when decisions on production levels are based on expected levels of sales, and income and demand are determined by production levels. Blinder and Maccini (1991) provide a good survey and bibliography of research in inventory cycles since Metzler's work.

The role of inventory investment in business cycle contractions has been well documented. Coincident declines in GDP and inventory investment are empirical regularities of postwar business cycles. Blinder and Maccini (1991) show that the average movement in inventory investment during recessionary periods in the postwar era account for 87 percent of Gross National Product (GNP) movement from peak to trough. Computed another way, the relative movement is even greater. Table 1 shows peak-to-trough movement in inventory investment compared to the peak-to-trough change in GDP in all postwar recessions.[1] The average percentage change in inventory investment to change in GDP is 113.8 percent. Admittedly, this method computes the difference between the highest quarterly increase in business inventory and the highest quarterly decrease in business inventory on an annualized basis, capturing the widest swing. However, it is evident that inventory investment has been a significant contributor to changes in GDP during contractions. Figure 1 compares the change in GDP and the change in business inventories since 1948. Recessions are shown by shaded bars.

Inventory level movement by itself is not the complete story. It is necessary to know whether movements are active responses to changes in the level of demand or reflect errors in forecasting. The ratio of inventory to sales, defined as total stocks divided by monthly sales, gives some indication of the nature of these movements. If we assume

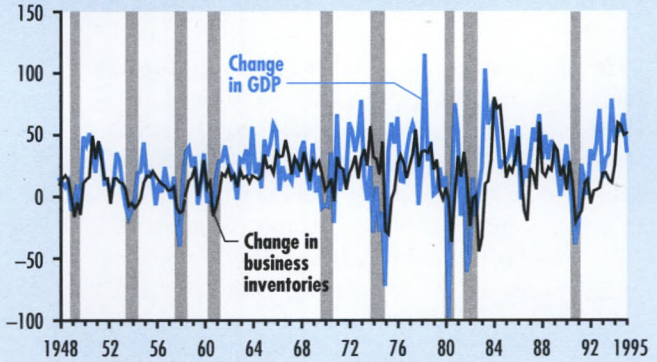
[1] The National Bureau of Economic Research (NBER) typically has identified a recession as a period with two consecutive quarters of decline in GDP. The peak of the cycle is the quarter prior to the first quarter of decline. The trough is the last quarter of negative growth. The peak-to-trough movement in inventory investment in Table 1 is the difference between the maximum and minimum inventory investment during the recession period.

that firms plan to maintain a relatively constant level of inventory to sales, major deviations in this ratio can give clues to whether movements are planned or unplanned. If inventory accumulation is accompanied by an increasing inventory-to-sales ratio, then the accumulation may be inadvertent because inventory is rising faster than sales. If inventory accumulation is accompanied by a constant inventory-to-sales ratio, then the accumulation may have been planned in response to increasing sales. An increase in inventory can also be accompanied by a decrease in the inventory-to-sales ratio, indicating that sales are increasing faster than inventories.

The total business inventory-to-sales ratio in the postwar period is shown in Figure 2, with recessions indicated by shaded bars. It is also evident from this figure that the ratio peaks around the contractions, making it a relatively reliable coincident indicator. Although it cannot be claimed that inventory changes cause the business cycle, any imbalance which occurs between expected and actual sales shows up in inventory, and correcting this imbalance can exacerbate the cycle. Even if we do not consider inventory investment to be a causative force but simply a barometer of forecast accuracy, most recessions appear to be marked by an inventory correction.

Figure 1

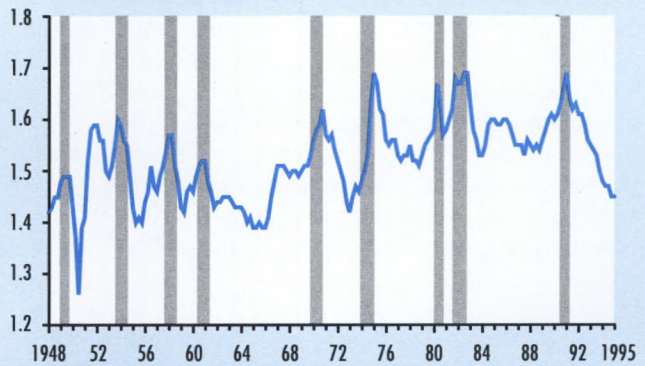
Change in GDP Compared to the Change in Business Inventories



SOURCE: U.S. Department of Commerce, Bureau of Economic Analysis.

Figure 2

Total Business Inventory-to-Sales Ratio



SOURCE: U.S. Department of Commerce, Bureau of Economic Analysis.

WHY HOLD INVENTORY IN THE FIRST PLACE?

Inventory stocks represent a major utilization of resources. At the end of 1994, manufacturing and trade inventories totaled $832 billion (1987 dollars) or 12.4 percent of annual sales. At the current prime rate, the opportunity cost of holding the 1994 level of inventory stocks amounts to more than $70 billion. This financing cost compares to the 1994 annual increase in GDP of roughly $200 billion. The capital tied up in financing inventory could also be converted into fixed investment in more productive capital equipment. But rational firms are motivated to hold inventory as long as the expected cost of holding it is less than the expected penalty (lost revenue or market share) for running out of stock. In other words, the optimal level of inventory in the face of uncertain sales and random supply interruption is not always zero, and there is a limit to the savings which can be realized by lowering inventory.

The motivations for holding inventories are diverse and firm-specific. Some firms minimize their delivery costs, some smooth production in the face of uncertain demand, and others stockpile against potential interruptions or anticipated price increases by suppliers. Most retailers are forced to hold inventory to accommodate the

INVENTORY MODELS

Production Smoothing

The accompanying figure on page 21 illustrates the production smoothing motivation when increasing marginal costs exist. If Q1 and Q 2 represent the demand in periods 1 and 2, respectively, then point A represents the average cost if Q1 is produced in period 1 and Q2 is produced in period 2. Point B represents the average cost if (Q1+Q2)/2 is produced in both periods, with the excess produced in period 1 carried over to period 2. The trade-off is between the cost of storage for one period versus the saving from smoothing.[1] The difference between A and B must be greater than the cost of holding inventory to justify smoothing. Note also that if mean demand is expected to decrease below current production for an extended period (that is, Q2 is current demand and Q1 is next period's expected demand), then it becomes optimal to reduce production and serve part of current demand from inventory. Thus, production smoothing motivation can lead to level changes if forecast sales change direction.

(S,s) Rule

If costs are linear, as in the case when marginal costs are constant, and there is a significant fixed cost of purchasing in each period, then it can be shown that "lumpy" adjustment is preferred to smoothing. An economic batch run, or a purchase which minimizes the total expected cost including the cost of storage of excess inventory, and the cost of lost sales can be determined. The inventory management technique used under these circumstances is referred to as (S,s) and entails determining maximum (S) and minimum (s) levels of inventory.[2] When inventories fall below (s), purchases are made to bring inventory up to (S), as long as inventories are between (S) and (s), nothing is done. The (S,s) parameters will define the upper and lower bound of inventory movement. It can be shown that the (S,s) margin is more sensitive to the mark-up of price over marginal cost than to interest rates.

[1] Holt, Modigliani, Muth and Simon (1960) provide the details of the production smoothing model.

[2] Scarf (1960) proves the optimality of the (S,s) rule under specific conditions.

wide range of preferences and sizes of consumers. Generally, inventories are a hedge against uncertainty or a means of minimizing production costs.

There are two competing models for inventory decisions, depending on the assumption about production costs. When firms operate in a region of increasing marginal costs, it becomes more economical to smooth production than to adjust to changing sales. When marginal costs are constant, but there are fixed costs associated with delivery or production, batch runs or bunching spread these fixed costs over larger quantities. (See the shaded insert above for discussion.) Wholesalers, retailers and manufacturing purchasers of raw materials and supplies are more likely to face non-negligible delivery costs and therefore more likely to use batch purchasing.

INVENTORY INNOVATIONS

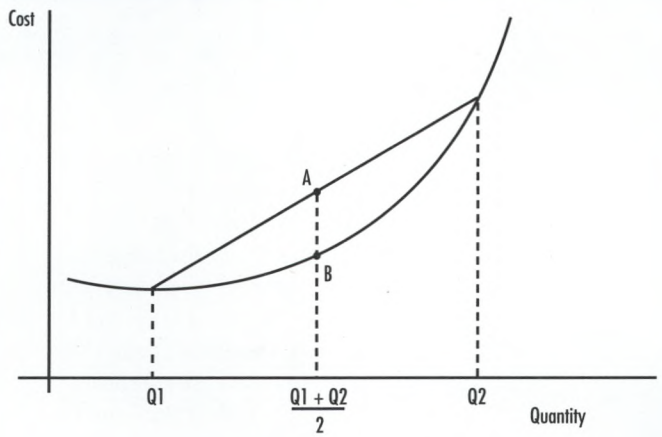
Is JIT Changing the Face of Inventory in America?

As businesses focused on streamlining operations in the 1980s, one of the targets has been inventory stocks. Over the last 15 years, there seems to have been major shifts in the methods used to manage inventory. In particular, many U.S. companies have studied and adopted the Japanese *kanban* (or JIT) method of inventory management. The objective of the JIT system is to minimize the stock of parts and components by having them delivered just in time for production, and to limit the inventory of finished goods by producing them just in time to fill demand. The monthly National Association of Purchasing Management survey indicates that as much as 26 percent

Just-in-Time Inventory

Just-in-time (JIT) inventory control attempts to match production as closely to sales as possible and thereby minimize the costs of holding inventory. This method, called *kanban* in Japan, is characteristic of Japanese industry in general and the auto industry in particular. JIT can be optimal when convex costs of production exist but storage costs exceed savings from smoothing or when linear costs exist but the low mark-up, low variance of sales, low fixed costs of delivery or high costs of storage result in low values of (S,s). If firms can meet demand without holding inventories, then inventories become superfluous. JIT can exist only in an atmosphere in which suppliers are reliable enough to minimize the risk of stock-outs. Larson (1991) argues that deregulation of the transportation industry has resulted in innovations which foster the use of JIT. Intuitively, deregulation, which reduces the economic lot-size of shipment, allows more continuous streams of shipment.

Convex Production Costs



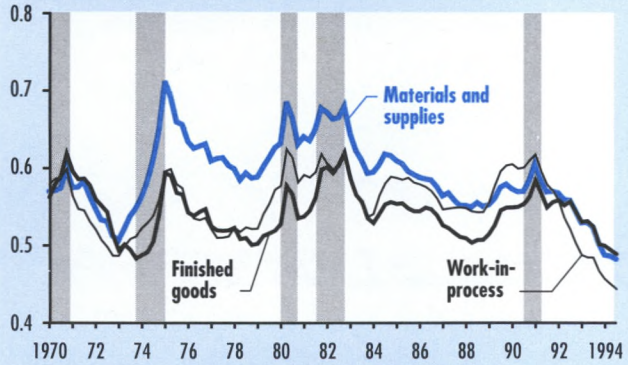
of the respondents reported purchasing materials "hand to mouth" in January 1995, compared to as little as 4 percent in February 1970. This suggests that the JIT philosophy has made major inroads into U.S. manufacturing. Bechter and Stanley (1992) find empirical evidence of improved inventory control along with faster speeds of adjustment to desired inventory levels.

Prima facie evidence of the success in reducing manufacturing inventory is also seen in the consistent decline in the aggregate inventory-to-sales ratio (shown in Figure 2), which has dropped from a peak of approximately 1.7 during the 1990 recession to 1.44 in December 1994 — the lowest in about 20 years. The manufacturing sector has been reducing inventory at all stages of production. Figure 3 shows the manufacturing sector inventory-to-sales ratios by stage of processing for 1970 to 1994. The work-in-process and materials and supplies are at a low point for the last two decades, after a steady decline since the early '80s. Some of this decline may be attributable to factors other than JIT. For instance, a closer look shows that materials and supplies increased rapidly relative to sales during the 1973-75 recession and did not return to earlier levels until recently. This could indicate an end to a post-oil-embargo tendency to stockpile, motivated by inflation expectations and sensitivity to interruptions.

Some industries have been more successful than others in lowering inventory levels relative to sales. Table 2 shows the summary statistics for the inventory-to-sales ratio by stage of processing for four manufacturing industries which have experienced significant declines in ratio. The December 1994 ratio is

Figure 3

Manufacturing Inventory-to-Sales Ratios by Stage of Processing



SOURCE: U.S. Department of Commerce, Bureau of Economic Analysis.

provided for comparison. Motor vehicle materials and supplies, and work-in-process inventory stocks have declined from peak ratios of over 70 percent of monthly sales each to 19 percent and 14 percent, respectively, in December 1994. In all four industries and for all stages of processing, December ratios are well below the mean for the entire period.

The use of JIT tends to shift the burden of responding to uncertainty to the suppliers and to require speedy delivery methods.[2] Some analysts believe that significant changes in the transportation industry, fostered in part by deregulation and increased competition, contributed to the viability of JIT. In particular, if a manufacturer wishes to maintain a continuous flow of materials, deliveries must take place more often in smaller batches. The deregulation in the trucking industry, which allowed competitive pricing for less-than-truckload deliveries, and increased competition in air freight help reduce the cost of smaller, more frequent deliveries.

JIT in the Auto Industry

The evolution of the structure of the U.S. automobile industry is relatively unique and was motivated primarily by the need to smooth production, combined with a limited ability to hold inventory (see Olney, 1989). The relationship between manufacturers, wholesalers and retailers ensures that the storage of finished goods occurs primarily at the retail and wholesale levels. The cost of financing high levels of inventory is a major cost of doing business. In the early years of the industry, finance companies took on the dual role of providing credit to wholesalers and buying consumer loans initiated by dealers. It seems appropriate, therefore, that the push to reduce inventory levels should take place in the auto industry. Minimizing inventory reduces financing needs and thus increases the competitive edge. The downside is greater vulnerability to interruptions such as strikes or to unanticipated surges in demand. U.S. automobile manufacturers appear to have embraced JIT and currently hold less than two weeks worth of sales in inventory, down from a high of 1.3 months.

Figure 4 shows the changes in inventory-to-sales ratios in the motor vehicle industry by stage of processing for the period 1970-94. It is apparent that there has been much success in reducing inventory levels over the last 10 years.

Figure 4 also reveals that the reduction occurred primarily at the work-in-process, and materials and supplies stages of production with very little change in the level of finished goods relative to sales. The burden of reduced inventory has been placed on the intermediate input stage of production.

As an example of the downside of lower inventory holdings, however, General Motors in 1994 experienced the shutdown of several assembly lines because of an interruption at a drivetrain component plant. If they had held higher levels of inventory, they would have been able to reduce the scale of the shutdown.

Bar Coding

The computer industry revolution and proliferation of bar coding has streamlined the inventory process in all sectors of the economy. Many retailers now use automatic scanning computer registers to record sales and track inventory immediately. These innovations have had the spillover effect of providing almost instant marketing information regarding the rate of sale or use of products. The increased use of bar code scanning and more sophisticated electronic systems

[2] The recent earthquake in Kobe, Japan, emphasized the potential disadvantage which this system produces, when many Japanese manufacturers, who were otherwise unaffected, had to shut down because of interruptions to suppliers and transportation.

Table 2

Inventory-to-Sales Ratio (January 1970 - December 1994)

Stage of Processing	Mean	Maximum	Minimum	Dec. 1994
Finished Goods				
Motor vehicles	0.147	0.261	0.099	0.115
Primary metals	0.635	1.032	0.372	0.545
Electrical	0.564	0.709	0.481	0.511
Non-Electrical	0.664	0.972	0.442	0.452
Work-in-Process				
Motor vehicles	0.298	0.730	0.137	0.139
Primary metals	0.862	1.338	0.564	0.677
Electrical	0.960	1.174	0.676	0.689
Non-Electrical	0.878	1.123	0.614	0.614
Materials & Supplies				
Motor vehicles	0.366	0.762	0.185	0.185
Primary metals	0.863	1.365	0.567	0.605
Electrical	0.679	0.893	0.541	0.554
Non-Electrical	0.647	0.886	0.437	0.514

over the last 10 years has led to more efficient retail (and wholesale) inventory management. This increased efficiency has not necessarily manifested itself as lower inventory levels, but allows more precise selection of stock items. In the retail sector, the inventory-to-sales ratio has actually increased slightly in contrast to the aggregate. The reasons for this increase are not obvious, but retailers must keep visible inventory on hand to stimulate sales, and therefore have less flexibility in inventory levels. In addition, an increase in the total number of stores[3] may have also contributed to the increase in aggregate retail inventory.

There have been efforts at limiting inventories at the retail level. "Quick Response" is the retail equivalent to JIT. Some retailers try to limit inventory by streamlining customer orders. The effectiveness of these efforts has been limited and so far appears to have had little impact on the level of aggregate retail inventory relative to sales.

Little (1992) uses quarterly manufacturing and trade data from 1968 through 1990 to test for structural changes in inventory management. Results of regressions of the data divided into two subperiods (1968:1 to 1982:3 and 1982:4 to 1990:4) support the notion that inventory management methods changed significantly beginning in the '80s. Similar work by Bechter and Stanley (1993) detects changes in the speed of adjustment and desired inventory-to-sales ratio after 1981 in a buffer-stock model.

The evidence supports the assertion that inventory innovations have impacted not only the quantity but also the quality of inventories held by allowing firms to more closely match patterns of use. Manufacturing has been more successful in reducing the quantity of inventory relative to sales, but the innovations in the wholesale and retail sectors should also limit the accumulation of unplanned inventory through more direct feedback of marketing information. As a result, the innovations in all three sectors should tend to limit the error portion of inventory accumulation.

IMPACT OF INNOVATIONS ON THE BUSINESS CYCLE

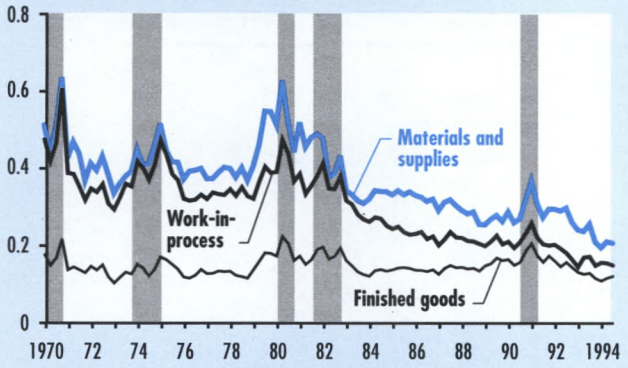
Inventory influences business cycle contractions primarily through unintended increases.[4] How do the structural shifts in inventory management affect unintended

[3] *The Economist* (March 4, 1995) reported in its retail survey that 1993 total shopping center space in the United States was 18.5 square feet per head, compared with 13.1 square feet per head in 1980, according to the Schroder Real Estate Associates.

[4] Some analysts suggest that higher-than-average growth during the recovery part of the cycle reflects planned inventory investment in anticipation of increased demand.

Figure 4

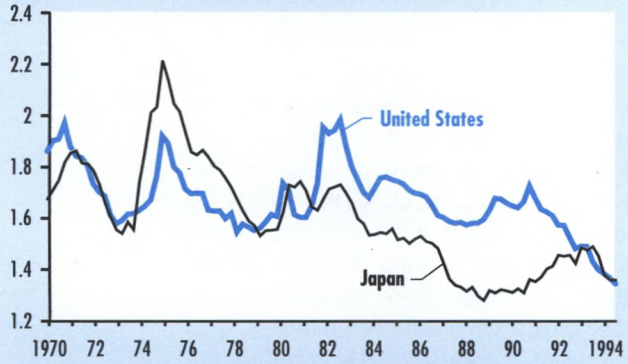
Motor Vehicles Inventory-to-Sales Ratio by Stage of Processing



SOURCE: U.S. Department of Commerce, Bureau of Economic Analysis.

Figure 5

Nominal Inventory-to-Sales Ratios for Japan and the United States



SOURCES: U.S. Department of Commerce (Bureau of Economic Analysis) and The Bank of Japan.

inventory build-up? Morgan (1991) suggests that a move to JIT produces a faster reaction to sales shocks and therefore will not result in the levels of unplanned accumulation previously observed. He also argues that as the use of JIT increases, the impact will be to lessen the inventory swings during recessions. Others have tried to directly assess the impact on the business cycle. Little (1992), for example, focuses on the transitory nature of the changes and suggests that the ongoing effort to reduce inventories were a drag on the recovery portion of the 1990-91 recession. The expected inventory accumulation after demand rebounded was offset by the continuing effort to reduce inventory-to-sales ratios. Bechter and Stanley (1993) use estimated parameters from their buffer-stock model to simulate inventory investment and conclude that the new parameters lead to larger inventory swings for a one-time shock in sales. Filardo (1995) uses an atheoretical vector autoregression (VAR) method and a model-based method to test empirically whether the changes in inventory management have muted the business cycle. He concludes there is no evidence of a reduced role for inventory in the business cycle. As Little (1992) suggests, however, the innovations are still being implemented and may not have saturated the market. In this case, there is an insufficient sample size to evaluate the business cycle impact empirically.

It is difficult to separate the effect of those firms using JIT from those which do not. One approach is to see if the industries that have converted to JIT now contribute less inventory investment during the recession. Primary metals, electrical machinery, non-electrical machinery and motor vehicles have shown significant decline in their inventory-to-sales ratios in the last 10-15 years. I looked at the 1980, 1982 and 1990 recessions to determine the contribution of these industries during the quarter with the biggest reduction in inventory. Together, the four industries contributed a net 22 percent to the third quarter 1980 change in business inventory, a net 29 percent to the fourth quarter 1982 change in business inventory, but only net 1.6 percent to the fourth quarter 1990 change in business inventory. The remaining manufacturing industries contributed 33 percent, 19 percent and 36 percent to the change in business inventories during these periods. These four industries that have reduced their inventory-to-sales ratios significantly over the past two decades contributed less to inventory swings in the 1990 recession than in 1980 or 1982. Despite the reduction in contribution by these industries, the change in business inventory contributed a higher proportion to the change in GDP during the 1990-91 downturn than in 1980 or 1981-82, but the magnitude of the decline in GDP was less in

1990-91 than in 1980 or 1981-82. On the surface, it appears that JIT may help reduce the magnitude of the inventory swing.

Another way to test the impact of JIT on business cycles is to compare the Japanese with the U.S. experience. First, we can look for evidence that Japan does maintain lower inventory levels. Figure 5 shows the inventory-to-sales ratio for the Japanese and U.S. manufacturing sectors. The Japanese ratio is lower than the United States during the 1980s, but both ratios have converged as the United States' decreased and Japan's increased somewhat.

Assuming that the lower inventory-to-sales ratio in Japan confirms the higher usage of JIT there, how does Japan's business cycle experience compare with the United States'? Unfortunately, an exact comparison is not possible because Japan has not recorded many periods of declining output. Using dates from Japan's Research Bureau Economic Planning Agency,[5] Japan's business cycles have had longer contractionary periods in the postwar era, averaging 16 months, compared with 11 months for the United States. Japan recorded 10 business cycles after World War II, compared to the United States' nine. The average duration of Japan's business cycles (50 months) and the expansion periods (33 months) were shorter than the United States' (63 months and 52 months, respectively).

Table 3 shows the changes in Japanese business inventory compared to changes in GDP during its last six contractions. Three of these six contractions had countercyclical inventory movement. Similar data for the United States (Table 1) shows unambiguous procyclical movement in business inventory. The data suggest that inventory changes may play a lesser role in GDP fluctuations in Japan than in the United States. How much of this is attributable to inventory management methods and how much is due to the difference in business cycle definition is uncertain.

Even if the use of JIT inventory management methods can dampen business cycles, this method is most applicable at the manufacturing level. The contribution of manufacturing, wholesale and retail inventories to total trade inventories has been changing over the last two-and-a-half decades. More recently, manufacturing's share has declined from 56.8 percent to 43.8 percent. Retail inventories have increased from a share of 24.3 percent to 31 percent. Wholesale inventories' share of the total has increased from 18.9 percent to 25.2 percent. The increased retail inventory-to-sales ratio and a greater retail share of the aggregate inventory may offset the gains in dampening the cycle from JIT at the manufacturing level.

Table 3

Changes in Japanese Inventory Investment During Business Cycle Troughs

Recession Period Peak to Trough[1]	Change in Real GDP[2]	Change in Inventory Investment[2]	Change in Inventory Investment as a Percentage of Change in Real GDP
1970:3 - 1970:4	–158.5	–527.0	332.5
1973:4 - 1974:1	–5296.8	7304.7	–137.9
1977:2 - 1977:3	1417.0	–499.5	–35.3
1980:1 - 1980:2	–453.0	4.3	–0.9
1985:4 - 1986:1	–3319.0	952.2	–28.7
1992:1 - 1993:4	–5097.0	–2807.4	55.1
Mean			30.8

[1] Peaks and trough correspond to peak and trough (or minimum growth) of real GDP during the contractions listed by the Research Bureau of the Economic Planning Agency of Japan, but do not always coincide with the peak and trough of the period.

[2] Billions of 1985 Yen (SAAR).

CONCLUSION

The data support anecdotal evidence that inventory management methods in the United States have changed significantly over the past decade or two. The result of these changes is evident in the reduced business inventory-to-sales ratio, driven almost entirely by lower inventories of work-in-process, and materials and supplies rather than finished goods. The impact

[5] The Japanese agency uses the Lucas (1977) definition, which loosely defines the business cycle in terms of deviation from trend growth. For most of the contractionary periods listed, Japan's GDP grew less than trend but did not experience a decline.

of these changes in inventory management techniques on business cycles is ambiguous. All other things being equal, inventory management innovations should reduce the probability of unintended accumulation. But as long as firms overestimate or underestimate future demand, inventory cycles will persist. And if cutbacks in production are required to reduce inventory, then the resulting reduction in income could result in lower demand and further inventory buildup. Inventory management innovations are not a panacea for taming business cycles, but in the long run these innovations can contribute to a faster response of production to changes in demand, which in turn can reduce the boom-bust cycle in the economy.

REFERENCES

Bechter, Dan M., and Stephen Stanley. "Economic Stability in the 1990s: The Implications of Improved Inventory Control," *Business Economics* (January 1993), pp. 35-8.

_____. "Evidence of Improved Inventory Control," Federal Reserve Bank of Richmond Economic Review (January/February 1992), pp. 3-12.

Blinder, Alan S. *Inventory Theory and Consumer Behavior.* University of Michigan Press, 1990.

_____ and Louis J. Maccini. "The Resurgence of Inventory Research: What Have We Learned?" *Journal of Economic Surveys* (No. 4, 1991), pp. 291-328.

Filardo, Andrew J. "Recent Evidence on the Muted Inventory Cycle," Federal Reserve Bank of Kansas City Economic Review (second quarter 1995), pp. 27-43.

Holt, Charles C., Franco Modigliani, John F. Muth and Herbert A. Simon. *Planning Production, Inventories, and Work Force.* Prentice-Hall, Inc., 1960.

Larson, Paul D. "Transportation Deregulation, JIT, and Inventory Levels," *The Logistics and Transportation Review* (June 1991), pp. 99-112.

Little, Jane Sneddon. "Changes in Inventory Management: Implications for the U.S. Recovery," Federal Reserve Bank of Boston *New England Economic Review* (November/December 1992), pp. 37-65.

Lucas, Robert E., Jr. "Understanding Business Cycles," in Karl Brunner and Allan H. Meltzer, eds., *Stabilization of the Domestic and International Economy.* Carnegie Rochester Conference Series, vol. 5. North-Holland, 1977.

Metzler, Lloyd A. "The Nature and Stability of Inventory Cycles," *Review of Economic Statistics* (February 1941), pp. 113-29.

Morgan, Donald P. "Will Just-In-Time Inventory Techniques Dampen Recessions?" Federal Reserve Bank of Kansas City *Economic Review* (March/April 1991), pp. 21-33.

Olney, Martha L. "Credit as a Production-Smoothing Device: The Case of Automobiles, 1913-1938," *The Journal of Economic History* (June 1989), pp. 377-91.

"Retailing" Survey, The Economist (March 4, 1995), pp. 3-18.

Scarf, Herbert E. "The Optimality of (S,s) Policies in the Dynamic Inventory Problem," in Kenneth. J. Arrow, Samuel Karlin and Patrick Suppes, eds., *Mathematical Methods in the Social Sciences,* 1959. Stanford University Press, 1960, pp. 196-202.

Alvin L. Marty is a professor of economics and finance at the Center for Business and Government, Baruch College, City University of New York. Daniel L. Thornton is an assistant vice president at the Federal Reserve Bank of St. Louis. Jonathan Ahlbrecht provided research assistance. The authors would like to thank, without implication, Phillip Cagan, David Laidler and Allan Meltzer for comments on an earlier draft.

Is There a Case for "Moderate" Inflation?

Alvin L. Marty and Daniel L. Thornton

The proposition that inflation is a monetary phenomenon is more widely embraced now than it was three decades ago. Moreover, it is more widely accepted that inflation is subject to long-run control by the central bank. In recent years, the central banks of the United Kingdom, New Zealand and Canada have placed increased emphasis on reducing their long-run inflation rates. In the United States, former Rep. Steven Neal, D.-N. Carolina (*House Joint Resolution 55*. January 5, 1993), and Sen. Connie Mack, R.-Florida, have proposed making stable prices the primary objective of the Federal Reserve.

Nevertheless, considerable opposition remains to making price stability the overriding objective of U.S. monetary policy. Some argue that the benefits of price stability do not warrant the cost of achieving it. For example, although extolling the virtues of price stability, Howitt (1990) is uncertain whether the benefits are worth the costs in terms of lost output (temporary, and perhaps permanent, due to hysteresis effects).

Although we are skeptical whether the empirical and theoretical analyses to date have correctly identified all of the benefits of price stability, this article addresses an issue that is logically prior to this one. Specifically, it addresses the question: If the inflation rate were zero, could society benefit from a higher rate of inflation? In other words, is moderate inflation preferable to price stability? Several arguments have been advanced that the economy benefits from moderate inflation. Recently, DeLong and Summers (1992) and Summers (1991) have suggested several rationales for why a central bank would choose moderate inflation over price stability as its long-run policy goal.

This article addresses four reasons that have been suggested to prefer moderate to zero inflation:

1. Moderate inflation enhances the stability of the economy.
2. Moderate inflation results in a higher steady-state level of output per person.
3. Moderate inflation increases the efficiency of inter-industry labor market adjustments.
4. Inflation enhances the efficacy of countercyclical monetary policy by allowing the real rate of interest to be negative, thereby stimulating effective demand in periods of recession.

The first two of these arguments are well-known to economists, but have received scant attention in public debates. Moreover, they are framed within specific, although quite different, theoretical models, so it is possible to provide a rather definitive evaluation of their merit. The remaining two arguments have received considerable attention, and may play a role in any public policy debate regarding the desirability of making price stability the primary monetary policy objective. The conceptual frameworks for these arguments are not well-specified, however, so we try to shore up their analytical footing by proposing specific interpretations.

Before proceeding, several issues should be clarified. First, the hypotheses that there are economic benefits from moderate inflation considered here implicitly argue against a so-called Friedman Rule (Friedman, 1969), that is, the "optimal" rate of money growth is one that generates steady-state deflation. Nevertheless, this article is not specifically about the Friedman Rule. Analyses of a Friedman Rule generally have been carried out in well-specified model economies.

Second, although the last two arguments for moderate inflation lack explicit theoreti-

cal foundations, this has not prevented them from achieving an intellectual status among some economists and policymakers. The lack of theoretical foundations forces our analysis to range from the fairly technical to the somewhat conjectural, so that we may not provide a definitive evaluation of these arguments. In this case, we are content to present an analysis of these arguments.

Third, the arguments for moderate inflation analyzed here are based on the assumption of a fully anticipated, steady-state inflation. Although such inflations do not characterize real-world economies, we make this assumption until it is relaxed when we discuss the reasons why price stability is preferable to moderate inflation.

Fourth, the phrase "moderate inflation" is not well-defined. Some might consider moderate inflation to be 2 to 3 percent. For others, any rate under 5 percent could be moderate. Still others may deem anything less than double-digit inflation moderate. We suggest 5 percent as the break point for moderate inflation in the United States.

Finally, although we used the phrases price stability and zero inflation interchangeably, we are aware that price stability is different and more stringent than zero inflation. Price stability implies that jumps in the price level are reversed; zero inflation need not.

THE CASE FOR MODERATE INFLATION

Is Stability of the Economy Enhanced by Moderate Inflation?

The first argument for moderate inflation is that certain stability conditions are sturdier at a high-money (nominal) rate of interest, making the economy less vulnerable to various shocks. Understanding this argument requires an understanding of the notion of stability upon which it rests. To illustrate, suppose that the real rate of interest suddenly rises, say, because of an increase in expected future profits. Given the underlying rate of inflation, this raises the money rate of interest, reducing the quantity of real money balances that individuals desire to hold. As individuals attempt to reduce their holdings of real balances, prices rise still further, fueling expectations of further inflation. This reduces the quantity of real balances demanded still further, giving rise to a further increase in prices and so on. The question is: Under what conditions will this sequence converge?

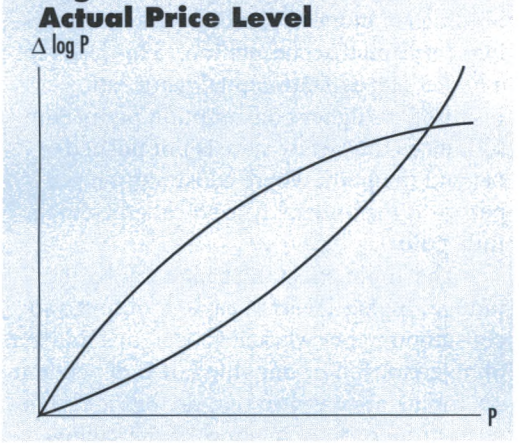
The answer is: Self-generating inflation cannot occur if, as the price level rises, its rate of change declines. Holding the growth of the money supply constant, this condition is illustrated in Figure 1, which shows two plots of the change in the log of the price level, against the price level itself, P. In one case, the slope of the curve rises with P. In this instance, the sequence will not converge and the ultimate solution is the trivial one; the demand for real money balances approaches zero. In the other case, the slope of the curve decreases as P increases, so that the sequence converges to a steady rate of inflation.

The argument that inflation enhances economic stability is an argument about the demand for money. To see this argument, assume that the demand for real money balances is solely a function of the nominal interest rate, whereby the nominal interest rate equals the constant real interest rate plus the actual rate of inflation (which is fully anticipated). Cagan (1956) showed that the rate of inflation decreases as the price level rises if $\alpha\beta < 1$. The parameter α is the semi-log slope of the demand for real money balances with respect to the nominal interest rate, that is, percent change in the demand for real money balances per percentage point change in the money interest rate ($d \ln(M/P)/di$). The parameter β is the rate at which individuals revise their expectations of inflation under adaptive expectations, $dE/dt = \beta(\pi - E)$, where π and E are the actual and expected rates of inflation, respectively.

If $\alpha\beta < 1$ and the expected rate of inflation is initially greater than the actual rate, expected inflation falls until a stable steady state is reached at which the expected rate of inflation is equal to the actual rate. If, however, $\alpha\beta$ is greater than unity and actual inflation is initially greater than expected inflation, both expected and actual inflation grow without limit with real balances falling.

Figure 1

Log of the Price Level Vs. Actual Price Level



On the other hand, if $\alpha\beta > 1$ and actual is initially less than expected inflation, both actual and expected inflation fall without limit. The steady state at which the actual and expected rate are equal is unstable.

Stability Through Moderate Inflation

If α and β are constants, the stability conditions will be invariant to the steady-state inflation rate. Consequently, the suggestion that the stability conditions are sturdier at non-zero rates of inflation is an argument that either α or β is inversely related to the steady-state inflation rate. Specifically, it was argued that α should be smaller at higher rates of money interest. That this may be so comes from noting that the elasticity of the demand for real money balances with respect to the money interest rates, e_m, is equal to αi, where i is the money interest rate. Thus, if e_m is constant, α will decline as the rate of inflation and, hence, the money interest rate rises.[1] All other things the same, the stability condition is more likely to be satisfied at higher rather than lower rates of money interest if the interest elasticity demand for real money balances is constant.

The widely used Cagan (1956) money-demand function assumes that α is independent of the nominal interest rate. Cagan's function significantly underpredicted real money balances during periods when prices were or had been relatively constant, but performed well during periods of high inflation.[2] This suggests that the underprediction was due to α being higher at low rates of inflation.

That α is inversely related to the steady-state inflation rate is plausible, but this does not imply that stability is more likely at higher rates of inflation. For example, it is plausible that individuals revise their expectations of inflation more rapidly at higher rates of inflation, that is, that β is positively related to the inflation rate. Indeed, Bruno (1989) provides some empirical support for a positive association between β and the rate of inflation. Consequently, it is not necessarily the case that stability is greater at high inflation rates. An inflation-induced fall in α might be just offset, or perhaps more than offset, by an inflation-induced rise in β. Using adaptive expectations, no general conclusion can be reached about the stability conditions and the steady-state inflation rate.

Getting on the Wrong Side of the Laffer Curve

Bruno and Fischer (1990) have revisited Cagan's stability condition in the context of financing a given budget deficit solely through seigniorage from money creation. Although assuming that the deficit is financed solely through money creation is not realistic in developed economies like the United States, where other forms of taxation are available, the Bruno and Fischer assumption is a useful theoretical device which allows stability conditions to be anchored by two equilibria on either side of the Laffer curve. The Laffer curve is the recognition that tax receipts do not increase continuously with the tax rate. Beyond some point, receipts decline as a further increase in the tax rate results in a significant erosion of the tax base. Consequently, except at that unique tax rate where tax revenue is maximized, there are two alternative tax rates and tax bases that generate the same tax revenue: a low tax rate and a high tax base, or a high tax rate and a low tax base. Bruno and Fischer demonstrate that if $\alpha\beta < 1$, a stable equilibrium is at the socially desirable low tax rate-high tax base point. If $\alpha\beta > 1$, an equilibrium is at the socially undesirable

[1] Note that e_m is constant if the money demand function is in double log form: $M/P=i^k$, so $e_m=\alpha i=k$.

[2] See Bailey (1956) and Friedman and Schwartz (1963) for a discussion of the issue of the empirical validity of Cagan's money-demand equation.

point on the Laffer curve. Consequently, the argument is not whether the system is stable or explosive, but whether equilibrium is achieved at a socially desirable point on the Laffer curve.

What If Expectations Are Rational?

From the condition that $E = \pi$, it is easy to see that rational expectations are the limiting case of adaptive expectations. Adaptive expectations approach rational expectations as $\beta \rightarrow \infty$. Consequently, if expectations are rational, the condition $\alpha\beta < 1$ cannot be satisfied for any value of α. In Cagan's world, the system explodes. In the Bruno and Fischer world of a fixed real deficit, a stable equilibrium (if it exists) is achieved at a high inflation rate on the wrong side of the Laffer curve. Under rational expectations, any affect on α is completely overwhelmed by β, which is infinite.

It appears that nothing remains of the argument for stability through inflation. In the case of adaptive expectations, any possible reduction in α due to inflation may be offset by an increase in β. If expectations are rational, an infinite β swamps any effect of inflation on α. Indeed, the stable equilibrium is at the socially undesirable side of the Laffer curve, that is, at a high rate of inflation (tax rate) and a low level of real cash balances (tax base). In particular, no argument can be made that moderate inflation produces stability on the socially desirable side of the Laffer curve.

Does Moderate Inflation Lead to a Higher Level of Output?

The second argument for moderate inflation, that it leads to a higher level of steady-state output and consumption, was first formulated by Tobin (1965). The essence of Tobin's model is that in a growing economy, non-interest bearing real money balances augment disposable income. Given that the propensity to save out of disposable income is less than unity, an increase in real balances, all of which must be saved, gives rise to smaller saving in the form of physical capital. In Tobin's portfolio-balance model, real money balances and physical capital are substitutes. A higher anticipated rate of inflation induces individuals to economize on their holdings of money balances, freeing up savings for capital accumulation. This leads to a higher steady-state capital/labor ratio, resulting in higher consumption per person so long as the steady state is not pushed beyond the point where consumption per person is maximized (the so-called Golden Rule point).

The Tobin effect, that higher inflation induces higher levels of capital, output and consumption per worker, is open to a number of objections. For one thing, it is dependent on Tobin's assumption that savings are a constant proportion of income. If the savings rate is directly reduced by higher inflation, the Tobin effect can be reversed—even in the framework of Tobin's model (Dornbusch and Frenkel, 1973). Moreover, the Tobin effect is model-specific. The effect is absent in Ramsey-type models, in which the marginal product of capital is tied to the representative agent's rate of time preference. In such models, the marginal product of capital defines a unique steady-state capital\labor ratio which is independent of the level of real money balances. It is now generally recognized that the results of both Tobin- and Ramsey-type models are sensitive to small changes in assumptions. Moreover, Orphanides and Solow (1990) show that different models or small changes in assumptions in a particular model deliver disparate conclusions about the Tobin effect. Consequently, it is impossible as a matter of pure theory to make a compelling case that inflation increases real output.

A crucial reason for the fragility of these results is that, by their very nature, these money-growth models are the wrong vehicles for analyzing the role of money in the economy and, hence, the effect of inflation on the economy. A striking example of this is provided by Tobin's model, which predicts that the highest level of output per person occurs in a barter economy, in which holdings of real money balances are nil. This result stems from not taking money seriously. Real money balances reduce transaction costs. They do this by overcoming the dou-

ble coincidence of wants associated with barter and by conveying information (for example, Brunner and Meltzer, 1971). Compared to a barter economy, the reduction in total transaction costs permits society to devote more of its scarce resources to production, raising output and the consumption of goods and leisure. By reducing marginal transaction costs, money also results in a higher level of trade and correspondingly higher levels of output. Although the development of Solow-type growth models was an important first step in the analysis of growing economies, it is not surprising that these one-commodity models fail to capture the important role that money plays in real-world economies.

Thus far, we have contrasted money and non-monetary barter economies. In principle, similar effects occur when individuals are induced by a rise in anticipated inflation to reduce their holdings of real balances. If inflation induces individuals to hold fewer real balances, even if one were to accept Tobin's argument that inflation increases output per person, any increase would be at the expense of a loss to society of the services of real balances. In fact, if money enters the production function, the Tobin effect may well be reversed; inflation then reduces output per person, as in Stockman (1981).

Inflation may not only reduce the steady-state level of per capita output, it may reduce the growth rate of output itself. For if capital is appropriately defined more broadly to include human capital, as is done in recent endogenous-growth models, inflation reduces investment in human capital, as well as in physical capital. Reduced investment in human capital lowers the growth of efficiency per person, which reduces the growth rate itself (for example, Lucas, 1988; King and Rebelo, 1990; and Dotsey and Ireland, 1993). A small but permanent reduction in the growth rate due to inflation has an adverse effect on output levels. This continual effect on output levels is more significant than any effect of inflation on the one-time altering of the level of output per capita explored in earlier exogenous-growth models.

Once we leave purely theoretical-growth models, and look at real-world economies, we must deal with how inflation interacts with real-world institutions. It has been shown that the interaction of inflation with a less-than-fully indexed tax system works to discourage capital accumulation (Feldstein, 1976, 1979; and Tatom, 1976).

The bottom line is that even within the framework of theoretical-growth models, the Tobin effect is subject to small changes in assumptions. When real-world institutions are included in the analysis, the weight of evidence is that inflation discourages capital accumulation. The Tobin effect is reversed.[3] When capital is defined more realistically to include human capital, the effect of inflation is to continually reduce the levels of output per person below what they would have been under stable prices.

Does Moderate Inflation Enhance Relative Real-Wage Adjustments?

The third argument for moderate inflation (Tobin, 1972; Schultze, 1985; Lucas, 1989; DeLong and Summers, 1992; and Summers, 1991) asserts that declines in the price of commodities and in the real wage of workers specialized to a particular industry can be made with less friction in a world with moderate inflation than in a world of stable prices. It is argued that under moderate inflation, the decline in a product's price and in the real wage rate of workers can be accomplished through a rise in prices and money wages elsewhere.

The belief that inter-industry adjustments are smoother under a regime of moderate inflation rests on the view that laborers prefer a rise in the prices of wage goods to an absolute reduction in money wages. But why should this be the case? Workers experience an identical decline in real wages in both cases.

One answer depends on the existence of a money illusion: A decline in real wages brought about by a rise in the prices of wage goods is incorrectly perceived as smaller than the same decline in the real wage that occurs through a reduction in money wages. The persistence of money illusion in a steady state of anticipated moderate inflation is difficult to rationalize. Moreover, recent evi-

[3] After a survey of theoretical models, Blanchard and Fischer (1989) conclude: "Calculations suggest, however, that the effects of changes in the inflation rate on capital accumulation in models of the type developed in this chapter are very small. If inflation has systematic effects on capital accumulation (and there is empirically a negative association), it is probably for reasons not included so far. One likely reason is that the tax system is not neutral with respect to inflation."

dence (McLaughlin, 1994; and Lebow, Stockton and Wascher, 1993) suggests there is no dearth of nominal wage cuts, even during periods of moderate inflation.

Furthermore, firms in a declining industry may adjust workers' compensation without cutting wages. Compensation includes benefits and perks which can be adjusted relatively easily relative to wages in cases in which workers have an irrational fear of nominal wage cuts. In any event, we believe that for the resistance to nominal wage cuts to be widespread, it must be motivated by considerations deeper than a pure money illusion.

One possible rationale for such resistance is that workers feel they have some control over money wages but no control over the general price level. Consequently, the same reduction in the real wage rate due to reduction in money wages brings into play factors that workers believe they can negotiate, in contrast to an increase in the prices of wage goods, which they are powerless to affect.

The second possible motivation would interact with the first. Workers may have less knowledge of demand than do employers. Consequently, when the industry demand declines, workers may be concerned that the employer is misrepresenting the true state of nature to force an unnecessary reduction in money wages. In this case, a fall in the real wage rate due to a rise in the prices of wage goods elsewhere avoids triggering a signal-extraction problem.

We have endeavored to make the best possible case for moderate inflation as a device for smoothing inter-industry wage adjustments, but in doing so, we have ignored the existence of a cushion on money wage declines even in a regime of stable prices. If we were to introduce technical progress, even under price stability, the average level of money wages would rise at a rate equal to the average increase in output per person. This provides a cushion mitigating the need for an absolute decline in the money wage.

Finally, we suggest the hypothesis that workers' resistance to nominal wage cuts is not independent of the inflation regime in which they live. Under stable prices, such cuts may become more frequent and workers will become more accustomed to and less distrustful of money wage cuts. Accordingly, any "lubricant" that moderate inflation may provide to ease labor market frictions will become increasingly unnecessary in a zero-inflation regime.

Although the claim that moderate inflation facilitates inter-industry wage adjustments cannot be definitively rejected, it does not rest on compelling theoretical or empirical foundations. In any event, monetary policy is an inappropriate and ineffective instrument for dealing with labor market problems, such as market frictions or the sub-optimality of the natural unemployment rate. The latter may be due to taxes on wages which make the after-tax real wage smaller than the before-tax marginal product of labor. The socially optimal amount of employment equates the disutility of labor to the before-tax real wage so that after-tax employment is sub-optimal. Moreover, high unemployment compensation increases time spent in "search unemployment." The drift to higher unemployment in Europe and Canada is unlikely to reflect a movement up a short-run Phillips curve produced by unanticipated deflation, but rather is due to an upward drift in the natural rate of unemployment. In this case, appropriate policies to reduce unemployment are reforms in taxes and unemployment compensation, not monetary policy.

Moderate Inflation Enhances the Countercyclical Efficacy of Monetary Policy

A fourth argument for moderate steady-state inflation is that it enhances the counter cyclical efficacy of monetary policy by enabling the Federal Reserve to make the real rate of interest negative. The argument that the efficacy of monetary policy is enhanced by a moderate rate of steady-state inflation stems from the recognition that the money rate can never be negative, so that in a non-inflationary environment, in which the real and money rates are equal, the best that monetary policy can do is to drive both the

real and nominal rates of interest to zero.[4] DeLong and Summers (1992) and Summers (1991) argue that in a zero inflation regime, monetary policy will be unable to produce a sufficiently large reduction in the real interest rate to restore full employment in times of large adverse shocks to aggregate demand.

Is this an important argument for moderate inflation? There are several reasons to think not. First, the argument is based on the belief that the monetary authority can exert considerable influence over real interest rates through the so-called liquidity effect, and that monetary policy works primarily, if not solely, through its ability to influence the real interest rate. According to this view, an expansionary monetary policy drives real interest rates down, inducing an increase in spending. But the extent and duration of the effect of monetary policy on short-term real interest rates is controversial, theoretically and empirically. The exchange between Ohanian and Stockman (1995) and Hoover (1995) highlights the difficulties with theoretical models of the liquidity effect. The empirical evidence on the liquidity effect is mixed. Work by Reichenstein (1987), Thornton (1988), Gordon and Leeper (1994) and Pagan and Robertson (1995) suggests the liquidity effect is relatively weak and short-lived, although research by Christiano and Eichenbaum (1991, 1992), Cook and Hahn (1989) and Romer and Romer (1990) suggest a more significant effect of monetary policy on real short-term interest rates.

Second, it is difficult to argue that sufficient investment opportunities will not exist unless the real rate is negative. The issue is whether the economic outlook can become sufficiently pessimistic that the expected real return on longer-term investments is negative. That DeLong and Summers (1992) and Summers (1991) have raised it again suggests that this old debate is far from settled. Bailey (1971) argues that there will always be some investments that yield a small nonnegative return, even if a depressed economy were not expected to return to its steady-state growth path for a period of 10 to 20 years. If investment opportunities increase sufficiently at very low but positive interest rates, the efficacy of monetary policy is not impaired by a zero lower bound on the real interest rate. Bailey's argument suggests that credit demand becomes very large (essentially infinite) at very low real interest rates, so that the real longer-term interest rates do not have to be negative to significantly increase investment.

Finally, despite the empirical evidence to the contrary, there appears to be a fairly widespread belief that the Federal Reserve exerts considerable influence over real short-term interest rates, but much less influence over longer-term interest rates (see, for example, Goodfriend, 1993; and Greenspan, 1993). If monetary policy cannot make the long-term rate negative, it is natural to ask: Is there any gain from the possibility that the Federal Reserve may be able to make short-term interest rates negative for temporary periods? In markets in which there are few impediments to the flow of funds between the long and short end of the market, consistency of expectations requires that the current long-term interest rate be equal to the expected average of future short-term rates plus a risk premium. The risk premium is affected by a number of things, including uncertainty about future short-term interest rates.

If the market believes that the policy does not signal an increase in policymakers' desired steady-state inflation rates, people know that today's policy must give rise to reversals later. Whether the difference in the magnitude of the decline and subsequent rise in short-term interest rates in the zero and moderate inflation regimes will result in significantly different paths for real long-term interest rates under the two regimes is impossible to determine, a priori. Indeed, it is as easy to conjecture scenarios in which there would be no difference in the response of long-term real interest rates under the two steady-state inflation regimes as it is to conjecture scenarios in which there would be a significant difference.[5]

Given that it is unlikely that moderate inflation will enable the Federal Reserve to have a significantly larger effect on long-term real interest rates, and that very low or zero

[4] No one would willingly trade a dollar for, say, 95 cents a year from now, so long as the same dollar could be held for a year at zero carrying cost.

[5] Indeed, Fuhrer and Madigan (1993) simulate the effect of more aggressive policies that result in negative short-term interest rates and find very small changes in long-term rates.

real interest rates are likely to be sufficient for the Fed to offset adverse aggregate demand shocks, the argument that moderate inflation enhances the efficacy of monetary policy seems doubtful. If some role for inflation uncertainty is factored in, the idea that moderate inflation enhances the efficacy of monetary policy becomes even more tenuous.

Why Zero Inflation Is Preferable

Although many estimate the output losses of moderate inflation to be modest, this issue is far from settled. In addition to the usual problems of measuring the permanent output losses, Dotsey and Ireland (1993) have shown that in a general-equilibrium analysis, the usual effects of inflation (the inefficient economizing on real money balances, substituting market activity for leisure, and redirecting resources from goods production to financial activities) compound to produce a significant output loss. Dotsey and Ireland's result stems in part from the fact that inflation lowers real output growth. Although the effects on output growth appear small, compounded over time, they are significant.

Another compelling reason to prefer zero inflation is that higher inflation is associated with increased variability of both inflation and relative prices. The increased variability of inflation and consequent inflation uncertainty shorten contract lengths, thereby increasing contract costs. The greater variability also contaminates price signals, so the price system conveys less information. As the variability of inflation (associated with higher inflation) increases, it becomes more difficult to determine whether a particular commodity price change reflects a movement in the general price level, or a real shift in supply or demand resulting from taste and productivity shocks.

In addition, inflation and the higher variability of the general price level cause a reallocation of resources from the production of goods to financial services for the sole purpose of hedging against inflation. Even if there is no reduction in conventional measures of output, inflation produces a distortion of output. The banking system and financial service industries expand relative to other employment of resources such as industrial output, and households sacrifice leisure to reduce their real balances when the inflation tax rises. These effects call into question the notion that, by penalizing the consumption of priced commodities, inflation reduces work effort and increases leisure.

Although it is difficult to quantify the degree to which inflation impairs the ability of the price system to signal correct information, there is no doubt that the price system allocates resources most efficiently in the absence of inflation.

Moreover, zero inflation is preferable to moderate inflation because inflation, even moderate inflation, distorts accounting, legal contracts and the tax system. Inflation also distorts the true cost of inventories, the depreciation of plant and equipment, as well as the time profile of real mortgage payments, and other fixed-dollar denominated contracts.

Of course, this analysis assumes that taxes and private contracts are not indexed against inflation. Why, it may be asked, don't the authorities index taxes against changes in the price level so that real payments are unaffected? Why, in turn, doesn't the private sector index wages and financial contracts to nullify the impact of price changes?

In fact, the tax code is now partly indexed against inflation. Indexation, however, is often taken as a signal that the authorities are giving up the battle against inflation. This was the basis for the outspoken opposition to indexation by former Federal Reserve Chairman Arthur Burns and why the Bank of Canada has opposed indexation. It is easy to construct examples in which inflation-mitigating schemes, such as indexation by reducing the marginal costs of inflation, lead to an increase in the aggregate inflation rate. Moreover, foregoing indexation may be a help in developing a reputation for credibly pursuing anti-inflation policies (see, for example, Fischer and Summers, 1989). For these reasons, it is not clear that indexation of tax codes is desirable.

In the private sector, indexation is unlikely to occur. At the heart of the diffi-

culty is a coordination problem. To be successful, indexation must be implemented by a large number of diverse firms facing different information. For example, are price changes due to nominal or real variables? It is well-known that indexing money wages to changes in prices due to real shocks is undesirable.[6] Are price changes permanent or transitory? Are the price changes industry-specific or global? It is unlikely that individuals will agree on the cause of a price change and then coordinate their actions. Given these obstacles, indexation in the private sector is difficult and, hence, fairly rare. Moreover, in instances in which private indexation is fairly widespread, as in Israel, it reduces the resolve to fight escalating inflation.

Finally, it is impossible to fully index real cash balances against inflation, as previously discussed, because inflation leads people to hold fewer cash balances and results in a loss of their services.[7] For these reasons, indexation is a frail reed on which to rest hopes of mitigating inflation's effects.

Price Stability as the Objective of Monetary Policy

Reducing an established moderate inflation trend may disrupt economic activity, producing temporary output and employment losses. Given an established moderate inflation rate, Howitt (1990, p. 104) argues that despite the desirability of zero inflation, the cost of achieving it probably outweighs the benefits. This argument against moving to price stability ignores the inflationary bias (and resulting uncertainty) that characterizes policy regimes motivated by concerns for transitional output and employment losses.

In the absence of a commitment to stable prices, a central bank concerned about transitional unemployment is likely to respond asymmetrically to shocks—temporary or permanent. This asymmetric behavior has clear implications for the price level in the case of demand shocks. A monetary authority concerned with transitional unemployment will be less willing to offset a demand shock that raises prices and employment than to offset an adverse demand shock that lowers prices and employment.

This asymmetric behavior extends to supply shocks as well.[8] Adverse shocks will be accommodated; favorable ones will be ignored. Although the price level depends on many factors, including the relative incidence of positive and negative shocks, concern for transitional unemployment leads a central bank to pursue policies that will cause the price level to be higher than it would be otherwise.

This asymmetric behavior creates an inflationary bias with the potential for accelerating inflation. As inflation increases, the monetary authority may be forced to tolerate transitional unemployment to bring the inflation down. Indeed, this appears to be what happened in the United States in the late 1970s and on a smaller scale in the late 1980s. The best way to avoid such disruptions is to commit to a policy of stable prices.

SUMMARY AND CONCLUSIONS

We have reviewed several arguments in favor of moderate inflation and we find them to be lacking theoretically and, in some instances, empirically. The first argument, that moderate inflation enhances economic stability, is subject to compelling objections. If expectations are adaptive, any decline in the semi-elasticity of money demand associated with a higher inflation may well be offset by a more rapid revision of inflationary expectations. If expectations are rational, this must be the case.

The argument that inflation leads to a higher level of output is based on theoretical models that are not robust to small specification changes. When real-world institutions are taken into account, the weight of the evidence is that inflation discourages capital accumulation. When capital is defined to include human capital, inflation may reduce not only the level of output per capita but its rate of growth as well.

Also suspect is the proposition that moderate inflation increases the efficacy of monetary policy by allowing the central bank to make the real rate of interest negative. Sufficient investment opportunities are likely to exist at very low but positive real

[6] The authorities in Israel indexed money wages to a price index which included imported goods. In fact, imported goods should be excluded since changes in import prices reflect changes in a real variable, the terms of trade. Later on, this mistake was rectified and terms of trade effects were excluded from the price index.

[7] In principle, to maximize the services of real balances, it would be desirable to have prices fall at the real rate of interest and set the money rate of interest to zero. People would then be induced to hold the satiety quantity of real balances. It would take us too far afield, however, to discuss the merits of deflation at the real rate in comparison with stable prices. Therefore, we confine our attention to a comparison of zero with moderate inflation.

[8] In the case of the oil shock in the 1970s (an adverse supply shock which tended to raise inflation and reduce output), a number of economists advocated a quantum increase in the stock of money to offset a potential increase in unemployment. On the other hand, how many voices were raised in favor of a reduction in the money stock when OPEC collapsed?

interest rates. Consequently, negative real rates are not required to make monetary policy effective. Also, even if positive inflation enabled the Fed to make real short-term interest rates negative, such actions may not lower long-term interest rates.

The proposition that moderate inflation eases inter-industry wage adjustments is weak too. One argument rests on the existence of a money illusion; we see no economic rationale for money illusion in the steady state. If the asserted resistance to nominal wage cuts is based on a deeper motivation, we suggest that it should disappear entirely as the regime of zero inflation persists. Moreover, the evidence suggests that nominal wage cuts are frequent even during periods of moderate inflation. Hence, the conjecture that workers resist nominal wage cuts lacks both theoretical and empirical justification.

Finally, we argue that a policy of living with inflation cannot be rationalized on the grounds that there are transitory output costs associated with reducing inflation. A policy motivated by concern for transitional unemployment is likely to have inflationary bias that will erode a commitment to any price objective.

REFERENCES

Bailey, Martin J. *National Income and the Price Level: A Study in Macroeconomic Theory.* McGraw Hill, 1971.

_____. "The Welfare Cost of Inflationary Finance," *Journal of Political Economy* (April 1956), pp. 93-110.

Blanchard, Olivier J., and Stanley Fischer. *Lectures on Macroeconomics.* MIT Press, 1989.

Brunner, Karl, and Allan H. Meltzer. "The Uses of Money: Money in the Theory of an Exchange Economy," *The American Economic Review* (December 1971), pp. 784-805.

Bruno, Michael. "Econometrics and the Design of Economic Reform," *Econometrica* (March 1989), pp. 275-306.

_____ and Stanley Fischer. "Seigniorage, Operating Rules, and the High Inflation Trap," *Quarterly Journal of Economics* (May 1990), pp. 353-74.

Cagan, Philip. "The Monetary Dynamics of Hyper Inflation," in Milton Friedman, ed., *Studies in the Quantity Theory of Money.* University of Chicago Press, 1956.

Christiano, Lawrence J., and Martin Eichenbaum. "Liquidity Effects and the Monetary Transmission Mechanism," *The American Economic Review* (May 1992), pp. 346-53.

_____ and _____. "Identification and the Liquidity Effect of a Monetary Policy Shock," National Bureau of Economic Research *Working Paper No. 3920* (1991).

Cook, Timothy, and Thomas Hahn. "The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s," *Journal of Monetary Economics* (November 1989), pp. 331-51.

DeLong, J. Bradford, and Lawrence H. Summers. "Macroeconomic Policy and Long-Run Growth," Federal Reserve Bank of Kansas City *Economic Review* (fourth quarter 1992), pp. 5-29.

Dornbusch, Rudiger, and Jacob A. Frenkel. "Inflation and Growth: Alternative Approaches," *Journal of Money, Credit and Banking* (February 1973), pp. 141-56.

Dotsey, Michael, and Peter Ireland. "On the Welfare Cost of Inflation in General Equilibrium," unpublished manuscript (August 1993), Federal Reserve Bank of Richmond.

Feldstein, Martin S. "The Welfare Cost of Permanent Inflation and Optimal Short-Run Economic Policy," *Journal of Political Economy* (August 1979), pp. 749-68.

_____. "Inflation, Income Taxes, and the Rate of Interest: A Theoretical Analysis," *The American Economic Review* (December 1976), pp. 809-20.

Fischer, Stanley, and Lawrence H. Summers. "Should Nations Learn to Live with Inflation?" *The American Economic Review* (May 1989), pp. 382-87.

Friedman, Milton. "The Optimum Quantity of Money," in *The Optimum Quantity of Money and Other Essays.* Aldine Publishing Co., 1969.

_____ and Anna Jacobson Schwartz. *A Monetary History of the United States 1867-1960.* Princeton University Press, 1963.

Fuhrer, Jeff, and Brian Madigan. "A Monetary Policy When Interest Rates are Bounded at Zero," unpublished manuscript (August 25, 1993), Board of Governors of the Federal Reserve System.

Goodfriend, Marvin. "Interest Rate Policy and the Inflation Scare Problem: 1979-1992," Federal Reserve Bank of Richmond *Economic Quarterly* (winter 1993), pp. 1-24.

Gordon, David B., and Eric M. Leeper. "The Dynamic Impacts of Monetary Policy: An Exercise in Tentative Identification," *Journal of Political Economy* (December 1994), pp. 1228-47.

Greenspan, Alan. "Statement before the Subcommittee on Economic Growth and Credit Formation of the Committee on Banking, Finance and Urban Affairs of the United States House of Representatives," *Federal Reserve Bulletin* (September 1993), pp. 849-55.

Hoover, Kevin D. "Commentary," this *Review* (May/June 1995), pp. 26-32.

House Joint Resolution 55. "Directing the Federal Reserve Open Market Committee of the Federal Reserve System to Adopt and Pursue Monetary Policies Leading To, and Then Maintaining Zero Inflation," 103rd Congress, 1st Session (January 5, 1993).

Howitt, Peter. "Zero Inflation as a Long-Term Target for Monetary Policy," in Richard G. Lipsey, ed., *Zero Inflation: The Goal of Price Stability*. C. D. Howe Institute, 1990.

King, Robert G., and Sergio Rebelo. "Public Policy and Economic Growth: Developing Neoclassical Implications," *Journal of Political Economy* (October 1990, Part 2), pp. S126-50.

Lebow, David E., David J. Stockton and William L. Wascher. "Inflation, Nominal Wage Rigidity, and the Efficiency of Labor Markets," unpublished manuscript (August 1993), Board of Governors of the Federal Reserve System .

Lucas, Robert E., Jr. "On the Mechanics of Economic Development," *Journal of Monetary Economics* (July 1988), pp. 3-42.

Lucas, Robert F. "The Bank of Canada and Zero Inflation: A New Cross of Gold?" *Canadian Public Policy* (March 1989), pp. 84-93.

McLaughlin, Kenneth J. "Rigid Wages?" *Journal of Monetary Economics* (December 1994), pp. 383-414.

Ohanian, Lee E., and Alan C. Stockman. "Theoretical Issues of Liquidity Effects," this *Review* (May/June 1995), pp. 3-25.

Orphanides, Athanasios, and Robert M. Solow. "Money, Inflation and Growth," in Benjamin M. Friedman and Frank H. Hahn, eds., *Handbook of Monetary Economics,* vol. I. North-Holland, 1990.

Pagan, Adrian R., and John C. Robertson. "Resolving the Liquidity Effect," this *Review* (May/June 1995), pp. 33-54.

Reichenstein, William. "The Impact of Money on Short-Term Interest Rates," *Economic Inquiry* (January 1987), pp. 67-82.

Romer, Christina D., and David H. Romer. "New Evidence on the Monetary Transmission Mechanism," *Brookings Papers on Economic Activity* (1990:1), pp. 149-98.

Schultze, Charles L. "Microeconomic Efficiency and Nominal Wage Stickiness," *The American Economic Review* (March 1985), pp. 1-15.

Stockman, Alan C. "Anticipated Inflation and the Capital Stock in a Cash-in-Advance Economy," *Journal of Monetary Economics* (November 1981), pp. 387-93.

Summers, Lawrence. "How Should Long-Term Monetary Policy Be Determined?" Panel Discussion, *Journal of Money, Credit and Banking* (August 1991, Part 2), pp. 625-31.

Tatom, John A. "The Welfare Cost of Inflation," this *Review* (November 1976), pp. 9-22.

Thornton, Daniel L. "The Effect of Monetary Policy on Short-Term Interest Rates," this *Review* (May/June 1988), pp. 53-72.

Tobin, James. "Inflation and Unemployment," *The American Economic Review* (March 1972), pp. 1-18.

_____. "Money and Economic Growth," *Econometrica* (October 1965), pp. 671-84.

David C. Wheelock is a senior economist at the Federal Reserve Bank of St. Louis. Paul W. Wilson is an associate professor of economics at the University of Texas at Austin. Heidi L. Beyer provided research assistance.

Evaluating the Efficiency of Commercial Banks: Does Our View of What Banks Do Matter?

David C. Wheelock and Paul W. Wilson

In the past 15 years, the banking industry has faced growing competition from other financial service firms and financial markets and, at the same time, has undergone substantial deregulation and change. Proponents of further deregulation, such as the removal of barriers to the commingling of commercial and investment banking, argue that such changes would enhance the efficiency and viability of American banks.

The impact of competitive and regulatory changes on banks can be judged by gross measures of performance, such as profitability and failure rates. Economists are also interested in how such changes affect the efficiency with which banks transform resources into various financial services. Inefficiency implies that resources are wasted, that is, that firms are producing less than the feasible level of output from the resources employed, or are using relatively costly combinations of resources to produce a particular mix of products or services. Thus, a goal of policymakers, as well as stockholders and managers, is to devise policies that improve the efficiency of commercial banks.

Unfortunately, economists do not agree upon the appropriate methodology for measuring the efficiency of banks. Several estimation techniques have been proposed, each with advantages and disadvantages. The problem is complicated by the myriad of different services that commercial banks perform. Researchers deal with complex issues in measuring bank production: Is a deposit an input to the production process, or an output? Should outputs be measured in terms of the number of a bank's accounts, the number of transactions it processes or the dollar amounts of its loans or deposits? Perhaps not surprisingly, estimates of commercial bank inefficiency vary considerably across studies that use different techniques, conceptions of bank production and data samples.

This article investigates the sensitivity of efficiency measures to broadly different conceptions of how banks operate. We use a single-estimation technique and a common pool of banks to compare efficiency measures based on alternative views of bank production. We find substantial differences in mean efficiency across models and low, though statistically significant, correspondence in the rankings of banks by efficiency scores across models.

First, we discuss why measuring commercial bank efficiency is useful, some alternative measures of efficiency and techniques for estimating efficiency. A description of the approach we take, our data and our results follow.

WHY DO WE CARE ABOUT THE EFFICIENCY OF COMMERCIAL BANKS?

The performance of firms is often described in terms of their efficiency. The measured efficiency of a production unit (a firm or plant) is generically interpreted as the difference between its observed input and output levels and the corresponding optimal values. An output-oriented measure of efficiency compares observed output with the maximum output possible for given input levels. Alternatively, an input-oriented

efficiency measure compares the observed level of inputs with the minimum input that could produce the observed level of output. These are measures of technical efficiency, and as such ignore the behavioral goals of the firm.

Measures of allocative efficiency compare the observed mix of inputs or outputs with the optimal mix that would minimize cost, maximize profit or obtain any other behavioral goal. Allocative efficiency can be combined with technical efficiency to measure overall efficiency. In addition, measures of technical efficiency can be used to construct measures of scale efficiency, which involve comparison of observed and optimal scale, or size, of the firm. One can also measure scope efficiency, which involves comparison of the cost of producing the observed mix of outputs in a single firm with the costs that would prevail if each output was produced in a separate firm. Researchers have found that banks suffer more from technical inefficiency than from scale or scope inefficiency (for example, Berger and Humphrey, 1991).

The efficiency of commercial banks is important for at least two reasons. First, efficiency measures are indicators of success, by which the performance of individual banks, and the industry as a whole, can be gauged. Banks face growing competition, both from other banks and from firms and markets outside the industry (see Wheelock, 1993), and presumably banks will be more successful in maintaining their business if they operate efficiently. Berger and Humphrey (1992) find that during the 1980s high-cost banks experienced higher rates of failure than more efficient banks. Similarly, in a study of bank failures during the 1920s, Wheelock and Wilson (1995) find that the less technically efficient a bank was, the greater its likelihood of failure.

A second reason to investigate the efficiency of commercial banks is the potential impact of government policies on efficiency. One might gauge the impact of a regulatory change by measuring its effect on commercial bank efficiency, or examine efficiency among banks in different states to measure the effect of differences in branching restrictions or other regulations. Recent proposals to end the Glass-Steagall separation of commercial and investment banking stem in part from a view that broader powers could enhance the efficiency of banks and other financial institutions. Obviously, this change could enhance the scope efficiency of banks if there are complementarities in the production of commercial and investment banking services. Conceivably, such change could also improve scale or overall efficiency. Improved efficiency is also one argument made in support of interstate branching and, indeed, Grabowski, Rangan and Rezvanian (1993) find that branch banking organizations are more efficient than multiple-office bank holding companies.

Other studies have considered whether bank mergers enhance efficiency. Using different approaches, Rhodes (1993) finds that mergers have not generally improved efficiency, though Fixler and Zieschang (1993) conclude the opposite. Shaffer (1993), on the other hand, evaluates potential mergers and concludes that they could significantly reduce inefficiency for many banks of less than $10 billion of assets.

The impact of ownership or management structure on efficiency has also been studied. Pi and Timme (1993), for example, find that banks whose chief executive officer also serves as board chairman are less efficient than other banks, and Mester (1993) shows that mutual savings and loan associations are more efficient than stock S&Ls.

MEASURING COMMERCIAL BANK EFFICIENCY

The efficiency of commercial banks has been studied using a variety of techniques and samples, and, as noted above, has been used to address numerous policy issues. Recent studies typically use techniques that accommodate the multiple outputs of banks and measure the efficiency of individual banks relative to a standard set by peer institutions. Readers interested in a survey of this research can refer to Berger, Hunter and Timme (1993).

To date, no technique for measuring efficiency has been generally accepted and different methodologies appear to generate

considerable differences in measured efficiency, even when common bank samples are used. Variants of four techniques are common in the literature. The "stochastic cost frontier" approach is an econometric methodology in which deviations of a firm's actual cost from predicted cost are presumed to be due to random error and inefficiency, each of which is assumed to have a particular statistical distribution (usually the normal distribution for the random error and a half-normal for inefficiency). The "thick frontier" approach is a variant in which deviations from predicted cost within the lowest average cost quartile of banks are assumed due to random error, and the differences between the predicted costs of banks in the highest and lowest quartiles are assumed to be due to inefficiency. The "distribution-free" approach is applicable when data for more than one year are available. It assumes that inefficiency is stable over time, while random errors average out over time. That is, a bank's inefficiency for a span of years is taken to be the mean of its measured inefficiency across all years within the period. Finally, "Data Envelopment Analysis" (DEA) is a non-parametric methodology in which linear programming is used to measure the distance of individual banks from the efficient, or "best-practice," frontier. All deviations from the efficient frontier are assumed to be due to inefficiency.

Researchers have found that estimates of inefficiency are sensitive to the choice of technique. Ferrier and Lovell (1990), for example, apply the stochastic cost frontier and DEA techniques to a common sample of banks and arrive at different estimates of inefficiency. Berger (1993) finds substantial differences in measured efficiency from two variants of the distribution-free approach.

A second reason why different studies of commercial bank efficiency often reach seemingly contradictory findings might stem from differences in how a banking firm is modeled. Regardless of which of the four measurement techniques is used, the researcher must specify a list of inputs and outputs. The question, "What do banks produce?" is not simple to answer. Banks provide a variety of services, from loans and deposit accounts to trust services, safe deposit box rentals, mutual fund sales and foreign exchange transactions. Moreover, changes in regulation, technology and customer demands have caused the types of services that banks perform to change over time. For example, banks now provide a variety of securities-related services, such as underwriting and mutual fund sales, which regulators forbid a few years ago. To tractably measure efficiency, researchers are forced to begin with simplified models of the banking firm. Unreliable estimates of efficiency can stem from the use of models that omit key features of bank production.

Some researchers view banks as producers of loans and deposit accounts, and measure output by either the number of transactions or accounts serviced. This view is referred to as the "production" approach. Others argue that a bank's output should be measured in terms of the dollar volume of loans or deposits it provides, a view known as the "intermediation" approach. Most studies of inefficiency use the intermediation approach, in part because the necessary data are more readily obtained. We are aware of only one recent study taking the production approach (Ferrier and Lovell, 1990), though in the 1970s and early 1980s such studies were more common (see Gilbert, 1984). The production approach focuses on operating costs and ignores interest expense. The intermediation approach, on the other hand, includes both operating and interest expenses, and hence may be of more interest for studying the viability of banks (see Berger, Hanweck and Humphrey, 1987; or Ferrier and Lovell, 1990). For analysis of the operating efficiency of banks, however, the production approach may be of interest.

Among those who use the intermediation approach are researchers who hold the view that banks produce various loans and other investments from deposits, other funding sources, labor and materials. This "asset" approach has been criticized because it ignores the fact that banks expend considerable resources supplying transactions and savings deposits (Berger and Humphrey, 1992).

Some researchers apply empirical criteria to determine what services to consider as bank outputs and what to consider as inputs.

Berger and Humphrey (1992), for example, classify activities for which banks create high added value, such as loans, demand deposits and time and savings deposits as important outputs, with labor, physical capital and purchased funds classified as inputs. Alternatively, Aly, Grabowski, Pasurka and Rangan (1990), Hancock (1991) and Fixler and Zieschang (1993) adopt a "user-cost" framework, whereby a bank asset is classified as an output if the financial return on the asset exceeds the opportunity cost of the investment, and a liability is classified as an output if the financial cost of the liability is less than its opportunity cost. Even though their details differ, the two approaches empirically tend to suggest similar classifications of inputs and outputs. The main exception is classification of demand deposits as an output in most user-cost studies, and as both an input and an output when the value-added approach is used (see Berger and Humphrey, 1992, for more detail).

Table 1 summarizes six recent studies of commercial bank production efficiency. Although representative, this list is far from exhaustive. These studies employ a variety of estimation techniques and include a variety of different inputs and outputs in modeling the banking firm. The studies typically report inefficiency measures by bank-size grouping and for more than one type of inefficiency, though for brevity we report just the mean overall inefficiency. The reported percentages indicate the extent to which the average bank overused inputs to produce a given level of output. Thus, the 35 percent inefficiency found by Aly and others (1990) indicates that the average bank could have produced the same level of output with just 65 percent of the input levels actually used. Measured inefficiency clearly varies with estimation technique, model specification and, probably, the sample of banks used by the researcher.

In the remainder of this article, we investigate the extent to which measures of efficiency and the rankings of individual banks depend on whether the intermediation approach or production approach is employed. Because we are interested in the impact of the approach taken on measured efficiency, we use a single technique—DEA—applied to a common pool of banks. Our findings might, of course, be different if we used another technique or sample, but the purpose of this article is to investigate how sensitive efficiency measures are to the model of bank production employed.

METHODOLOGY

We trace our measures of efficiency to the work of Debreu (1951) and Farrell (1957). Boles (1966) was one of the first to use linear programming methods to measure efficiency in production using their ideas. Other extensions have collectively come to be named Data Envelopment Analysis (DEA), a term coined by Charnes, Cooper and Rhodes (1978). Lovell (1993) summarizes this literature.

Details about the efficiency measures used in this article are contained in the shaded insert on page 6. The essential ideas, however, are illustrated in Figure 1, which considers the case of a sample of firms producing a single output from two inputs, x_1 and x_2. Suppose firms A, B and C each produce a given level of output; A and B lie on the production frontier XX′, while C lies in the interior of the production set. The frontier XX′ is the set of all combinations of inputs which can produce the same level of output, and where the reduction of at least one input necessarily causes output to fall. Hence, firms A and B are regarded as efficient, whereas firm C is regarded as inefficient. Inefficient firms such as C may lie in the interior of the production set due to imperfect information, managerial incompetence or perhaps other reasons. For firm C, input weak technical efficiency (IWE) is defined as the ratio of distances OC′/OC in Figure 1. By reducing the input quantities used by firm C by this amount, the firm could move to point C′ and would be considered efficient in the IWE sense.

Next, we define input overall efficiency (IOE). In terms of Figure 1, the isocost line is given by PP′. For firm C, the IOE score is given by the ratio of distances OC″/OC. Although the point C″ lies outside the production set boundary, and hence is not feasible, input costs at C″ are the same as at B, which is a feasible point. Hence, if firm C

Table 1

Selected Studies of Commercial Bank Production Inefficiency

Study, Technique, Approach	Inputs	Outputs	Sample	Results
Aly and others (1990); DEA; intermediation	labor, physical capital, loanable funds[1]	real estate loans, commercial loans, consumer loans, all other loans, demand deposits	random, 322 banks, 1986 data	overall inefficiency: 35%
Berger and Humphrey (1991); thick frontier; intermediation	labor, physical capital, purchased funds[2]	demand deposits, retail time and savings deposits, real estate loans, commercial loans, installment loans	all banks, 1984 data	total inefficiency[3]: 24% (branching states), 19% (unit banking states)
Elyasiani and Mehdian (1990); DEA intermediation	labor, physical capital, demand deposits, time and savings deposits	securities held, real estate loans, commercial loans, all other loans	191 banks with assets over $300 million, 1980 data	technical inefficiency: 10%
Ferrier and Lovell (1990) DEA and stochastic cost frontier; production	labor, occupancy costs, expenditure on material	number of: demand deposit accounts, time and savings deposit accounts, real estate loans, installment loans, commercial loans	575 banks, 1984 data	overall inefficiency: 21% (DEA), 26% (stochastic cost frontier)
Hunter and Timme (1995) distribution free; intermediation	labor, physical capital,[4] purchased funds, transactions accounts,[4] non-transactions accounts under $100,000[4]	commercial and security loans, consumer loans, all other loans, non-interest income	317 banks with assets over $1 billion, 1985-90 data	overall inefficiency: 30-54% (depending on model); 23-36% (omitting 1% extreme values)
Kaparakis and others (1994); stochastic cost frontier; intermediation	labor, physical capital, interest bearing deposits under $100,000, non-interest bearing deposits,[4] purchased funds	loans to individuals, real estate loans, commercial loans, other[5]	5,548 banks with assets over $50 million, 1986 data	overall inefficiency: 10%

[1] time and savings deposits, notes and debentures and other borrowed funds.
[2] federal funds purchased, time deposits over $100,000, foreign deposits and other borrowed funds.
[3] includes inefficiencies due to excessive deposit interest paid and purchased fund interest paid.
[4] input treated as "quasi-fixed," that is, not variable in the short run.
[5] fed funds sold, securities held, securities and other assets in trading accounts.

were to become efficient in the IOE sense, its input mix would have to be altered; the IOE score, however, can be obtained by considering the hypothetical proportionate reductions of inputs represented by point C″.

In terms of Figure 1, allocative efficiency for firm C is given by the ratio of distances OC″/OC′. Allocative inefficiency arises from using a combination of inputs that does not minimize total cost, as opposed to technical inefficiency, which is a proportionate overuse of all inputs.

Finally, we can determine scale efficiency by comparing IWE computed under the assumption that the firm is operating at constant returns-to-scale with IWE obtained previously. A score of unity implies that the firm is operating under constant returns. While a score other than 1 does not translate

A MATHEMATICAL DESCRIPTION OF EFFICIENCY MEASUREMENT

We use measures of efficiency discussed by Färe, Grosskopf and Lovell (1985). First, we compute the input weak technical efficiency (IWE) score for the ith firm in a sample by solving the linear programming problem:

$$\begin{aligned} &\min W_i \\ &\text{subject to} \quad Xq_i \le W_i x_i \\ &\qquad\qquad\quad Yq_i \ge y_i \\ &\qquad\qquad\quad \vec{1}q_i = 1 \\ &\qquad\qquad\quad q_i \in \Re_+^N, \end{aligned} \tag{1}$$

where n firms produce s outputs using m inputs, q_i is a $(N \times 1)$ vector of weights to be computed for the ith firm, $0 < W_i \le 1$ is a scalar, x_i is a $(m \times 1)$ vector of inputs for the kth firm, y_i is a $(s \times 1)$ vector of outputs for the kth firm, $X = [x_1,...,x_N]$ is a $(m \times N)$ matrix of observed inputs, $Y = [y_1,...,y_N]$ is a $(s \times N)$ matrix of observed outputs and $\vec{1}$ is a $(1 \times N)$ vector of ones.

The minimand W_i in equation 1 measures the input weak efficiency of the ith firm. The inequality constraints in equation $\vec{1}$ define a reference technology with strong disposability of inputs; constraining the weights in q to sum to unity allows the reference technology to exhibit variable returns to scale. For the ith firm, W_i gives the proportion by which inputs can be reduced to move the firm from the interior of the production set onto the piecewise-linear boundary of the production set corresponding to the reference technology in 1.

Next, we compute input overall efficiency (IOE) score O_i for the ith firm by first solving the linear program:

$$\begin{aligned} &\min_{x_i^*} p_i x_i^* \\ &\text{subject to} \quad Xq_i \le x_i^* \\ &\qquad\qquad\quad Yq_i \ge y_i \\ &\qquad\qquad\quad \vec{1}q_i = 1 \\ &\qquad\qquad\quad q_i \in \Re_+^N \\ &\qquad\qquad\quad x_i^* \in \Re_+^m \end{aligned} \tag{2}$$

where X, Y, x_i and y_i are defined as in equation 1, p_i is a $(1 \times m)$ vector of input prices, and x_i^* is an $(m \times 1)$ vector of efficient inputs to be computed. The IOE score may be defined as

$$O_i \equiv p_i x_i^* / p_i x_i. \tag{3}$$

The constraints in equation 2 are similar to those in 1. The same reference technology is defined by the constraints in 2, but instead of proportionately reducing inputs until the ith firm lies on the reference technology, inputs are further reduced proportionately until the firm lies on the isocost plane tangent to the production set boundary.

An allocative efficiency score, A_i, may be defined by dividing the IOE score by the IWE score:

$$A_i \equiv O_i / W_i. \tag{4}$$

The efficiency scores obtained from 1 measure technical efficiency as the distance to the relevant isoquant, but do not consider where the firm is situated along the variable–returns production frontier. To measure scale efficiency, equation 1 must be recomputed for each firm, first assuming constant returns to scale by removing the restriction $Cq_i = 1$, and then assuming non-increasing returns-to-scale by imposing the restriction $Cq_i \le 1$. In the case of IWE, this produces efficiency scores W_i^{CRS} and W_i^{NIRS}, respectively, for the ith firm. The scale efficiency score corresponding to 1 is then defined as

$$S_i \equiv W_i^{CRS} / W_i. \tag{5}$$

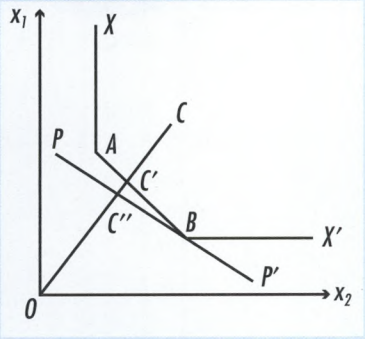
Clearly, $0 < S_i \le 1$ since $W_i^{CRS} \le W_i^{NIRS} \le W_i$. If $S_i = 1$, then the ith firm is scale-efficient, that is, the firm is operating at the point of constant returns on the production frontier. If $S_i < 1$, then the firm is scale-inefficient due to either decreasing returns if $W_i^{NIRS} = W_i$, or increasing returns if $W_i^{NIRS} < W_i$.

easily into a specific percentage deviation from constant returns, the scores are useful for ranking firms by the extent of their inefficiency.

Each of the efficiency scores described above measures efficiency in an input orientation; efficiency is measured by holding output fixed and determining the maximum feasible reduction in inputs. Efficiency can also be measured by holding inputs fixed and determining the maximum feasible expansion of outputs. Since the efficiency measures we use do not imply underlying assumptions regarding the behavior of firms, the choice between input and output orientations is somewhat arbitrary; one might compute both types of efficiency measures to get more information than can be obtained from either the input or output orientations alone. Note that both IWE and IOE are *radial* measures of efficiency, that is, in each case efficiency is measured along a ray emanating from the origin and passing through the firm in input-output space. Consequently, the efficiency scores are independent of the units of measurement used for both inputs and outputs, which is advantageous since units of measurement may always be defined arbitrarily. Färe and others (1985) observe that some DEA formulations do not share this property.

Figure 1

Measuring Technical, Allocative and Overall Efficiency



EMPIRICAL IMPLEMENTATION

For our empirical analysis of commercial bank efficiency, we use a sample of banks participating in the Federal Reserve System's Functional Cost Analysis (FCA) program for 1993. Participants in this program supply information about their operations and costs which are not generally available for banks, and which are necessary to measure efficiency using the production approach. After eliminating observations with missing values and observations for depository institutions other than commercial banks, data for 269 banks remain.

Because participation in the FCA program is voluntary, the banks in our sample may not be representative of the industry as a whole. For example, whereas the average total assets at the end of 1993 for FCA program banks was $163.6 million, with a range from $8.0 million to $2,602.8 million, average total assets were $300.7 million, with a range from $1.0 million to $108,223.0 million, for all U.S. commercial banks (as reported in the Federal Deposit Insurance Corporation Reports of Condition, that is, the "Call Reports"). The average return on assets of 1.15 percent for the banks in our sample, however, was approximately the same as the average for all banks (1.12 percent). Nevertheless, because our sample of banks is not random, the efficiency measures calculated here should not be interpreted as reflecting the efficiency of commercial banks in general.

For the production approach to modeling bank activities, we construct variables using definitions from Ferrier and Lovell (1990):

Outputs:

y_1 = number of demand deposit accounts
y_2 = number of time deposit accounts
y_3 = number of real estate loans
y_4 = number of installment loans
y_5 = number of commercial loans

Inputs:

x_1 = number of employees
x_2 = occupancy costs and expenditure on furniture and equipment
x_3 = expenditure on materials

Table 2

Descriptive Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
y_1	7902.91	13091.76	469.00	173362.00
y_2	7618.43	10680.64	413.00	106821.00
y_3	998.23	1503.99	0.00	13456.00
y_4	3134.01	6918.63	44.00	87794.00
y_5	899.63	1744.37	0.00	23998.00
y'_1	40764316.60	60020700.09	1387963.00	653519000.00
y'_2	99272775.89	144141890.79	5103000.00	1616691000.00
y'_3	42038331.61	54512376.38	0.00	377449000.00
y'_4	16254832.77	43838042.24	127890.00	617136000.00
y'_5	28553775.52	64197194.17	0.00	895471000.00
x_1	88.02	144.09	3.89	1730.07
x_2	776304.50	1260920.37	790.00	13090834.00
x_3	330106.61	576912.53	11859.00	6948552.00
w_1	30827.13	5905.65	19222.38	67832.63
w_2	0.0054	0.0025	0.0001	0.0270
w_3	0.0023	0.0007	0.0010	0.0057
w'_1	30827.13	5905.65	19222.38	67832.63
w'_2	52.06	48.43	0.75	715.17
w'_3	21.88	11.69	2.36	131.86
v_1	17045.20	37140.71	166.00	431227.00
v_2	54417.53	96144.16	213.00	1279962.00
v_3	16755.40	29282.60	93.00	270480.00
v_4	5980.03	12048.51	100.00	137300.00
u_1	110490.98	158765.47	5138.00	1894477.00
u_2	18323.86	43312.12	200.00	410249.00
u_3	89.32	150.08	4.00	1856.00
u_4	3008.39	5252.85	13.00	47511.00
p_1	0.0347	0.0048	0.0143	0.0489
p_2	0.0392	0.0110	0.0066	0.1000
p_3	31.11	6.18	20.38	77.91
p_4	0.34	0.31	0.08	2.83

Input prices:

w_1 = total expenditure on salaries and fringe benefits/x_1

w_2 = x_2/level of deposits

w_3 = x_3/level of deposits

For the intermediation approach to bank production, we construct variables using definitions from Kaparakis, Miller and Noulas (1994):

Outputs:

v_1 = loans to individuals for household, family, and other personal expenses

v_2 = real estate loans

v_3 = commercial and industrial loans

v_4 = federal funds sold, securities purchased under agreements to resell, plus total securities held in trading accounts

Inputs:

u_1 = interest-bearing deposits except certificates of deposit greater than $100,000

u_2 = purchased funds (certificates of deposit greater than $100,000, federal funds purchased, and securities sold plus demand notes) and other borrowed money

u_3 = number of employees

u_4 = premises and fixed assets

Input prices:

p_1 = average interest cost per dollar of u_1

p_2 = average interest cost per dollar of u_2

p_3 = average annual wage per employee

p_4 = average cost of premises and fixed assets.

In addition, Kaparakis and others also define quasi-fixed input, non-interest bearing deposits, for which there is no corresponding price. Other studies adopting the intermediation approach have ignored this item, as we do in the results reported below. Including non-interest bearing deposits as a fifth input when measuring technical or scale efficiency seems to have little effect on the results.

To form a specification midway between the production approach represented by the Ferrier and Lovell (1990) specification and the intermediation approach represented by the Kaparakis and others (1994) specification, we define $y'_1,\ldots,y'_5$ as the dollar amount of each account or loan corresponding to $y_1,\ldots,y_5$, respectively. Because outputs are now measured in dollar amounts, this model is best classified as representing the intermediation approach, even though the choice of variables is based on Ferrier and Lovell (1990). In addition, we define an alternative price system, w'_1, w'_2, w'_3, for the Ferrier and Lovell specification, where $w'_1 = w_1$, and w'_2 and w'_3 are computed similarly to w_2 and w_3 except that level of deposits is replaced by the *number* of time and demand deposits. This seems to us to make the mapping of inputs and outputs under the production approach more consistent. We report summary statistics for each of the variables in Table 2.

RESULTS OF EFFICIENCY MEASUREMENT

We compute the various efficiency measures for the five models summarized below:

Model	Inputs	Outputs	Input Prices
1	x_1-x_3	y_1-y_5	w_1-w_2
2	x_1-x_3	y_1-y_5	$w'_1-w'_2$
3	x_1-x_3	$y'_1-y'_5$	w_1-w_2
4	x_1-x_3	$y'_1-y'_5$	$w'_1-w'_2$
5	u_1-u_4	v_1-v_4	p_1-p_4

Models 1 and 2 correspond to the Ferrier and Lovell (1990) specification, with alternative price definitions. Models 3 and 4 provide a bridge to the intermediation approach by replacing the number of accounts and loans in the output variables with dollar amounts. Model 5 is the Kaparakis and others (1994) specification.

Table 3 presents the mean scores for each type of efficiency described in the preceding section. Note that since the same inputs and outputs are used in models 1 and 2, and models 3 and 4, the technical and scale efficiency scores are the same for these models. For each efficiency measure, Table 3 also shows the standard deviation of the scores across the 269 banks in the sample, the number of banks having an efficiency score of unity (labeled "Number Efficient"), that is, the number of banks operating on the efficient frontier, as well as 90 and 95 percent confidence intervals for the mean. Given the large number of banks with efficiency scores of unity, particularly in the case of technical efficiency, and since all of the efficiency scores are defined to lie between zero and 1, the underlying distributions of the individual efficiency scores are clearly non-normal. Results from Atkinson and Wilson (1995), however, suggest that our sample size of 269 is easily large enough for us to rely on the asymptotic normality of the sample means implied by the central limit theorem, and thus to compute confidence intervals based on a normal distribution.

In several cases, the confidence intervals for the means reported in Table 3 overlap.

Table 3

Efficiency Scores (269 observations)

Model	Mean	Standard Error	Number Efficient	90% Confidence Interval		95% Confidence Interval	
Technical efficiency (W_k):							
1,2	0.6348	0.0141	42	0.6115	0.6581	0.6071	0.6626
3,4	0.7675	0.0103	52	0.7505	0.7846	0.7472	0.7879
5	0.8088	0.0107	75	0.7911	0.8265	0.7877	0.8299
Scale efficiency (S_k):							
1,2	0.8833	0.0067	32	0.8723	0.8943	0.8701	0.8964
3,4	0.9452	0.0047	49	0.9374	0.9530	0.9359	0.9545
5	0.9414	0.0057	7	0.9319	0.9509	0.9301	0.9527
Allocative efficiency (A_k):							
1	0.7698	0.0082	7	0.7562	0.7833	0.7536	0.7860
2	0.4992	0.0108	4	0.4814	0.5170	0.4780	0.5204
3	0.7924	0.0068	7	0.7812	0.8036	0.7790	0.8057
4	0.6340	0.0108	7	0.6162	0.6518	0.6128	0.6552
5	0.7838	0.0080	13	0.7706	0.7970	0.7680	0.7996
Overall efficiency (O_k):							
1	0.4835	0.0116	7	0.4644	0.5026	0.4607	0.5063
2	0.3356	0.0128	4	0.3146	0.3567	0.3105	0.3607
3	0.6053	0.0095	7	0.5897	0.6209	0.5867	0.6239
4	0.4928	0.0121	7	0.4727	0.5128	0.4689	0.5166
5	0.6320	0.0107	13	0.6144	0.6496	0.6110	0.6529

We test for significant differences among the means of each efficiency measure across different models. At the 0.05 significance level, we are unable to reject the null hypothesis of equivalent means in the following cases: (1) scale efficiency for models 3, 4 and 5; (2) allocative efficiency for models 1 and 5, and models 3 and 5 (we do reject the null hypothesis when comparing allocative efficiency among models 1 and 3); and (3) overall efficiency for models 1 and 4. In all other instances, we reject the null hypothesis of no difference. Even the seemingly innocuous modification of redefining the input prices between models 1 and 2, and between 3 and 4 has a large effect on mean allocative and overall efficiency. Note also that for the most extreme comparison, models 1 (the production view) and 5 (the intermediation view), we reject the null hypothesis of equal levels of technical and overall efficiency. This suggests that, at least for this sample of banks, average efficiency does depend on the view of bank production assumed by the researcher. We find that average technical and overall efficiency is higher under the intermediation approach (model 5) than under the production approach (model 1). Our finding for overall inefficiency of 37 percent using model 5 is similar to what Aly and others (1990) found for their sample, though substantially greater than what Kaparakis and others (1994) found for theirs (see Table 1).

It is possible to determine whether a particular bank lies on the increasing (IRS), constant (CRS) or decreasing (DRS) returns portion of the technology. Table 4 shows

the results of this analysis, considering only banks that were found to be technically efficient.[1] Thus, for example, 16 banks, or 38.1 percent of all technically efficient banks, operated on the constant-returns portion of the technology under models 1 and 2.

We test the null hypothesis of no association among the rows and columns of the matrix represented by Table 4 using Pearson's chi-square test, the likelihood ratio chi-square statistic, and Fisher's exact test.[2] For the entire matrix, all three tests reject the null hypothesis of no differences in the proportions in each row and column. However, when we perform pairwise tests by deleting individual rows from Table 4, we fail to reject the null hypothesis of no difference for models 3, 4 and 5, and for models 1, 2 and 3, 4. Each of the three tests fail to reject at the 90 percent level.

In the case of models 1, 2 and 5, we reject the null hypothesis of no difference in the proportions at greater than 99 percent. Thus, while we find evidence of similarity in terms of returns-to-scale when comparing models 3, 4 with either models 1, 2 or 5, models 1, 2 and 5 appear different in terms of returns-to-scale. More banks appear to be operating under constant returns-to-scale when the intermediation approach is taken (model 5) than when the production approach is used (model 1). Since returns-to-scale at a given location on the production frontier depend upon the shape of the variable-returns technology, these results indicate that the technology implicitly estimated by models 3, 4 is similar to the technologies implied by models 1, 2 and 5, which in turn are significantly different. This is consistent with our view of models 3, 4 as a bridge between the production approach represented by models 1, 2 and the intermediation approach represented by model 5. The result also suggests that differences between the two approaches might be due not only to use of number of accounts and loans versus dollar amounts, but also to the treatment of time deposits as an output or an input.

In addition to comparing mean efficiency scores, we use the Wilcoxon matched-pairs signed-ranks test, a sign test for equality of medians and Kendall's τ-statistic to further examine the similarity of efficiency scores across the five models. We report the results of these tests in Table 5.[3]

Table 4

Returns to Scale

Model	CRS	DRS	IRS
1,2	16	24	2
	(38.1%)	(57.1%)	(4.8%)
3,4	31	20	1
	(59.6%)	(38.5%)	(1.9%)
5	48	21	6
	(64.0%)	(28.0%)	(8.0%)

The Wilcoxon matched-pairs signed-ranks test analyzes the equality of distributions without making assumptions regarding the form the distributions might take. The values shown in parentheses in the second column of Table 5 give the two-tailed normal probabilities associated with the test statistic. Hence, we fail to reject the null hypothesis of identical distributions when comparing scale efficiency scores from models 3, 4 and 5, when comparing allocative efficiency scores from models 1 and 5, and from models 3 and 5, and when comparing overall efficiency scores from models 1 and 4. In all other cases, we reject the null hypothesis. It appears that, for the most part, the distributions of the various efficiency scores do vary across models.

The sign test for equivalence of medians yields a two-tailed binomial probability, which we also report in Table 5. In only two instances do we not reject the null hypothesis of equal medians: when comparing scale efficiency scores from models 3, 4 and 5, and when comparing allocative efficiency scores from models 3 and 5. These results are consistent with our finding that, in most cases, average efficiency varies across models.

Finally, rather than comparing the distributions of efficiency scores from different models, we use the scores to *rank* banks in terms of their estimated efficiency. Kendall's τ-statistic measures the correlation among the ranks of banks from two models and provides a statistical test of the null hypothesis of no association between two sets of

[1] Since we are using an input orientation, we could also examine whether inefficient banks would lie on the increasing-, constant- or decreasing-returns portion of the technology if inputs were proportionately contracted to move the bank to the frontier. However, since the path a bank might actually take to reach the frontier if the sources of inefficiency were removed depends upon behavioral goals, we ignore technically inefficient banks here.

[2] Details on these computations may be found in the *Stata Reference Manual: Release 3.1*, Stata Corporation (1993).

[3] See Snedecor and Cochran (1989) and Kendall and Gibbons (1990) for a discussion of these tests. Computational details are given in the *Stata Reference Manual: Release 3.1*, Stata Corporation (1993).

Table 5

Statistical Comparison of Efficiency Scores
(probability values in parentheses)

Model	Wilcoxon Signed-Ranks	Sign Test	Kendall's τ
Technical efficiency (W_x):			
(2,2) / (3,4)	–8.24 (0.0001)	0.0000	0.2513 (0.0000)
(1,2) / 5	–8.90 (0.0001)	0.0000	0.0909 (0.0262)
(3,4) / 5	–3.25 (0.0012)	0.0034	0.1081 (0.0082)
Scale efficiency (S_k):			
(1,2) / (3,4)	–9.46 (0.0001)	0.0000	0.2333 (0.0000)
(1,2) / 5	–7.58 (0.0001)	0.0000	0.1114 (0.0065)
(3,4) / 5	–0.32 (0.7482)	0.4644	0.1805 (0.0000)
Allocative efficiency (A_k):			
(1,2)	12.88 (0.0001)	0.0000	0.0068 (0.8680)
(1,3)	–3.58 (0.0003)	0.0003	0.4179 (0.0000)
(1,4)	8.40 (0.0001)	0.0000	-0.0964 (0.0185)
(1,5)	–1.45 (0.1481)	0.0327	0.0934 (0.0225)
(2,3)	–13.39 (0.0001)	0.0000	0.0267 (0.5135)
(2,4)	–12.28 (0.0001)	0.0000	0.5661 (0.0000)
(2,5)	–13.03 (0.0001)	0.0000	0.0342 (0.4036)
(3,4)	10.23 (0.0001)	0.0000	-0.0047 (0.9092)
(3,5)	0.32 (0.7515)	0.6258	0.0627 (0.1255)
(4,5)	–10.49 (0.0001)	0.0000	0.1648 (0.0001)

Table 5 (cont.)

Statistical Comparison of Efficiency Scores
(probability values in parentheses)

Model	Wilcoxon Signed-Ranks	Sign Test	Kendall's τ
Overall efficiency (O_k):			
(1,2)	12.31 (0.0001)	0.0000	0.5874 (0.0000)
(1,3)	–9.08 (0.0001)	0.0000	0.2518 (0.0000)
(1,4)	–1.06 (0.2903)	0.0327	0.3072 (0.0000)
(1,5)	–9.17 (0.0001)	0.0000	0.1329 (0.0012)
(2,3)	–12.74 (0.0001)	0.0000	0.1704 (0.0000)
(2,4)	–11.62 (0.0001)	0.0000	0.5000 (0.0000)
(2,5)	–12.36 (0.0001)	0.0000	0.0405 (0.3222)
(3,4)	9.76 (0.0001)	0.0000	0.4128 (0.0000)
(3,5)	–2.67 (0.0076)	0.0015	0.1034 (0.0115)
(4,5)	–9.01 (0.0001)	0.0000	0.1831 (0.0000)

rankings. The statistic is approximately normally distributed, with zero expected value and with variance

$$VAR(\tau) = \frac{4N+10}{9N(N-1)},$$

where N gives the number of observations. By definition, the statistic lies between -1 and +1, taking a value of +1 if rankings are in complete agreement, or -1 if the ranks are completely reversed.

Kendall and Gibbons (1990) suggest that the τ-statistic may also be viewed as a measure of concordance. Any two pairs of ranks (u_i,v_i) and (u_j, v_j), $i, j = 1,...,N$, $i \neq j$, are defined as concordant if $v_i < v_j$ when $u_i < u_j$ or $v_i > v_j$ when $u_i > u_j$. Similarly, they are defined as discordant if $v_i < v_j$ when $u_i > u_j$ or $v_i > v_j$ when $u_i > u_j$. The total number of pairs is

$N(N-1)/2$, and τ can be shown to be equivalent to the proportion of concordant pairs minus the proportion of discordant pairs.

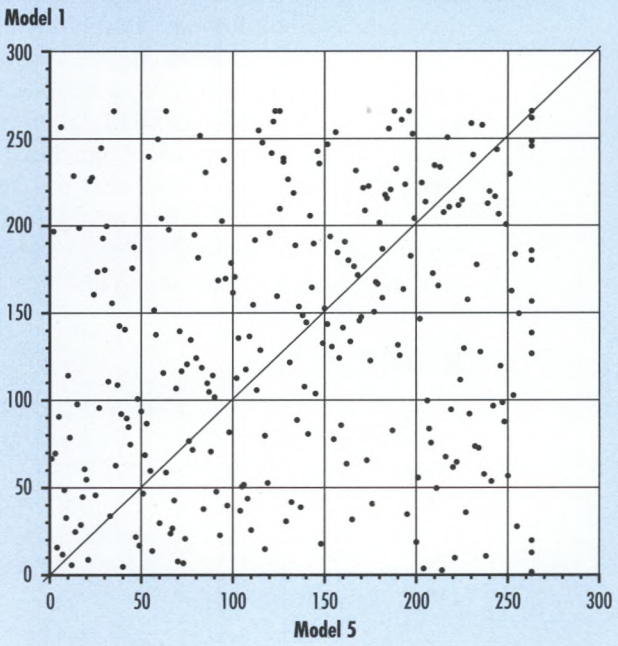
The last column of Table 5 gives the τ-statistic for the various pairs of models for each measure of efficiency, along with significance levels as shown in parentheses. We fail to reject the null hypothesis of no association among ranks in only five instances when comparing allocative efficiency scores. (In particular, we fail to reject the null hypothesis for the following pairs of models: 1, 2; 2, 3; 2, 5; 3, 4; and 3, 5, and in only one case when comparing overall efficiency scores—for models 2 and 5). In all other cases, we reject the null hypothesis of no correlation.

Note, however, that when we reject the null hypothesis of no association, the τ-statistic is usually rather small in absolute terms; the largest value of the statistic shown in Table 5 is 0.5874 (in the case of overall efficiency for models 1 and 2). As is typical in classical hypothesis testing, rejection of the null hypothesis does not necessarily imply acceptance of an alternative hypothesis. That is, our statistical test may reject the hypothesis that the rankings are not associated, but that does not necessarily imply that the rankings are associated–the test is simply not that powerful. Figure 2 plots the rankings of overall efficiency scores for model 1 against those for model 5. Note that the value of Kendall's τ-statistic for this comparison is significantly different from zero at the 0.0012 level, indicating that the two sets of rankings are not discordant. The low value of the test statistic (0.1329), however, suggests that neither are they concordant.

Our results based on the Wilcoxon matched-pairs signed-ranks test and the sign test for equivalence of medians are consistent with our observations on the differences of mean efficiency scores across the models discussed earlier. Taken together, our results indicate that different model specifications are likely to produce different measures of the level of inefficiency among a sample of banks, but not necessarily dissimilar *rankings* of individual banks in terms of measured efficiency. For our data, the rankings are similar enough to reject the null hypothesis of no association, but in many cases are far from being in complete concordance. Concordance is relatively high for the technical and scale efficiency measures, which do not rely on price data. The introduction of price data to measure allocative and overall efficiency might also introduce more sources of noise or error.

Figure 2

Rankings of Overall Efficiency Scores



CONCLUSION

Like other studies of commercial bank efficiency, we find considerable inefficiency among banks in our sample. Other studies have found substantial variation in efficiency measures in applying different estimation techniques to a common pool of banks. We find that measured efficiency also depends on the researcher's conception of what banks do. In this article, we measure various types of production efficiency under two very different views of banking. We find that, on average, technical and overall efficiency is higher under the intermediation view of the banking firm than under the production view. Mean allocative efficiency is, however, similar

under the extreme versions of each approach. Under the intermediation view, we also find somewhat less scale inefficiency and more banks operating on the constant-returns portion of the efficient frontier. Despite the differences in mean measured efficiency across the different conceptions of how banks operate, however, we find some similarity in the rankings of efficiency scores of individual banks. Further research will, of course, be necessary to determine how sensitive these findings are to the particular dataset and estimation techniques employed in this article.

REFERENCES

Aly, Hassan Y., Richard Grabowski, Carl Pasurka and Nanda Rangan. "Technical, Scale, and Allocative Efficiencies in U.S. Banking: An Empirical Investigation," *Review of Economics and Statistics* (May 1990), pp. 211-18.

Atkinson, Scott, and Paul W. Wilson. "Comparing Mean Efficiency and Productivity Scores Across Small Samples: A Bootstrap Methodology," *Journal of Productivity Analysis* (forthcoming).

Berger, Allen N. "Distribution-Free Estimates of Efficiency in the U.S. Banking Industry and Tests of the Standard Distributional Assumptions," *Journal of Productivity Analysis* (1993), pp. 261-92.

_____, G. A. Hanweck and D. B. Humphrey. "Competitive Viability in Banking: Scale, Scope, and Product Mix Economies," *Journal of Monetary Economics* (1987), pp. 501-20.

_____ and David B. Humphrey. "Measurement and Efficiency Issues in Commercial Banking," in Zvi Griliches, ed., *Output Measurement in the Services Sector.* University of Chicago Press, 1992, pp. 245-79.

_____ and _____. "The Dominance of Inefficiencies Over Scale and Product Mix Economies in Banking," *Journal of Monetary Economics* (August 1991), pp. 117-48.

_____, William C. Hunter and Stephen G. Timme. "The Efficiency of Financial Institutions: A Review and Preview of Research, Past, Present and Future," *Journal of Banking and Finance* (April 1993), pp. 221-49.

Boles, J. "Efficiency Squared: Efficient Computations of Efficiency Indexes," *Western Farm Economics Association Proceedings* (1966), pp. 137-42.

Charnes, A., W. W. Cooper and E. Rhodes. "Measuring the Efficiency of Decision Making Units," *European Journal of Operational Research* (1978), pp. 429-44.

Debreu, Gerard. "The Coefficient of Resource Utilization," *Econometrica* (1951), pp. 273-92.

Elyasiani, Elyas, and Seyed M. Mehdian. "A Nonparametric Approach to Measurement of Efficiency and Technological Change: The Case of Large U.S. Commercial Banks," *Journal of Financial Services Research* (July 1990), pp. 157-68.

Färe, Rolf, Shawna Grosskopf and C. A. Knox Lovell. *The Measurement of Efficiency of Production.* Kluwer-Nijoff Publishing, 1985.

Farrell, M. J. "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society, Series A* (1957), pp. 253-81.

Ferrier, Gary D., and C. A. Knox Lovell. "Measuring Cost Efficiency in Banking: Econometric and Linear Programming Evidence," *Journal of Econometrics* (October/November 1990), pp. 229-45.

Fixler, Dennis J., and Kimberly D. Zieschang. "An Index Number Approach to Measuring Bank Efficiency: An Application to Mergers," *Journal of Banking and Finance* (April 1993), pp. 437-50.

Gilbert, R. Alton. "Bank Market Structure and Competition: A Survey," *Journal of Money, Credit and Banking* (November 1984), pp. 617-45.

Grabowski, Richard, Nanda Rangan and Rasoul Rezvanian. "Organizational Forms in Banking: An Empirical Investigation of Cost Efficiency," *Journal of Banking and Finance* (April 1993), pp. 531-38.

Hancock, Diana. *A Theory of Production for the Financial Firm.* Kluwer Academic Publishers, 1991.

Hunter, William C., and Stephen G. Timme. "Core Deposits and Physical Capital: A Reexamination of Bank Scale Economies and Efficiency with Quasi-Fixed Inputs," *Journal of Money, Credit and Banking* (February 1995), pp. 165-85.

Kaparakis, Emmanuel I., Stephen M. Miller and Athanasios G. Noulas. "Short-run Cost Inefficiency of Commercial Banks: A Flexible Stochastic Frontier Approach," *Journal of Money, Credit and Banking* (November 1994), pp. 875-93.

Kendall, M., and J. D. Gibbons. *Rank Correlation Methods.* Oxford University Press, 1990.

Lovell, C. A. Knox. "Production Frontiers and Productive Efficiency," in Hal Fried, C. A. Knox Lovell and Shelton S. Schmidt, eds., *The Measurement of Productive Efficiency: Techniques and Applications.* Oxford University Press, 1993, pp. 3-67.

Mester, Loretta J. "Efficiency in the Savings and Loan Industry," *Journal of Banking and Finance* (April 1993), pp. 267-86.

Pi, Lynn, and Stephen G. Timme. "Corporate Control and Bank Efficiency," *Journal of Banking and Finance* (April 1993), pp. 515-30.

Rhoades, Stephen A. "Efficiency Effects of Horizontal (In-Market) Bank Mergers," *Journal of Banking and Finance* (April 1993), pp. 411-22.

Shaffer, Sherrill. "Can Megamergers Improve Bank Efficiency?" *Journal of Banking and Finance* (April 1993), pp. 423-36.

Snedecor, G. W., and W. G. Cochran. *Statistical Methods.* Iowa State University Press, 1989.

Stata Reference Manual: Release 3.1. State Corporation, 1993.

Wheelock, David C. "Is the Banking Industry in Decline? Recent Trends and Future Prospects from a Historical Perspective," this *Review* (September/October 1993), pp. 3-22.

_____ and Paul W. Wilson. "Explaining Bank Failures: Deposit Insurance, Regulation and Efficiency," *Review of Economics and Statistics* (forthcoming).

All nonproprietary and nonconfidential data and programs for the articles published in *Review* are available to our readers and can be obtained from three sources:

1. **FRED (Federal Reserve Economic Data), an electronic bulletin board service.** You can access FRED by dialing (314) 621-1824 through a modem-equipped personal computer. For a free brochure on FRED, please call (314) 444-8809.

2. **Research Department,** Federal Reserve Bank of St. Louis, Post Office Box 442, St. Louis, MO 63166-0442. You can request data and programs on either disk or hard copy. Please include the author, title, issue date and page numbers with your request.

3. **Inter-university Consortium for Political and Social Research (ICPSR).** Member institutions can request data through the CDNet Order facility. Nonmembers should write to: ICPSR, Institute for Social Research, P.O. Box 1248 Ann Arbor, Michigan 48106, or call (313) 763-5010.

Federal Reserve Bank of St. Louis
Post Office Box 442
St. Louis, Missouri 63166-0442