Reflections on Monetary Policy

J. Alfred Broaddus, Jr.

It is a pleasure and indeed an honor to be with you this evening. I must confess that when I recall the long line of distinguished economists who have delivered the Sandridge lecture, I wonder whether I am really worthy of this opportunity. But in any case I am grateful for it and will strive to make the most of it.

I have worked at the Federal Reserve Bank of Richmond for just about a quarter of a century, and for virtually all of that time I have been involved in one way or another in the formation of monetary policy. For most of that period I was an advisor to the president of the Richmond Fed, and for the last two years I have served as president myself. Given this background, I believe the most useful thing I can probably do this evening is to make a few remarks about monetary policy and some of the major issues the Fed is facing in conducting policy currently, in the context of my experience with the policymaking process over the years.

The last 25 years have been extraordinarily eventful ones for monetary policy in many ways. In this period there were fundamental changes in attitudes among policymakers, financial market participants, and the public regarding the appropriate role of monetary policy and also about some of the procedures used by the Fed in implementing policy decisions. The major factor triggering this reevaluation without any doubt was the inflation that began at the end of the 1960s and peaked at about 13 percent at the beginning of the 1980s. This rise in inflation was unprecedented in recent peacetime American history; it was largely unexpected by the public and the Fed; and it severely challenged widely held assumptions about the economy and inflation prevailing at the time.

This article is adapted from the Sandridge Lecture delivered by J. Alfred Broaddus, Jr., president of the Federal Reserve Bank of Richmond, at the annual meeting of the Virginia Association of Economists in Richmond on March 16, 1995. Mr. Broaddus wishes to thank his long-time colleague, Timothy Cook, for substantial assistance in preparing the address. The address also draws on several published and unpublished articles by other members of the Bank's staff. The views expressed are those of the author and not necessarily those of the Federal Reserve System.

The inflation in the 1970s was followed by a severe recession in the early 1980s and, subsequently, a sharp deceleration in the rate of inflation to approximately 4 to 5 percent. Most recently, as you know, inflation has been running at about a 3 percent rate, which is the lowest rate since I began my career at the Fed. If the 1970s taught us the necessity of containing inflation, I would say that the major lesson of the 1980s was the importance of (1) having a long-run strategy to achieve that goal and (2) maintaining the public's confidence in that strategy or, to use the currently popular jargon, maintaining the *credibility* of the strategy.

Tonight I want to look back over my years at the Fed, explain to you how developments over this period have influenced thinking about monetary policy and how it should be conducted, and share some of my own views with you. My purpose is not so much to convince you of the wisdom of my views, although I certainly hope you find at least some of them persuasive, but to give you perhaps a fuller appreciation of the fundamental issues facing monetary policy today.

1. THE ORIGIN OF THE FEDERAL RESERVE AND ITS MANDATE

Let me begin with just a few introductory comments about the Fed. Most if not all of you are probably familiar with the Fed; nonetheless, a brief review may increase your appreciation of some of the points I will be making.

The Federal Reserve was established by Congress in 1914. Initially, the Fed's main purpose or "mandate" was to cushion short-term interest rates from liquidity disturbances arising from banking panics or from seasonal changes in the demand for credit. In later years, however, the Fed's mandate was broadened to include a wide range of macroeconomic goals. Currently, Section 2A of the Federal Reserve Act instructs the Fed to "maintain long-run growth of the monetary aggregates commensurate with the economy's long-run potential to increase production, so as to promote effectively the goals of maximum employment, stable prices, and moderate long-term interest rates." Moreover, in carrying out monetary policy, the Fed is to "[take] account of past and prospective developments in employment, unemployment, production, investment, real income, productivity, international trade and payments, and prices."

The Fed has a measure of independence within the government in that it makes its month-to-month policy decisions without the direct involvement of Congress, but it is fully accountable to Congress. Accordingly, the Fed reports formally to Congress on monetary policy every six months, and the chairman of the Fed's Board of Governors and other System officials testify before congressional committees on monetary policy issues as well as other matters throughout the year. The body within the Fed that actually formulates and carries out monetary policy is called the Federal Open Market Committee.

It is made up of the seven members of the Board of Governors located in Washington and, at any particular time, five of the twelve regional Federal Reserve Bank presidents. I am a voting member of this committee every third year. I voted last year and will vote again in 1997.

A final point I would make is that the Fed's *policy instrument*—the particular variable it controls on a week-to-week basis to achieve its ultimate policy objectives—is the interest rate on reserves that private banks lend to one another, generally referred to in financial markets as the "federal funds rate." Changes in this rate trigger adjustments in other interest rates, in money and credit flows, and ultimately in broad macroeconomic variables—particularly the aggregate level of prices in the economy.

So the key points about the Fed are (1) that it is a creature of Congress, which has ultimate authority over it; (2) that, as such, it receives its "mandate" from Congress, regarding what it should try to achieve with monetary policy; (3) that the policymaking Federal Open Market Committee has some degree of independence in making its short-run policy decisions, although over time these decisions are subject to congressional review; and (4) that the Fed's policy instrument is the federal funds rate.

2. PREVAILING VIEWS REGARDING MONETARY POLICY IN THE 1960s

With these points about the Fed in mind, let me review policy over the last 25 years, obviously in a very summary fashion. When I began my career at the Fed in 1970, there were three widely held views about the economy that strongly influenced monetary policy and the procedures used to implement it. First, most economists believed that there was a Phillips Curve trade-off between unemployment and inflation in the long run as well as the short run. As you all know, this famous curve summarizes the inverse empirical correlation between unemployment and inflation especially evident in the 1950s and 1960s. The implication for monetary policy, in the eyes of many, was that the Fed could exploit the trade-off the curve seemed to indicate: that is, it could seek a lower level of inflation at the cost of higher unemployment; conversely—and perhaps more to the point—it could seek lower unemployment at the cost of higher inflation.

The second widely held assumption was that economists knew enough about the structure of the economy and the way businesses and consumers behave to permit the Fed to make policy decisions that would eliminate, or at least greatly diminish, the amplitudes of business cycles. This confidence had been fostered by the relatively steady economic growth that had characterized the 1960s and by the neo-Keynesian macroeconomic theories dominant at the time.

The third important idea commonly held in the 1960s was that the welfare costs of inflation were small and that, in any case, they were pretty much limited to the "shoe-leather costs" associated with economizing on money balances in moderately inflationary periods. Of course, there had not been much inflation since the Korean War in the early 1950s, so this belief that inflation was a relatively benign phenomenon probably reflected the absence of any significant recent experience with inflation.

3. INFLATION IN THE 1970s AND ITS EFFECTS

Each of these three views fell victim—largely, if not completely—to developments in the 1970s. During this period there were three major cycles of rising inflation, each more severe than the one before. Each of these accelerations of inflation, in turn, was followed by a sharp tightening in monetary policy and a recession. The most memorable episode occurred in 1979 and 1980, when the Consumer Price Index rose at an annual rate exceeding 12 percent. Confronted with this situation, the Fed took actions that raised short-term interest rates to unprecedented levels, and the worst recession in the postwar period followed, lasting fully six quarters between mid-1981 and the end of 1982.

Sharp increases in oil prices in the mid- and late 1970s no doubt contributed to inflation, but in the long run we know that monetary policy determines the rate of inflation; consequently, inflation could not have risen so sharply over this period without the Fed's acquiescence. There are a number of explanations for the Fed's loss of control over inflation in this period, but in retrospect the breakdown is not terribly surprising. If one combines the notions (1) that the Fed can trade off higher inflation for lower unemployment, (2) that the costs of inflation are small and, moreover, (3) that the Fed has sufficient knowledge about the economy's structure to fine-tune economic activity, it is not difficult to see how the Fed could be led to make monetary policy decisions that had an inflationary bias.

In any case, our experience in the 1970s had a profound impact on conventional thinking about inflation and monetary policy. Most obviously it provided much new data that was, to put it mildly, inconsistent with the Phillips Curve relationship observed in the 1960s. It was in the 1970s, of course, that the term "stagflation" arose to describe a combination of high inflation and *low* growth. In recent years substantial research has been done on the long-run relationship between growth and inflation—much of it based on cross-country data—and I think it is fair to say that on balance there is no compelling evidence that higher inflation is associated with higher growth. Indeed, the research suggests that the relationship may be inverse. The implication, of course, is that inflationary monetary policy is *not* conducive to economic growth; indeed, the opposite may be true.

The 1970s inflation also made people realize that the costs of inflation are much greater and more varied than had been thought earlier. We now understand much better than we did before that inflation creates arbitrary and unfair redistributions of income and wealth that cause social tensions and weaken the fabric of society. Inflation also distorts the signals that prices send in our market economy, which produces serious inefficiencies in the allocation of resources and reduces economic growth. Further, inflation needlessly causes people to spend additional time and energy managing their personal finances. Finally, the 1970s experience illustrated all too well that the public distress caused by rising inflation is inevitably followed by corrective monetary policy actions that depress economic activity, often—as in the early 1980s—severely.

The third consequence of our experience with inflation in the 1970s was a healthy diminution in our confidence that we knew enough about the structure of the economy and the way it functions to fine-tune economic activity and eliminate recessions. As you know, this diminished confidence in our ability to guide economic activity has been mirrored by important developments in monetary theory over the last two decades. I cannot review these developments here in any detail, but most monetary economists now believe that the economy will inevitably be buffeted by various unexpected "shocks" from a variety of directions—such as the energy sector or the stock market—and that it simply is not feasible, and probably not desirable, for the Fed to try systematically to offset the effects of these shocks on the economy. Indeed, we have to be very careful that our efforts to cushion the effects of such shocks do not create rising inflation and thereby exacerbate the economy's problems. As in the practice of medicine, our first responsibility is to do no harm.

I hasten to add that this recognition of the limitations of monetary policy does not relieve the Fed from making short-run policy decisions. And inevitably these decisions will be affected by current developments in the economy. My main point here is that we now realize that these short-run decisions must be consistent with a feasible and credible longer-term policy strategy and that we should not compromise this strategy in a futile attempt to fine-tune the economy.

4. THE IMPACT OF THE 1970s EXPERIENCE ON POLICY PROCEDURES

The 1970s inflation pointed to two fundamental weaknesses in the Fed's overall conduct of monetary policy—weaknesses that to some extent are still present today. First, the System did not have a clear and unambiguous longer-run objective. As inflation accelerated in the mid- and late 1970s, it became apparent that to contain inflation the Fed needed to set targets for some nominal variable that it could control over time and that was clearly linked to inflation over time.

It was in this period that the Federal Open Market Committee first began to set numerical targets for growth rates of the money supply. Initially, the committee set short-run targets for internal use only. Subsequently, in response to a congressional resolution in 1975, the committee began voluntarily to announce quarterly targets for the growth rates of several definitions of the money supply. Finally, the Humphrey-Hawkins Act of 1978 required the Fed to set money-growth targets on a fourth-quarter-to-fourth-quarter basis and to present them formally to Congress. Unfortunately, there was a major flaw in the targeting procedure—commonly referred to as "base drift"—which in fact remains to this day. Base drift occurs because the base level of the money supply used in calculating each new annual target is not the *target* level set for the fourth quarter of the preceding year, but the *actual* level achieved in that period. Therefore, target misses are forgiven when new targets are set, which allows the base level to drift, either upward or downward. In the late 1970s persistent upward base drift led to a prolonged period of unacceptably rapid growth in the monetary aggregates, which in my judgment was a major factor contributing to the subsequent double-digit inflation.

The second weakness in the conduct of policy highlighted by the inflation of the 1970s was the tenuous link between the Fed's month-to-month policy decisions and the emerging longer-run money supply objectives. Under the procedure in place through much of the decade, the Fed was supposed to tighten policy if money growth exceeded the annual target ranges, but the response in any instance was entirely at the Open Market Committee's discretion and, in practice, responses were uncertain and unpredictable. Many economists would agree that the Fed did not react aggressively enough to the persistent above-target growth registered in the latter years of the decade. To deal with this problem, in October 1979 the Fed instituted a new, so-called nonborrowed reserve operating procedure. This procedure was quite complicated in its actual implementation, and I will not try to explain it in any detail tonight. Suffice it to say that the innovation was a monetary policy milestone because for the first time in the Fed's history, its operating procedure caused short-term interest rates to rise *automatically* in response to excessive money growth.

The nonborrowed reserve procedure was abandoned in October 1982, mainly because of increasingly significant practical problems in defining the money supply accurately in a period of rapid technological and institutional change and financial innovation—a problem that has continued to this day. Since then, month-to-month operating decisions have become once again entirely discretionary. I am uncomfortable with this procedure, needless to say, and I will return to this point in a few minutes.

5. DISINFLATION IN THE 1980s: THE IMPORTANCE OF CREDIBILITY

The prolonged recession in the early 1980s was followed by a pronounced disinflation, and by the end of 1983 the inflation rate had fallen to approximately 4 percent, where it remained for the next ten years. In recent years the rate has fallen further to approximately 3 percent. Against the background of these developments, one can say that the decade of the 1980s was a relatively tranquil period for monetary policy—certainly by comparison to the preceding decade. But the more recent period was not without its own lessons for policy. If the 1970s taught the Fed that the costs of inflation are significant and that it must commit itself clearly and fully to a low-inflation policy, the years since have underlined the necessity of maintaining the *credibility* of this policy—by which I mean maintaining the public's confidence that controlling inflation is not a sometime thing but a *permanent* feature of the Fed's overall longer-term monetary strategy.

We now understand more clearly than before the vital role credibility plays in minimizing the cost of reducing inflation and eventually stabilizing the price level. In practical terms, maintaining credibility means the Fed must react promptly to rising inflation expectations. If the Fed's policy actions suggest an indifference to higher expected inflation, the public will lose confidence in its strategy, and workers and firms will demand higher wages and charge higher prices in a perfectly natural effort to protect wages and profits from inflationary erosion. The longer the Fed waits to respond to deteriorating inflation expectations, the more likely it will need eventually to raise real short-term interest rates sharply with potentially depressing effects on business activity. In a nutshell, low credibility makes it more costly from an economic perspective to pursue an anti-inflation strategy.

A few years ago one of my colleagues at the Richmond Fed, Marvin Goodfriend, wrote a widely read article[1] in which he referred to episodes of sharply rising inflation expectations as "inflation scares," and use of that term has now become rather general. Inflation scares can be captured by a variety of financial market indicators, but in my view the most reliable is the long-term bond rate, and this is the indicator I watch most closely to gauge the credibility of our anti-inflation strategy. A sharp rise in long-term rates—as occurred, for example, in the first half of 1994—is a strong signal that inflation expectations have risen and the credibility of our policy has declined, and it is a sign that demands a response from the Fed. The Fed *has*, in fact, reacted to inflation scares more promptly in recent years than earlier, and I believe that this has been one of the hallmarks of recent monetary policy.

[1] Marvin Goodfriend, "Interest Rate Policy and the Inflation Scare Problem: 1979–1992," Federal Reserve Bank of Richmond *Economic Quarterly*, vol. 79 (Winter 1993), pp. 1–24.

6. PRINCIPLES FOR MONETARY POLICY

This completes my review of monetary policy over the last quarter century. I hope it has helped you appreciate why I believe so strongly that the Fed can make its maximum contribution to the economy's growth and productivity by providing a stable price environment in which private individuals, households, and business firms can thrive. For me, the broadest lesson of our experience in the seventies and the eighties is that the overriding goal of monetary policy should be the elimination of inflation, and by that I mean achieving a condition where changes in the general price level are no longer a significant factor in the economic decisions of individuals and businesses.

In this regard, it seems clear that we should not be satisfied with the current 3 to 4 percent inflation rate. One frequently hears the argument that the benefits of achieving price-level stability do not justify the costs. I disagree strongly with this assertion, because I do not believe that a 3 to 4 percent inflation rate could ever be a credible monetary policy objective in the way that price-level stability could. A Fed commitment to aim for 3 or 4 percent inflation—despite its relatively moderate level by recent historical standards—would lack credibility because financial markets and the public quite understandably would fear that eventually the Fed would tolerate higher inflation to achieve some short-term objective. In technical terms, the "time inconsistency" problem in conducting monetary policy, which is one of the most important elements in the recent professional literature on policy, would be much more compelling in a policy regime with a 3 to 4 percent inflation objective than in a regime firmly committed to price stability. This suspicion, in turn, would create uncertainty regarding future inflation, and the attendant increase in risk obviously could harm the economy in a variety of ways.

So my first core belief about monetary policy is that the Fed should remain committed to a policy of eventually achieving true price-level stability and strengthen that commitment in any way it can. My second core belief is that the System needs to maintain the credibility of this policy, which implies—among other things—that its policy procedures and short-run policy actions must be consistent, to the greatest extent possible, with its long-term price stability objective. (I will make some specific points in this regard in a minute.) As I have already noted, by maintaining credibility the Fed can make its anti-inflationary strategy less costly in the transition to price stability and therefore more likely to be successful.

If I have been persuasive this evening, you may think that the two monetary policy principles I have put forward—a policy of price stability and maintenance of the credibility of that policy—are obvious and that there is little left to say. Unfortunately, we still have a substantial distance to go in putting these principles fully into practice. To see that our price stability objective lacks full credibility, one has only to open the newspaper and look at the current level

of the long-term U.S. Treasury bond rate, which is still well over 7 percent. Since it is doubtful that real long-term bond rates ever rise above 4 percent, this means that market participants, on average, currently expect a long-run inflation rate of *at least* 3 percent.

What are the reasons for this lack of credibility? I think there are a number, and I would like to close my remarks tonight by identifying some of them and sharing some ideas about what might be done to deal with them.

As I have indicated before, I believe the most pressing problem the Fed faces in conducting monetary policy currently is the lack of a clear policy mandate from Congress. As I explained earlier, the current mandate contained in the 1978 Humphrey-Hawkins law makes the Fed responsible for a laundry list of economic outcomes having to do with employment, productivity, international trade, and so forth, in addition to the price level. A revised mandate instructing the Fed to focus squarely on achieving price stability almost certainly would enhance the contribution of monetary policy to the nation's long-run economic growth and productivity—indeed, because it would do so, it would *increase*, not reduce, the likelihood that the laudable objectives of the Humphrey-Hawkins law will be achieved.

Five years ago Congressman Steve Neal of North Carolina introduced in Congress an amendment to the Federal Reserve Act proposing just such a mandate. This resolution would have instructed the Federal Open Market Committee to pursue a policy strategy that would "reduce inflation gradually in order to eliminate inflation by not later than 5 years from the date of enactment of [the] legislation and [to] then adopt and pursue monetary policies to maintain price stability." We at the Federal Reserve Bank of Richmond wholeheartedly supported the Neal amendment, as did many others in the Federal Reserve System, as an operationally feasible means of increasing the credibility of the Fed's anti-inflationary strategy. Unfortunately, the amendment did not pass.

Since Congress has not seen fit to pass the amendment, my personal view—and I need to emphasize here that I am speaking strictly for myself—is that the Fed should explicitly and publicly announce that it is adopting the language of the amendment as its longer-term strategic policy goal. In my judgment this step would put the Fed's reputation clearly on the line, which would directly increase the credibility of our strategy. Moreover, as I have already suggested, such a step would be fully consistent with the present Humphrey-Hawkins mandate since price stability would permit the economy to achieve maximum growth in output and employment over time. In this regard, I might note that the value of price stability as a primary monetary policy objective is increasingly recognized around the world. In recent years the central banks in Canada, New Zealand, and the United Kingdom have actually specified explicit *numerical* inflation targets. Since the Neal amendment does not specify numerical targets, its adoption by the Fed would be a step short of these actions abroad. Nonetheless, the amendment's language is sufficiently clear to commit the Fed

firmly to attaining price stability in a specific time frame and hence contains all the ingredients necessary to enhance the System's credibility. Moreover—and this is an especially important point—adoption of the amendment language as its long-term objective would increase the Fed's flexibility in dealing with short-term economic disturbances since appropriate short-term actions could be taken without (or with much less) concern about the potential loss of long-term credibility.

A second area requiring attention is our operating procedures. As I mentioned in my earlier historical review, only in the three-year period from October 1979 to October 1982 has the Fed used an operating procedure that automatically linked movements in our policy instrument—namely the federal funds rate—to a longer-term policy goal, in this case growth rates of the monetary aggregates. As I noted earlier, that procedure was abandoned, largely because of the technical difficulties that arose in defining an operationally reliable measure of the money supply in a period of rapid technological and institutional change—difficulties that unfortunately still confront us. Currently, we still set annual targets for the money supply, but these targets have little effect on our month-to-month policy decisions, which are made pretty much in the same discretionary fashion that characterized the pre-1979 period. This is another important reason why, in my judgment, our credibility is not as full as it could be and should be.

What the Fed needs, in my view, is an operating procedure that clearly links our short-run policy actions directly to our longer-run inflation goals or to some other nominal variable such as nominal gross domestic product. Regrettably, at this point no such procedure exists that commands sufficient confidence to be used in practice. Many economists both inside and outside the Fed are working actively on this problem, however, and I have confidence that somewhere down the road we will come up with an acceptable operating procedure that more systematically and efficiently links our instrument to our goals. In the meantime, the Fed must retain the independence to take the short-run policy actions that it believes are most likely to be consistent with its long-run objectives—recognizing, of course, that it is responsible for and accountable for the consequences of these decisions.

A final and very important point I would make is that the Fed has a strong obligation to educate the public about the cost of inflation and the limitations of activist short-term monetary policies. In my review, I explained how the inflation of the 1970s led me and many others to conclude that some of the views regarding inflation and monetary policy in the 1960s were not valid. Unfortunately, in my opinion, many people still believe that a long-run as well as a short-run trade-off between inflation and unemployment exists, that the costs of inflation are small, and that the Fed can fine-tune economic activity. The persistence of these views—particularly when they are held by people with political power—naturally diminishes the credibility of our anti-inflation

strategy, especially given that our mandate is so imprecise. It would be a tragedy if the lessons of the last 25 years were forgotten and the nation needlessly experienced another devastating boom-bust cycle like the one in the 1979–82 period. So I think we in the Fed have an obligation to speak out on these issues. My remarks here tonight have been an effort in that direction, and I hope that I have added at least a bit to your appreciation of some of the fundamental issues facing monetary policymakers today.

Neighborhoods and Banking

Jeffrey M. Lacker

The economic condition of some of our low-income neighborhoods is appalling. Are banks responsible? Critics blame the banking industry for failing to meet the credit needs of poorer neighborhoods. Some claim that bankers pass up worthwhile lending opportunities because of racial or ethnic bias. Others argue that a market failure causes banks to restrain lending in low-income neighborhoods. They claim that joint lending efforts by many banks in such neighborhoods would be profitable, but no single bank is willing to bear the cost of being the pioneer.

The central statute regulating the relationship between bank lending and neighborhoods, the Community Reinvestment Act of 1977 (CRA, or "the Act"), was inspired by the critics' view that banks discriminate against low-income communities.[1] The Act directs the bank regulatory agencies to assess the extent to which a bank meets "the credit needs of its entire community, including low- and moderate-income neighborhoods." In a similar spirit, the Home Mortgage Disclosure Act (HMDA) requires depository institutions to disclose mortgage originations in metropolitan areas by census tract. The annual HMDA reports routinely show large disparities in mortgage flows to minority and white neighborhoods, bolstering the critics' case.

Defenders of the banking industry attribute the disparity in credit flows to differences in the creditworthiness of potential borrowers, information that is

This article originally appeared in this Bank's *1994 Annual Report*. The author has benefited from comments by Marvin Goodfriend, Tom Humphrey, Tony Kuprianov, Stacey Schreft, and John Weinberg. The views expressed are the author's alone and do not necessarily reflect the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] I will use the term "banks" throughout to refer to commercial banks and thrifts. Credit unions are currently exempt from the CRA.

unavailable from the HMDA reports. They view the CRA as a burdensome interference in otherwise well-functioning credit markets and as a regulatory tax on banking activity. They argue that the decay of low-income neighborhoods, while deplorable, is beyond the capacity of the banking industry alone to repair.[2]

The CRA is currently attracting renewed attention. Public release of expanded HMDA reports, along with widely publicized research suggesting bank lending discrimination, has sparked complaints that banks neglect low-income neighborhoods. Critics now assert that regulators have been too lax in implementing the CRA, and they press for regulations based on measures of bank lending in low-income neighborhoods. In response, federal banking agencies recently adopted revisions to the regulations implementing the CRA that would base a bank's assessment in part on quantitative measures of lending in low-income neighborhoods (Board of Governors of the Federal Reserve System 1994). Banks' defenders argue that the regulations were already too burdensome and that numerical measures inevitably will come to resemble lending quotas. Banks will be induced to make loans to uncreditworthy borrowers, risking losses to the deposit insurance funds and, ultimately, to taxpayers.

This essay reexamines the rationale for the CRA. A reconsideration seems worthwhile in light of the dire condition of our poor neighborhoods on the one hand, and the demonstrable risks to banks and taxpayers on the other. After a review of the empirical literature relevant to critics' claims, I will argue that there is little conclusive evidence that banks fail to meet the credit needs of low-income neighborhoods per se. Instead, the CRA regulations should be understood as a transfer program, aimed at redistributing resources to low-income neighborhoods. The basic goal of the CRA to improve conditions in distressed neighborhoods is obviously a worthy one. But the lending and community investment obligations impose an implicit tax on the banking industry for which there is little justification. Nonprofit community development organizations (CDOs) also redistribute resources through subsidized lending in low-income neighborhoods and represent an alternative to imposing a potentially unsustainable burden on banks. Directing investment toward low-income neighborhoods could be better accomplished by carefully subsidizing existing institutions that specialize in community development, rather than by imposing a burdensome and potentially risky implicit tax on the banking system.

[2] I will use throughout the essay the less cumbersome term "low-income neighborhoods" to refer to the low- and moderate-income neighborhoods that are the focus of the CRA. The newly proposed CRA regulations define low-income neighborhoods as census tracts with median household income less than 50 percent of the median household income of the metropolitan statistical area (MSA). Moderate-income neighborhoods are defined as census tracts with median household income between 50 and 80 percent of the median household income of the MSA.

1. DO BANKS REDLINE?

The legislative history of the Community Reinvestment Act makes clear that the Act was based on the premise that banks engage in "redlining." Senator William Proxmire, principal sponsor of the CRA, defined redlining during debate on the Senate floor:

> By redlining . . . I am talking about the fact that banks and savings and loans will take their deposits from a community and instead of reinvesting them in that community, they will invest them elsewhere, and they will actually or figuratively draw a red line on a map around the areas of their city, sometimes in the inner city, sometimes in the older neighborhoods, sometimes ethnic and sometimes black, but often encompassing a great area of their neighborhood.[3]

The term "redlining" dates back to the 1930s, when the Home Owners Loan Corporation (HOLC) and the Federal Housing Administration (FHA) used detailed demographic and survey analysis to classify city neighborhoods for lending risk.[4] The agencies adopted standardized appraisal and underwriting practices that embodied the common real estate practice of the time of rating neighborhoods in part on the basis of their current and prospective racial and ethnic composition (Jackson 1985). Blocks with the lowest of four grades were color-coded red on secret maps. A 1939 FHA *Underwriting Manual* warned that "if a neighborhood is to retain stability, it is necessary that properties shall continue to be occupied by the same social and racial classes."[5] While government agencies retreated from explicitly racial policies after the 1948 U.S. Supreme Court decision against racial deed covenants, neighborhood racial composition apparently continued to affect appraisals into the 1970s.[6]

As evidence of continuing redlining, legislators cited the results of numerous studies in the early 1970s by community groups and local governments. The availability of HMDA data in the mid-1970s spurred further redlining research in the academic and policy communities. Although critics often cite discrimination against older or lower-income neighborhoods, research has addressed almost exclusively redlining on the basis of a neighborhood's racial composition. The studies documented large disparities in mortgage lending activity, which led critics of banks to conclude that they had unfairly restricted

---

[3] *Congressional Record*, daily ed., June 6, 1977, S. 8958, cited in Dennis (1978). Senator Proxmire's definition of redlining also reflects the doctrine of localism in banking—the idea that the savings of a community should be invested locally rather than where returns are highest. See Macey and Miller (1993) for a critique.

[4] See Woelfel (1994) for a description of the HOLC and the FHA.

[5] Quoted in Jackson (1985), p. 207.

[6] In 1977 the American Institute of Real Estate Appraisers removed discriminatory racial references from their textbook as part of an agreement settling a federal lawsuit. See Art (1987), p. 1078.

loan supply in predominantly minority neighborhoods and thus had failed to serve the credit needs of their communities.[7]

This first-generation research failed to show, however, that supply rather than demand was responsible for the lending disparities. A basic premise of the redlining hypothesis is that banks curtail the supply of credit to a neighborhood for noneconomic reasons such as racial composition. Many factors that influence the demand for mortgage credit by qualified borrowers also vary across neighborhoods: income and wealth levels, owner-occupancy rates, and housing turnover rates, for example. Moreover, many of these factors are known to be correlated with the racial composition of a neighborhood. Without controlling for differences in the demand for credit, there is little one can say about constraints on the supply of credit to minority neighborhoods.

Subsequent redlining research sought to remedy this problem using information on the economic characteristics of neighborhoods and individual loan applicants. When such information is taken into account, mortgage flows and loan approval rates appear unrelated to neighborhood racial composition. For example, Schill and Wachter (1993) estimate models of banks' loan approval decisions. In their simplest model, the neighborhood racial composition is significantly related to approval probability, but when neighborhood characteristics such as median income, vacancy rate, and age of the housing stock are included, neighborhood racial composition is no longer important. Similarly, Canner, Gabriel, and Woolley (1991) find that after controlling for individual and neighborhood measures of default risk, there is no evidence of discrimination based on the racial composition of neighborhoods. Several other studies confirm these findings.[8] Research thus has failed to uncover any evidence that banks discriminate against neighborhoods on the basis of racial composition.[9]

[7] See Canner (1982) and Benston (1979) for surveys.

[8] See Avery and Buynak (1981), Holmes and Horvitz (1994), King (1980), Munnell et al. (undated), and Schill and Wachter (1994). Some studies have reported evidence of redlining, but in these the controls for individual characteristics are limited or absent. Bradbury, Case, and Dunham (1989) use data at the neighborhood level, but they employ a problematic credit flow variable that includes commercial as well as residential transactions. They do not control for individual economic characteristics. Calem and Stutzer (1994) also use neighborhood-level data, and so do not control for individual economic characteristics. Avery, Beeson, and Sniderman (1993) rely on HMDA data and census tract information, and so are unable to control for applicant wealth or creditworthiness. Although it is conceivable that future research will turn up evidence of redlining, it seems unlikely; the fact that studies with better controls for individual economic characteristics obtain smaller or negligible estimates of the effect of racial composition on mortgage outcomes suggests that the estimates we have are biased upward.

[9] Critics also have charged that banks redline older and lower-income neighborhoods (see Art [1987], for example), but age of the housing stock and borrower income are both plausibly related to lending risk. As a result, statistical research of the type referred to above is unable to distinguish between legitimate underwriting practices and redlining these neighborhoods. I am unaware of any attempt to disentangle the two.

2. DO BANKS DISCRIMINATE AGAINST INDIVIDUALS?

Redlining is distinct from racial discrimination against individuals because not all minority applicants live in redlined neighborhoods.[10] Although research has found little evidence of discrimination against minority *neighborhoods*, recent research has uncovered evidence consistent with discrimination against individual minority loan *applicants*. The most widely publicized evidence comes from the HMDA data. In 1989, Congress amended the Home Mortgage Disclosure Act to require lenders to report the disposition of every mortgage loan application, along with the race or national origin, gender, and annual income of each applicant. Numerous press reports have focused on the disparities between whites and minorities in the fraction of applicants denied credit. For example, the 1993 data show that for conventional home purchase loans, 34 percent of African-American applicants and 25 percent of Hispanic applicants were denied credit, while only 15 percent of white applicants were denied credit (Federal Financial Institutions Examination Council 1994).

By themselves, however, simple tabulations of HMDA data are inconclusive for the same reason that raw mortgage flow data are misleading. The HMDA data report applicant income, but not credit history or other economic characteristics. Without controlling for applicant creditworthiness, the disparity in mortgage loan denial rates in the HMDA data could reflect the disadvantaged economic status of minorities, rather than noneconomic discrimination by banks. It is well known that racial and ethnic groups differ significantly on many dimensions of creditworthiness. For example, the average minority individual has lower income, lower net worth, and lower financial asset holdings than does the average white American. Furthermore, minority mortgage applicants are more likely to have adverse credit histories and to request larger loans relative to property value, factors associated with higher default risk.[11] In short, differentiating between racial discrimination and racial disparities in creditworthiness is difficult.

[10] For example, in 1992, 39.2 percent of minority individuals lived outside of census tracts in which over half of the population was minority (derived from Canner, Passmore, and Smith [1994]).

[11] On racial disparities in income and economic status, see, for example, Kennickell and Shack-Marquez (1992), Jaynes and Williams (1989), or the Symposium in the Fall 1990 issue of the *Journal of Economic Perspectives*. Munnell et al. (1992) report that loan-to-value ratios and adverse credit history variables are higher for minority applicants; see also Carr and Megbolugbe (1993). Canner and Luckett (1990) report that households headed by a minority are significantly more likely to have missed a debt payment, even after controlling for other household characteristics.

3. THE BOSTON FED STUDY

A recent study by economists at the Federal Reserve Bank of Boston has gone the farthest toward solving this problem.[12] They asked banks and mortgage companies for detailed information from the loan applicant files for a sample of Boston HMDA data for 1990. They obtained data on housing expenses, total debt payments, net wealth, credit and mortgage payment histories, appraised property values, whether properties were single- or multi-family dwellings, whether applicants were self-employed, and whether applicants were denied private mortgage insurance. Combining this information for a sample of 3,062 individual applicants with applicant race and the unemployment rate in the applicant's industry, they estimated the probability of a particular mortgage loan application being denied.

The study's major finding is that after controlling for the financial, employment and credit history variables they were able to observe, race still had a highly significant effect on the probability of denial. The results imply that minority individuals with the characteristics of an average white applicant have a 17 percent denial rate compared to an 11 percent denial rate for white applicants with the same characteristics. Moreover, the Boston Fed study suggests that whatever discrimination takes place is of a subtle form. Whereas applicants of all races with unblemished credentials were almost certain to be approved, the study found that the vast majority of applicants had some imperfection. As a result, lenders have considerable discretion to consider compensating factors in evaluating creditworthiness. The Boston Fed researchers suggest that "lenders seem more willing to overlook flaws for white applicants than for minority applicants" (Munnell et al. 1992, p. 3).

These findings are consistent with the widely held view that lending discrimination is common in housing markets. A recent survey found that 69 percent of African Americans and 33 percent of whites do not feel that African Americans have an equal opportunity for credit loans and mortgages. Housing discrimination also has been the focus of housing market audit studies, in which matched pairs of testers, one white and one minority, respond to advertisements for rental or sales properties. Such studies have found evidence of differential treatment based on race, such as African Americans not being shown certain available properties. The few pilot studies on home mortgage lending discrimination at the pre-application stage are too small to be conclusive. Anecdotal reports of lending discrimination are sometimes cited as well.[13]

---

[12] See Munnell et al. (1992).

[13] The survey data are from National Conference (1994). On audit studies in housing, see Fix and Struyk (1993), but particularly the critique by Heckman and Siegelman (1993). Cloud and Galster (1993) survey home mortgage lending audit studies, along with anecdotal reports of lending discrimination. The application of audit methodology to lending discrimination is inhibited

4. INTERPRETING THE BOSTON FED RESULTS

Although anecdotes and evidence from audit studies are suggestive, the Boston Fed study remains the most rigorous evidence available of home mortgage lending discrimination.[14] Despite the study's sophistication, however, considerable uncertainty remains concerning its interpretation. Some researchers have questioned the reliability of the data and the empirical model underlying the study.[15] Although the critiques are far from definitive, replication of the study's results using different data sets obviously would increase confidence in its findings. Seldom is a single retrospective study taken as conclusive, particularly in the social sciences, and the Boston Fed study is the only research on lending discrimination that explicitly controls for individual applicants' credit history. Further research would be especially valuable in view of the plausible alternative hypotheses that are consistent with the Boston Fed results.

One such alternative view is that the variables in the study measuring creditworthiness are imprecise or incomplete and fail to capture completely the judgment of a hypothetical unbiased loan officer. If there is any random discrepancy between applicants' true creditworthiness and their creditworthiness as measured by model variables, there is likely to be a bias in measuring discrimination. When true creditworthiness is inaccurately measured, it is very difficult to distinguish racial discrimination from unmeasured racial disparities

---

by laws prohibiting applying for a mortgage under false pretenses. Audit methodology is thus limited to the more subjective problem of differential treatment at the pre-application stage.

[14] Several redlining studies examined data for outcomes of individual mortgage applications. Some found that minority applicants were less likely than whites to obtain a mortgage loan, even after controlling for neighborhood economic characteristics. See Avery, Beeson, and Sniderman (1993), Shafer and Ladd (1981), Canner, Gabriel, and Woolley (1991), and Schill and Wachter (1993, 1994). None of these studies controlled for applicant credit history, and so they suffer from the same omitted-variable problem that plagues the analysis of the HMDA data. In related research, Hawley and Fujii (1991), Gabriel and Rosenthal (1991), and Duca and Rosenthal (1993), using data from the 1983 Survey of Consumer Finances, find that after controlling for individual characteristics, minorities are more likely than whites to report having been turned down for credit. Information on individual creditworthiness was quite limited, however, again leaving these studies vulnerable to the omitted-variable problem.

[15] Horne (1994) reports on reexaminations of some of the loan files at the FDIC institutions participating in the study. Although he reports a large number of data errors, he does not reestimate the model, so no conclusion is possible about the effect of those errors. In addition, files were selected for reexamination in a way that would bias any reestimation. Liebowitz (1993) claims in an editorial page essay in *The Wall Street Journal* that correcting selected data-coding errors eliminates the finding of discrimination, but Carr and Megbolugbe (1993) and Glennon and Stengel (1994) document that the discrimination finding persists after systematic data-cleaning, suggesting bias in the way Liebowitz corrects errors. See also Browne (1993a). Zandi (1993) claims that omission of a variable assessing whether the institution reports that the applicant met their credit guidelines was responsible for the estimated race effect. Carr and Megbolugbe (1993) confirm that including this variable reduces the estimated race effect somewhat, but note that this subjective assessment by the lending institution is significantly related to an applicant's race, even after controlling for the objective economic characteristics of the applicant. See also Browne (1993b). Schill and Wachter (1994) also study the Boston Fed data set.

in creditworthiness. If true creditworthiness is associated with applicant race, the model will indicate that race affects the probability of denial, even if race plays no direct causal role. If true creditworthiness is lower on average for minority applicants, then there will be a bias toward finding discrimination against minorities.[16]

The fact that measured creditworthiness is statistically associated with race suggests that this condition holds. Regulatory field examiners report that it is often difficult to find matched pairs of loan files corroborating discrimination detected by a statistical model or summary statistics. Examination of applicant files often reveals explanatory considerations that are not captured by any model variables. The credit history variables in the Boston Fed study are simple functions of the number of missed payments or whether the applicant ever declared bankruptcy, and do not reflect the reasons for any delinquencies. Evaluating explanations of past delinquencies is at the heart of credit underwriting; some will indicate poor financial management skills or unstable earnings, while others will reflect response to unusual one-time financial shocks or inaccurate credit reports. It seems quite plausible, therefore, that the Boston Fed findings are an artifact of our inability to capture complex credit history information in a tractable quantitative representation.

Another hypothesis consistent with the evidence from the Boston Fed study is that minority borrowers are more likely to default than equally qualified white borrowers, so lenders implicitly use race as an indicator of creditworthiness in marginal cases, above and beyond the information provided by income, balance sheets, or credit history. Such behavior, often called "statistical discrimination," might be economically rational, though still illegal. The statistical discrimination and measurement error hypotheses are closely related because both assume that the outside analyst does not observe true creditworthiness. The distinction is that under the measurement error hypothesis the loan officer observes true creditworthiness, while under the statistical discrimination hypothesis the loan officer does not directly observe credit quality but uses race as a proxy.

A recent study of mortgage default data supports these alternative explanations. The study found that an African-American borrower is more likely to default than a white borrower, even after controlling for income, wealth, and other observable borrower characteristics.[17] Why would a minority borrower

[16] If a true explanatory variable is measured with noise, its regression coefficient will be biased toward zero. In that case, any other variable correlated with the true explanatory variable will be significant in the regression, even though it may play no direct causal role in explaining the behavior in question. Thus, measurement error is a very serious problem in statistical inference. See Johnston (1963) for a discussion of measurement error and Cain (1986) for a discussion of the implications for detecting discrimination.

[17] Berkovec et al. (1994) use data on more than a quarter of a million FHA mortgages originated during 1987–1989. Their data do not include information on the borrower's credit history, but they estimate that including credit history would reduce the estimate of the race effect

be more likely to default than an equally qualified white borrower? Mortgage defaults often are attributable to "trigger events," such as involuntary job loss or large unexpected health care costs, that sharply reduce the borrower's ability to repay.[18] Most people are poorly insured against such risks, and it seems plausible that minorities experience these events more often than whites.[19] For example, unemployment rates are higher for minorities than for whites, but more important, the probability of involuntary job loss is higher for minorities (Jackson and Montgomery 1986; Blau and Kahn 1981; Flanagan 1978). Minority household holdings of financial assets are far smaller on average, reducing their ability to withstand uninsured financial shocks (Kennickell and Shack-Marquez 1992). Minorities tend to be less healthy on average and are more likely to lack health insurance (National Center for Health Statistics 1994). There seems to be no research on whether these differences in the likelihood of trigger events persist after controlling for income, wealth, credit history, and other factors observable at the time of the application. But it seems plausible that these risk factors can explain the disparity in mortgage default rates and can thereby account for disparities in loan approval rates. This line of reasoning suggests that disparities outside lending markets—in labor markets, for example—might well be responsible for what appears to be lending discrimination.[20]

One other consideration that lends support to these alternative explanations of the Boston Fed results is the presumption that competitive forces should act to eliminate unprofitable discriminatory practices. If some lenders discriminate on noneconomic grounds, they ought to systematically lose business over time as long as there are some lenders that do not discriminate. The discriminatory lenders may end up serving only part of the market, but nondiscriminatory lenders would be eager to fill the void.[21]

To summarize, the empirical evidence on bank lending discrimination based on an applicant's race seems inconclusive. A skeptic with a strong prior belief in the ability of market forces to restrain unprofitable discrimination could easily remain unconvinced by the Boston Fed study. On the other hand, critics

---

by only 30 percent. Barth, Cordes, and Yezer (1979, 1983) and Evans, Maris, and Weinstein (1985) also show that race is significantly related to default probabilities, but the omission of important variables weakens the interpretation of their results.

[18] See Quercia and Stegman (1992) for a review of recent literature on mortgage default.

[19] Cochrane (1991) reports evidence that households are poorly insured against involuntary job loss and long-term absences due to illness.

[20] See Jaynes (1990) for a survey of the labor market status of African Americans.

[21] Calomiris, Kahn, and Longhofer (1994) suggest that a lack of "cultural affinity" between white loan officers and minority applicants may explain findings of discrimination. A lack of affinity might reduce the reliability and accuracy of loan officers' subjective evaluations, leading to higher standards for African Americans at predominantly white banks. The cultural affinity hypothesis, however, has trouble explaining the higher rejection rates found at minority-owned banks.

with a strong prior belief in the prevalence of lending discrimination will find striking confirmation in the Boston Fed study. Between these two extremes lies a range of reasonable assessments.[22]

What does the empirical evidence on discrimination, such as it is, imply about appropriate public policy? Discrimination against mortgage applicants on the basis of an individual's race calls for vigorous enforcement of fair-lending laws. However, the lack of evidence of discrimination against neighborhoods per se raises questions about the need for a lending obligation aimed at neighborhoods. Not all minority applicants have low incomes or live in low-income neighborhoods, so the connection between racial discrimination against individuals and lending to low-income neighborhoods is doubly obscure. The evidence that we do have, which suggests the possibility of racial discrimination against *individuals* but not neighborhoods, provides little reason for a law like the CRA that targets lending to low-income *neighborhoods*.[23]

5. IS THERE SOME OTHER SOURCE OF MARKET FAILURE?

Lacking evidence of bank discrimination against neighborhoods, is there some other rationale for a government-imposed lending obligation? Could CRA-induced lending be socially desirable even though banks would otherwise find it unprofitable? In other words, is there a market failure affecting lending in low-income neighborhoods?[24]

Many writers have pointed out that low-income housing markets are frequently characterized by "spillover effects" because the physical condition and appearance of one property affects the desirability of nearby properties. This leads to a strategic interaction among property owners; improvements to a house in a well-maintained block are worthwhile but would have little value if the rest of the block is poorly maintained or vacant. A run-down neighborhood might be worth renovating from society's point of view, yet no single property owner has an incentive to invest. This strategic interaction extends to potential lenders as well. Each bank judges an applicant in isolation, ignoring the effect on nearby properties. Taking the poor condition of neighboring properties as

[22] Public policy toward neighborhoods and banking could be aided greatly by research on the root cause of mortgage defaults: Why is it that trigger events such as health problems or involuntary job loss are so poorly insured? Such research might allow us to distinguish between competing explanations of disparities in credit flows across neighborhoods and ethnic groups. Furthermore, we might find that reducing disparities in the incidence of trigger events is more effective than affirmative lending obligations that encourage banks to ignore such disparities.

[23] For a similar view, see the Shadow Financial Regulatory Committee (1994).

[24] Market failure can occur in situations with spillover effects, since one person does not have to pay for the effect of their decision on the well-being of others, as when polluters do not pay for the damage caused by their emissions.

given, the loan might appear to be a poor risk, even though simultaneous loans to improve all properties might be worthwhile. All would be better off if lenders could coordinate their decisions and agree to lend, since those loans would be profitable. But in the absence of coordination, each bank's reluctance to lend confirms other banks' reluctance to lend and becomes a self-fulfilling prophecy of neighborhood decline. In these circumstances, a genuine market failure could be said to occur.[25]

Spillovers seem quite important in affluent residential and commercial markets as well. The preeminence of location in valuing suburban homes epitomizes the importance many homebuyers place on the characteristics of the surrounding neighborhood. Office buildings often are clustered to take advantage of common services or homogeneity of appearance. What is striking about spillovers in more affluent real estate markets is that they do not seem to cause any serious market failure; private market mechanisms seem quite capable of coordinating investment decisions. For example, suburban housing is often developed in large parcels of very similar homes, ensuring the first buyers that subsequent investment will not blemish the neighborhood. The development is coordinated by a single entity that either builds all the homes or enforces homogeneity through building restrictions and deed covenants.

From this perspective, it is hard to see just what would impede similar market mechanisms in low-income neighborhoods. A substantial part of the economic role of a real estate developer is to coordinate investment decisions, internalizing the spillovers inherent in closely neighboring investments. If a coordinated investment in a low-income neighborhood is in society's best interest, why wouldn't a private developer assemble the capital to finance the investment?

Several notable differences between the suburbs and low-income, inner-city neighborhoods might explain why coordinating investments is more difficult or costly in city neighborhoods. Low-income urban neighborhoods tend to be older, higher-density areas, while development in the suburbs is often on virgin tracts of undeveloped land. Assembling control of the requisite property rights is arguably less costly for the latter. Another factor affecting the ease of assembling property rights is the higher incidence in the cities of governmental encumbrances such as rent controls or tax liens. The greater incidence of crime in urban areas also inhibits development by making it more costly to provide residents with a given level of security.

Disparities between urban and suburban markets in the costs of coordinating investments, however, do not necessarily provide a rationale for government stimulus of low-income community development. The expense of keeping crime out of a neighborhood, for example, is a real social cost that deserves to

[25] Guttentag and Wachter (1980) present this externality argument.

be addressed directly, and there is no reason to encourage people to ignore it in their investment decisions. Similarly, government restrictions on property rights distort decisions, although usually with the aim of benefiting some particular group. These distortions impose genuine costs, and it is hard to see why we should encourage people, including lenders, to discount them. In sum, these very real costs do not, by themselves, represent a market failure.

Lang and Nakamura (1993) describe a more subtle type of spillover. The precision of appraisals, they argue, depends on the frequency of previous home-sale transactions in the neighborhood. A low rate of transactions makes appraisals imprecise, which increases mortgage lending risk in the neighborhood, reducing mortgage supply, and thereby reducing the frequency of transactions. A neighborhood can get stuck in a self-reinforcing condition of restricted mortgage lending and low housing stock turnover. The key impediment to efficiency in this story is the failure of lenders and homebuyers to account for the social benefit of their transaction on others' ability to make accurate appraisals in the future.

While this argument seems theoretically plausible, some important problems remain. For example, it is not clear what limits this phenomenon to low-income neighborhoods. Affluent housing markets are quite prone to transitory declines in transactions volume, but rarely seem to get stuck in a depressed condition. And again, it is not clear why market mechanisms would be unable to coordinate transactions in low-income neighborhoods as they do in many other real estate markets. On the whole, then, it seems difficult to argue that lending in low-income neighborhoods is any more beset by market failures than lending in affluent neighborhoods.

6. IS REDISTRIBUTION THE PURPOSE OF THE CRA?

If the CRA cannot be rationalized as a corrective for lending discrimination or some other identifiable market failure, then the CRA must be essentially a redistributive program that should be justified by equity rather than efficiency considerations. Indeed, the desire to simply transfer resources to low-income neighborhoods is understandable in view of their appalling condition. But how should such a transfer be carried out?

The CRA has been likened to a tax on conventional banking linked to a subsidy to lending in low-income neighborhoods (White 1993; Macey and Miller 1993). Although banks are examined regularly for compliance with CRA regulations and receive public CRA ratings, enforcement relies on the power of the regulatory agencies to delay or deny an application for permission to merge with or acquire another bank or to open a new branch. The prospect of having an application delayed or denied, along with the public relations value of a high

CRA rating, provides banks with a tangible incentive for CRA compliance.[26] According to this view, by tilting banks' profit-loss calculations, the CRA regulations give banks an incentive to make loans they would not otherwise have made. To the extent that banks are induced to make loans and investments they would not otherwise have found profitable, the CRA regulations encourage banks to subsidize lending in low-income neighborhoods. Investments at concessionary rates and CRA-related outlays, such as for marketing programs and philanthropic contributions, directly reduce a bank's net earnings. The gap between the cost of these loans to borrowers and what they would have cost in the absence of the CRA represents a transfer to the low-income neighborhood.

Two questions naturally arise, though, if the CRA is viewed as a redistributive program. First, why should we provide low-income neighborhoods with an enhanced credit supply rather than unencumbered cash payments? Second, why should the banking industry be the source for such transfers?

7. WHY SUBSIDIZE LENDING IN LOW-INCOME NEIGHBORHOODS?

If the goal is to make the residents of low-income neighborhoods better off, why not provide unrestricted transfer payments? Economists generally argue that unrestricted income transfers are more efficient than equally costly transfers tied to particular goods or services. This efficiency arises from the expanded choices available to recipients. Community development subsidies via enhanced mortgage lending, in contrast, tie transfers to borrowing and homeownership. Why encourage low-income households to take on more debt? And why should subsidies to residents of low-income neighborhoods be tied to their ownership of housing?

A plausible argument can be made for targeting subsidies to low-income homeowners as a way to rectify the baneful housing and lending policies of the past. A variety of explicit policies at both public and private institutions in the first half of this century encouraged the flight of white middle-class residents from inner cities to the suburbs. Metropolitan real estate boards adopted explicitly racial appraisal standards and attempted to prevent members from integrating neighborhoods (Helper 1969). The FHA provided a significant stimulus to homeownership, but agency underwriting policies and housing standards strongly favored newly constructed homes in all-white suburbs (Jackson 1985). It recommended racially restrictive deed covenants on properties it insured until the Supreme Court ruled them unenforceable in 1948. The banking industry

[26] The recent proposed revision to the regulations implementing the CRA would allow regulators to seek enforcement action in cases of "substantial noncompliance," the lowest possible CRA rating.

apparently adopted similar underwriting policies (Jackson 1985, p. 203). Some researchers cite these policies as important in the creation and persistence of racial segregation and the concentration of poverty in the inner cities.[27]

This rationale for the CRA invokes the notion of corrective justice, the normative idea that compensation should be made for past inequities.[28] The discriminatory practices of earlier times depressed the welfare of low-income minority communities by raising the cost of home mortgages there relative to more affluent suburban communities, although the lack of evidence of redlining in recent years suggests that noneconomic cost differentials have largely been removed. Subsidies that lower the cost of home mortgage lending in low-income minority communities—in contrast to unrestricted cash payments—transfer resources to precisely the same groups that the earlier discriminatory policies transferred resources *from*—nearly creditworthy low-income homeowners. As Peter Swire (1994) notes, "Only a very small subset of the effects of discrimination [in housing markets] can be traced with enough specificity to permit legal redress" (p. 95). Thus, it may be quite difficult to target unrestricted income transfers to individuals directly harmed by past discriminatory practices. Mortgage lending subsidies that mirror the implicit tax of historic home lending discrimination might be the most efficient way of compensating those who were harmed.

8. SHOULD BANKS SUBSIDIZE LENDING?

Why should depository institutions be singled out for the affirmative obligation imposed by CRA regulations? Why do other lending intermediaries such as mortgage, finance, and life insurance companies escape obligation? More broadly, why should financial intermediaries bear the burden rather than society as a whole? Senator Proxmire provided a partial answer when introducing the original Act by noting that a bank charter "conveys numerous benefits and it is fair for the public to ask something in return."[29] The CRA, in this view, is a quid pro quo for the special privileges conferred by a bank charter, which incidentally explains why the Act links assessment to a bank's "application for a deposit facility." To the extent that CRA obligations are unprofitable or are equivalent to charitable contributions, apparently they are to be cross-subsidized from the stream of excess profits otherwise generated by the bank charter.

---

[27] See, for example, Wilson (1987) or Massey and Denton (1993). Homeowner preferences apparently play a role as well.

[28] In a paper devoted to legal and economic analysis of the CRA, Swire (1994) discusses corrective justice as a "noneconomic" rationale for the CRA. He also discusses "distributive justice," which would also rationalize transfers but would not necessarily suggest they take the form of subsidized lending.

[29] U.S. Congress (1977), p. 1. See also Fishbein (1993).

The difficulty with this role for the CRA is that cross-subsidization may be infeasible (White 1993). The competitive environment facing banks has changed greatly since passage of the CRA in 1977. Over the last two decades the legal and regulatory restrictions on competition among banks have been substantially reduced, a trend that will continue with the implementation of the Interstate Banking Efficiency Act of 1994. Perhaps more important, rapid changes in financial technology are eroding the advantages of banks relative to nonbank competitors. Consequently, imposing a unique burden on the banking industry might only diminish banks' share of intermediated lending. The regulatory burden ultimately would fall on bank-dependent borrowers in the form of higher loan rates and on bank-dependent savers in the form of lower deposit rates. And to the extent that lending induced by the CRA regulations increases the risk exposure of the deposit insurance funds, taxpayers who ultimately back those funds bear some of the burden as well.

Senator Proxmire suggested a practical reason banks are asked to shoulder the CRA burden when he remarked that "there is no way the Federal Government can solve the problem [of revitalizing the inner cities] with its resources."[30] From this perspective, the CRA imposes a tax on banks to avoid an explicit general tax increase. But a general tax increase is usually less costly to society than an equal-sized tax on a single industry because spreading the burden over a wider base minimizes the resulting distortions in economic activity. From this perspective, imposing the CRA tax on banks rather than the economy as a whole involves an excess social cost.

Compelling banks to provide subsidized lending in low-income neighborhoods might be warranted nevertheless if banks have a unique comparative advantage in doing so. The cost savings from such a comparative advantage might justify incurring the excess social cost of the CRA burden on banks. But if no comparative advantage can be identified, we ought to consider alternative means of providing subsidized lending that avoid the excess cost of a tax levied solely on banks.

9. COMMUNITY DEVELOPMENT ORGANIZATIONS PROVIDE SUBSIDIZED LENDING

Community development organizations (CDOs) are institutions that promote investment in target neighborhoods, working closely with homebuyers, private lenders, businesses, government agencies, and private donors.[31] They primarily

[30] See U.S. Congress (1988), p. 7.

[31] See Wells and Jackson (1993) for a survey of community development lending, and see Board of Governors of the Federal Reserve System (1993) for a survey of community development lending by banks.

arrange loans for development projects and homeowners, and their costs are generally funded by grants and donations. Their goal of revitalizing decaying neighborhoods matches exactly the avowed purpose of the CRA. CDOs represent an alternative to channeling subsidized lending through the banking system.

Neighborhood Housing Services of Baltimore (NHSB) is one such organization.[32] The main focus of the NHSB is promoting occupant homeownership, improving the physical appearance of neighborhoods, and "stabilizing" the real estate market. The NHSB has targeted four different Baltimore neighborhoods since its inception in 1974. Within a neighborhood, it often targets particular blocks by systematically searching for owner-occupants for each property on the block. When it finds a suitable buyer for a property, NHSB often arranges for extensive renovations, handles the design and bidding, and selects a contractor.

A great deal of the work of NHSB involves lending. It provides extensive education and counseling to help prospective borrowers qualify for loans. This assistance can involve establishing bank accounts, repairing credit records, documenting sources of income, learning about home purchase and mortgage application procedures, and saving for a down payment. Qualification often requires a number of sessions lasting nearly a year or more. Counseling serves as a screening process—NHSB officials often talk of seeking a "match" between a property and a borrower. After the purchase, counselors provide advice to financially strapped borrowers and may help them renegotiate payment schedules.

10. COMMUNITY DEVELOPMENT LENDING IS DIFFERENT

The activities of NHSB are different in many ways from the usual for-profit home mortgage lending that banks perform. NHSB coordinates a package of home purchase financing for a borrower that is generally more complex than typical arrangements. A first mortgage is obtained, sometimes from a conventional lender, often on conventional terms, but occasionally through a special mortgage program tailored to low-income borrowers. NHSB also makes first mortgages from its own loan fund. Some NHSB loans are sold in a secondary market run by a national organization, Neighborhood Housing Services of America. A second mortgage is usually crucial to the package since borrowers generally have just a minimal amount of cash. NHSB arranges for the second mortgage, usually from its own loan fund. Further funding may be available

[32] Neighborhood Housing Services of Baltimore, Inc., is a private nonprofit organization and is affiliated with a network of over 200 Neighborhood Housing Services organizations nationwide. NHSB also operates an affiliated organization, Neighborhood Rental Services, that renovates rental property.

from a "Closing Cost Loan Program" it administers. Loan terms often are designed to retire the junior debt first before retiring principal on the first mortgage. NHSB officials often refer to their supplemental financing as "soft second" money, since they are sometimes willing to reschedule payments if the borrower suffers an adverse financial shock.

The NHSB goes to great lengths to minimize the credit risks posed by its clients. Extensive information about borrowers emerges in the early counseling stage. Borrowers are carefully selected for the right "fit" with the property in the sense that the payments will be affordable. Borrowers generally are required to save a down payment of at least $1,000, which provides an equity interest in the home and helps demonstrate the discipline required to manage mortgage payments. NHSB also closely monitors the neighborhood and encourages close connections between residents through community clean-up projects, neighborhood organizations, and crime patrols. This helps NHSB learn early on about a borrower's financial difficulty before a costly mortgage default, generally the last stage of financial distress for a conventional borrower. In addition, renovations are designed in part to minimize the chance of costly repairs—new furnaces and appliances are often installed, even when existing units satisfy city housing codes. Active post-purchase counseling helps minimize the ex post costs of financial distress.

Second, the NHSB spends much time coordinating investment in targeted neighborhoods. A primary goal of NHSB is to achieve a "generally good physical appearance" in a neighborhood. It tries to develop vacant properties, rehabilitate existing properties, and improve commercial areas. It encourages owner occupancy in the belief that owners who occupy their own home spend more on maintenance and improvements. It tries to influence local government spending on amenities such as streets, sidewalks, and public lands. Sometimes it helps arrange the departure of taverns or other "undesirable" businesses. In short, much of NHSB's activity involves trying to overcome just the sort of neighborhood externalities discussed earlier in this essay.

Third, NHSB lending requires substantial subsidies. Its counseling, monitoring, and coordination activities are quite labor-intensive, and home purchase transactions are often subsidized. Operating and program expenditures are funded out of federal, state, and local grants and private donations. Officials admit that they often "overimprove" a house, undertaking renovations that cost more than the resulting increase in market value. NHSB officials also recognize that their second-mortgage loans are not "bankable" in that no private lender would lend on the same terms. In fact, loans sold to Neighborhood Housing Services of America, a national umbrella group, are backing for notes sold to institutional investors who agree to receive a below-market rate of return on their "social investment."

11. SHOULD BANKS DO COMMUNITY DEVELOPMENT LENDING?

The community development lending performed by CDOs is the type of subsidized lending encouraged by CRA regulations. As suggested above, however, the community development activities of CDOs like NHSB differ in many respects from traditional banking. Do banks have any comparative advantage in providing community development lending? Furthermore, how many of these activities are banks capable of performing safely?

First, the concessionary lending done by NHSB seems inappropriate for insured depository institutions. Although CRA regulations require that lending be "sound," the regulations also encourage concessionary investments and charitable contributions toward community development. Banks get CRA credit for offering higher loan-to-value ratios and other "more flexible" lending terms, which can only mean more risky lending terms. In fact, in the newly adopted CRA regulations, concessionary community development investments are included alongside low-income neighborhood lending in assessing CRA compliance. This approach threatens to blur the distinction between concessionary and for-profit lending and could induce banks to make underpriced or excessively risky loans. In the absence of convincing evidence that banks pass up economically viable lending opportunities in low-income neighborhoods, the attempt to stimulate additional bank lending to these neighborhoods risks saddling the banking industry with a large portfolio of poorly performing mortgages if it has any effect at all. Since these debts would carry regulators' implicit imprimatur, forbearance in the event of widespread losses would be hard to avoid, as in the case of sovereign debt in the 1980s.

Maintaining a clear boundary at banks between concessionary and for-profit lending is thus crucial to the clarity and integrity of regulatory supervision. Examiners need to know whether a portfolio is intended to be profitable or philanthropic. Allowing government-insured banks to carry concessionary lending on their books hides the cost, unless the subsidy is explicitly recognized up front through higher loan loss reserves or discounting the value of the loan for interest rate subsidies. Funding concessionary lending explicitly out of retained earnings or bank capital subjects transfers to at least minimal accounting safeguards, ensures timely recognition of costs, and makes their redistributive nature clear. Better yet, concessionary community development lending could be conducted separately through a community development subsidiary of a bank's holding company. This would have the advantage of keeping such lending programs separate from the bank's conventional lending, making the evaluation of both portfolios easier.

One impediment to community development lending by banks or bank holding companies, however, is the extensive counseling that appears crucial to lending by NHSB and other CDOs. Unlike CDO counselors, bank loan

officers face regulatory constraints on their ability to communicate with borrowers; under the Equal Credit Opportunity Act, they cannot tell an applicant what to do to qualify for a loan without triggering a formal application with the required documentation and disclosures. As a result, NHSB counselors learn far more about borrowers than would bank loan officers. Because the screening inherent in these programs appears to be essential to the viability of community development lending, banks often contract with community development groups to perform pre-application counseling. Thus, even bank holding company subsidiaries may require external assistance to perform community development lending.[33]

Would banks have any comparative advantage in community development lending that would motivate a community development requirement for banks? The experience of the NHSB suggests the answer is no. NHSB counselors have extensive contact with local bank lending officers and appear well informed about specialized loan programs available and the constraints associated with conventional for-profit lending. In addition, NHSB has extensive contact with residents through ongoing work with neighborhood associations, and thus sometimes has better information about borrowers than would a bank. If anything, then, CDOs would seem to have a comparative advantage over banks in the community development lending encouraged by the CRA regulations.

Banks have made substantial contributions of funds to community development, much of it under agreements negotiated with community groups.[34] Do banks have any special advantage at making such contributions? Perhaps their working involvement with local community development groups helps them compare and evaluate organizations. Bankers often speak of trying to select "truly responsible" organizations.[35] On the other hand, banks and other lenders appear to be a minority among NHSB's contributors. Most are corporations, individuals, and foundations in the Baltimore area, and it seems unlikely that they learned about NHSB through joint lending arrangements. Also, the national network of Neighborhood Housing Services organizations, along with explicit certification programs, assures some uniformity, making evaluation

[33] Other community development activities of the NHSB seem difficult for banks as well. For example, much of the coordinating activity that seems vital to the CDO approach involves finding owner-occupants that are viewed as beneficial to the neighborhood. Such discrimination among buyers or borrowers would pose legal problems for a bank real estate subsidiary.

[34] Allen Fishbein (1993) of the Center for Community Change estimates that around $35 billion has been "committed" by banks and savings and loans since the late 1970s under agreements with community groups. The banking agencies officially view commitments for future action as "largely inapplicable to an assessment of the applicant's CRA performance" (Garwood and Smith 1993, p. 260).

[35] "Our job, quite frankly, is to choose partnerships with organizations that do not have hidden agendas, are truly responsible and have an appreciation of our limitations" (Milling 1994, p. 7).

easier for outside investors and contributors. Thus, it is unclear why banks would have any advantage in evaluating subsidy recipients.

To summarize, there does not seem to be a compelling rationale for imposing a costly lending obligation on banks. Ultimately such an obligation is a tax on bank-dependent borrowers and depositors. Similarly, there seems to be scant economic justification for looking to banks for the concessionary investments encouraged by the CRA regulations.

12. WHERE DO WE GO FROM HERE?

Our low-income neighborhoods nevertheless remain in appalling condition. Community development lending seems to be a promising way of channeling resources toward improving conditions in these neighborhoods. The evidence summarized in this essay, however, suggests that the CRA is not an efficient vehicle for revitalizing decayed neighborhoods, despite its laudable goals.

An alternative to the CRA is to fund community development subsidies directly out of general tax revenues. The Community Development Banking Act (CDBA), signed into law in September 1994, provides federal funding for community development. This Act creates a new government corporation, called the Community Development Financial Institutions Fund, charged with providing financial and technical assistance to specialized, limited-purpose community development financial institutions (CDFIs), and authorizes expenditures of $382 million over four years.[36] Explicit appropriation for community development has distinct advantages over drawing subsidies from banks. Removing the implicit tax burden on banks would reduce existing distortions in financial flows and avoid the risks of concessionary lending. By directing assistance through organizations that have community development as their sole mission, monitoring and evaluation of such assistance would become transparent.

The CDBA leaves considerable uncertainty, however, about important aspects of the Fund's operation.[37] For example, the CDBA requires that a CDFI have "a primary mission of promoting community development," without defining the latter term. Other key provisions depend on undefined concepts like "significant unmet needs for loans or equity investments." More fundamentally, distributing public money to a network of small, information-intensive

[36] Funds can be provided in the form of grants, equity investments, loans, or deposits, and must be matched dollar for dollar by private funds. The Fund is prohibited from holding over 50 percent of the equity of a CDFI and may not provide more than $5 million to any one CDFI during any three-year period. Up to one-third of the appropriation may be applied toward a depository institution's deposit insurance premium. The appropriation covers administrative costs as well. Many similar efforts have been funded in smaller amounts in the past. See Wells and Jackson (1993). Macey and Miller (1993) also argue that direct funding of community development would be superior to the CRA as it is currently implemented.

[37] See Townsend (1994) for a critique of an earlier draft of the Community Development Banking Act.

lending organizations can create adverse incentives in much the same way that deposit insurance can distort bank behavior. Moreover, the oversight and reporting provisions in the CDBA are notably less detailed than current banking legislation, and formal regulations have been left to the Fund to establish. Consequently, much will depend on the way in which the CDBA is implemented; in particular, effective screening and monitoring is essential. Nevertheless, the CDBA or something similar to it seems to be more promising than the CRA for dealing with the plight of the nation's low-income neighborhoods.

REFERENCES

Art, Robert C. "Social Responsibility in Bank Credit Decisions: The Community Reinvestment Act One Decade Later," *Pacific Law Journal*, vol. 18 (1987), pp. 1071–1139.

Avery, Robert B., Patricia E. Beeson, and Mark Sniderman. "Accounting for Racial Differences in Housing Credit Markets," Working Paper 9310. Cleveland: Federal Reserve Bank of Cleveland, December 1993.

Avery, Robert B., and Thomas M. Buynak. "Mortgage Redlining: Some New Evidence," Federal Reserve Bank of Cleveland *Economic Review,* Summer 1981, pp. 18–32.

Barth, James B., Joseph J. Cordes, and Anthony M. J. Yezer. "FHA Mortgage Insurance and High-Risk Mortgage Lending: Some Lessons for Policy," *Housing Finance Review,* vol. 2 (April 1983), pp. 93–107.

_____. "Financial Institution Regulations, Redlining and Mortgage Markets," in *Proceedings of a Conference on the Regulation of Financial Institutions* (Federal Reserve Bank of Boston Conference Series No. 21, October 1979), pp. 101–43.

Benston, George J. "Redlining Research: A Review and Critical Analysis Discussion," in *Proceedings of a Conference on the Regulation of Financial Institutions* (Federal Reserve Bank of Boston Conference Series No. 21, October 1979), pp. 144–95.

Berkovec, James, Glenn Canner, Stuart Gabriel, and Timothy Hannan. "Race, Redlining, and Residential Mortgage Loan Performance," *Journal of Real Estate and Economics,* vol. 9 (November 1994), pp. 263–94.

Blau, Francine D., and Lawrence M. Kahn. "Causes and Consequences of Layoffs," *Economic Inquiry,* vol. 19 (April 1981), pp. 270–96.

Board of Governors of the Federal Reserve System. "Community Reinvestment Act: Joint Notice of Proposed Rulemaking" (Docket No. R-0822, September 27, 1994).

_____. *Report to Congress on Community Development Lending by Depository Institutions*. Washington: Board of Governors, 1993.

Bradbury, Katherine L., Karl E. Case, and Constance R. Dunham. "Geographic Patterns of Mortgage Lending in Boston, 1982–1987," *New England Economic Review*, September/October 1989, pp. 3–30.

Browne, Lynn Elaine. Letter to the Editor, *The Wall Street Journal*, September 21, 1993a.

_____. "Critique of Boston Lending Study Missed the Mark," *American Banker*, September 9, 1993b, p. 21.

Cain, Glen G. "The Economic Analysis of Labor Market Discrimination: A Survey," in Earl Ashenfelter and Richard Layard, eds., *Handbook of Labor Economics*. New York: North Holland, 1986.

Calem, Paul, and Michael Stutzer. "The Simple Analytics of Observed Discrimination in Credit Markets." Manuscript. Federal Reserve Bank of Philadelphia, June 7, 1994.

Calomiris, Charles W., Charles M. Kahn, and Stanley D. Longhofer. "Housing-Finance Intervention and Private Incentives: Helping Minorities and the Poor," *Journal of Money, Credit, and Banking*, vol. 26 (August 1994, Part 2), pp. 634–74.

Canner, Glenn B. *Redlining: Research and Federal Legislative Response*. Staff Studies 121. Washington: Board of Governors of the Federal Reserve System, 1982.

_____, Stuart A. Gabriel, and J. Michael Woolley. "Race, Default Risk and Mortgage Lending: A Study of the FHA and Conventional Loan Markets," *Southern Economic Journal*, vol. 58 (July 1991), pp. 249–62.

Canner, Glenn B., and Charles A. Luckett. "Consumer Debt Repayment Woes: Insights from a Household Survey," *Journal of Retail Banking*, vol. 12 (Spring 1990), pp. 55–62.

Canner, Glenn B., Wayne Passmore, and Dolores S. Smith. "Residential Lending to Low-Income and Minority Families: Evidence from the 1992 HMDA Data," *Federal Reserve Bulletin*, vol. 80 (February 1994), pp. 79–108.

Carr, James H., and Isaac F. Megbolugbe. "The Federal Reserve Bank of Boston Study on Mortgage Lending Revisited," Fannie Mae Working Paper. Washington, D.C., 1993.

Cloud, Cathy, and George Galster. "What Do We Know About Racial Discrimination in Mortgage Markets?" *Review of Black Political Economy*, vol. 22 (Summer 1993), pp. 101–20.

Cochrane, John H. "A Simple Test of Consumption Insurance," *Journal of Political Economy,* vol. 99 (October 1991), pp. 957–76.

Dennis, Warren L. "The Community Reinvestment Act of 1977," Working Paper 24. West LaFayette, Ind.: Credit Research Center, Krannert Graduate School of Management, Purdue University, 1978.

Duca, John V., and Stuart S. Rosenthal. "Borrowing Constraints, Household Debt, and Racial Discrimination in Loan Markets," *Journal of Financial Intermediation,* vol. 3 (October 1993), pp. 77–103.

Evans, Richard D., Brian A. Maris, and Robert I. Weinstein. "Expected Loss and Mortgage Default Risk," *Quarterly Journal of Business and Economics,* vol. 24 (Winter 1985), pp. 75–92.

Federal Financial Institutions Examination Council. Press Release. October 26, 1994.

Fishbein, Allen J. "The Community Reinvestment Act After Fifteen Years: It Works, But Strengthened Federal Enforcement is Needed," *Fordham Urban Law Journal,* vol. 20 (1993), pp. 293–310.

Fix, Michael, and Raymond J. Struyk. *Clear and Convincing Evidence: Measurement of Discrimination in America*. Washington: Urban Institute Press, 1993.

Flanagan, Robert J. "Discrimination Theory, Labor Turnover, and Racial Unemployment Differentials," *Journal of Human Resources,* vol. 13 (Spring 1978), pp. 187–207.

Gabriel, Stuart A., and Stuart S. Rosenthal. "Credit Rationing, Race, and the Mortgage Market," *Journal of Urban Economics,* vol. 29 (May 1991), pp. 371–79.

Garwood, Griffith L., and Dolores S. Smith. "The Community Reinvestment Act: Evolution and Current Issues," *Federal Reserve Bulletin,* vol. 79 (April 1993), pp. 251–67.

Glennon, Dennis, and Mitchell Stengel. "An Evaluation of the Federal Reserve Bank of Boston's Study of Racial Discrimination in Mortgage Lending," Working Paper 94-2. Washington: Office of the Comptroller of the Currency, Economic & Policy Analysis, Washington, D.C., April 1994.

Guttentag, Jack M., and Susan L. Wachter. "Redlining and Public Policy," *New York University, Solomon Brothers Center for the Study of Financial Institutions Monograph Series in Finance and Economics,* 1980-1.

Hawley, Clifford B., and Edwin T. Fujii. "Discrimination in Consumer Credit Markets," *Eastern Economic Journal,* vol. 17 (January–March 1991), pp. 21–30.

Heckman, James J., and Peter Siegelman. "The Urban Institute Audit Studies: Their Methods and Findings," in Michael Fix and Raymond J. Struyk, eds., *Clear and Convincing Evidence: Measurement of Discrimination in America.* Washington: Urban Institute Press, 1993.

Helper, Rose. *Racial Policies and Practices of Real Estate Brokers.* Minneapolis: University of Minnesota Press, 1969.

Holmes, Andrew, and Paul Horwitz. "Mortgage Redlining: Race, Risk, and Demand," *Journal of Finance,* vol. 49 (March 1994), pp. 81–99.

Horne, David K. "Evaluating the Role of Race in Mortgage Lending," *FDIC Banking Review,* vol. 7 (Spring/Summer 1994), pp. 1–15.

Jackson, Kenneth T. *The Crabgrass Frontier.* New York: Oxford University Press, 1985.

Jackson, Peter, and Edward Montgomery. "Layoffs, Discharges and Youth Unemployment," in Richard B. Freeman and Harry J. Holzer, eds., *The Black Youth Employment Crisis*. Chicago: University of Chicago Press, 1986.

Jaynes, Gerald David. "The Labor Market Status of Black Americans: 1939–1985," *Journal of Economic Perspectives,* vol. 4 (Fall 1990), pp. 9–24.

_____, and Robin M. Williams, Jr., eds. *A Common Destiny: Blacks and American Society*. Washington: National Academy Press, 1989.

Johnston, J. *Econometric Methods*. New York: McGraw-Hill, 1963.

Kennickell, Arthur, and Janice Shack-Marquez. "Changes in Family Finances from 1983 to 1989: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin,* vol. 78 (January 1992), pp. 1–18.

King, A. Thomas. "Discrimination in Mortgage Lending: A Study of Three Cities," *New York University, Solomon Brothers Center for the Study of Financial Institutions Monograph Series in Finance and Economics,* 1980-4.

Lang, William W., and Leonard I. Nakamura. "A Model of Redlining," *Journal of Urban Economics,* vol. 33 (March 1993), pp. 223–34.

Liebowitz, Stan. "A Study That Deserves No Credit," *The Wall Street Journal,* September 1, 1993.

Macey, Jonathan R., and Geoffrey P. Miller. "The Community Reinvestment Act: An Economic Analysis," *Virginia Law Review,* vol. 79 (March 1993), pp. 291–348.

Massey, Douglas S., and Nancy A. Denton. *American Apartheid*. Cambridge, Mass.: Harvard University Press, 1993.

Milling, R. King. "Banks Must Lead for Cities to Thrive," *Stone Soup,* vol. 12 (Summer 1994), pp. 6–7.

Munnell, Alicia H., Lynn E. Browne, James McEneaney, and Geoffrey M.

B. Tootell. "Mortgage Lending in Boston: Interpreting the HMDA Data," Working Paper 92-7. Boston: Federal Reserve Bank of Boston, 1992.

Munnell, Alicia H., Geoffrey M. B. Tootell, Lynn E. Browne, and James McEneaney. "Is Discrimination Racial or Geographic?" Manuscript. Federal Reserve Bank of Boston, undated.

National Center for Health Statistics. *Health, United States,* 1993. Hyattsville, Md.: Public Health Service, 1994.

National Conference. *Taking America's Pulse: The Full Report of the National Conference Survey on Inter-Group Relations*. New York: The National Conference, 1994.

Quercia, Roberto G., and Michael A. Stegman. "Residential Mortgage Default: A Review of the Literature," *Journal of Housing Research,* vol. 3 (1992), pp. 341–79.

Schafer, Robert, and Helen F. Ladd. *Discrimination in Mortgage Lending*. Cambridge, Mass.: The MIT Press, 1981.

Schill, Michael H., and Susan M. Wachter. "Borrower and Neighborhood Racial and Income Characteristics and Financial Institution Mortgage Application Screening," *Journal of Real Estate Finance and Economics,* vol. 9 (December 1994), pp. 223–39.

_____. "A Tale of Two Cities: Racial and Geographic Disparities in Home Mortgage Lending in Boston and Philadelphia," Working Paper 148. Philadelphia: Wharton Real Estate Center, June 1, 1993.

Shadow Financial Regulatory Committee. *Statement of the Shadow Financial Regulatory Committee on Proposed Revisions to Community Reinvestment Regulations,* Statement No. 105, February 14, 1994.

Swire, Peter P. "The Persistent Problem of Lending Discrimination: A Law and Economics Analysis." Manuscript. University of Virginia, August 26, 1994.

Townsend, Robert M. "Community Development Banking and Financial Institutions Act: A Critique with Recommendations." Paper presented to the 30th Annual Conference on Bank Structure and Competition, May 11–13, 1994.

U.S. Congress, Senate. *Community Reinvestment Act.* Hearing, 100 Cong. 2 Sess. Washington: Government Printing Office, 1988.

_____. *Community Credit Needs.* Hearing, 95 Cong. 1 Sess. Washington: Government Printing Office, 1977.

Wells, F. Jean, and William Jackson. "Community Development Lenders: Policy Options and the Track Record," Congressional Research Service Report for Congress, May 11, 1993.

White, Lawrence J. "The Community Reinvestment Act: Good Intentions

Headed in the Wrong Direction," *Fordham Urban Law Journal,* vol. 20 (1993), pp. 281–92.

Wilson, William Julius. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* Chicago: University of Chicago Press, 1987.

Woelfel, Charles J. *Encyclopedia of Banking and Finance,* 10th ed. Chicago: Probus, 1994.

Zandi, Mark. "Boston Fed's Bias Study Was Deeply Flawed," *American Banker,* August 19, 1993, p. 13.

When Geometry Emerged: Some Neglected Early Contributions to Offer-Curve Analysis

Thomas M. Humphrey

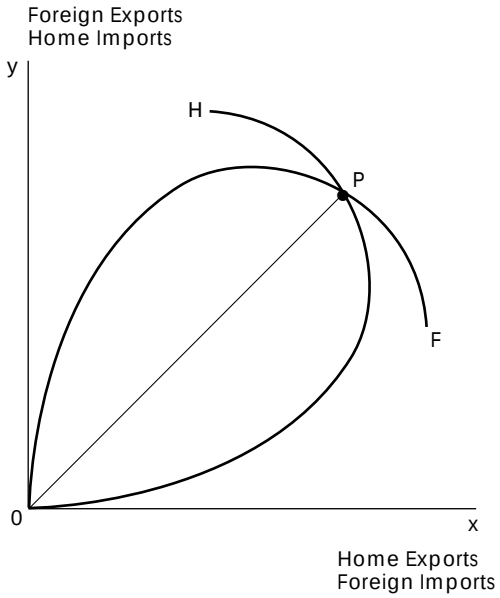
In his 1952 *A Geometry of International Trade,* Nobel Laureate James Meade presented the definitive modern version of the celebrated reciprocal-demand, or offer-curve, diagram of the trade theorist. The diagram features curves depicting alternative quantities of exports and imports countries are willing to trade at all possible prices (see Figure 1).

Let two countries, home and foreign, trade two goods, x and y. Measure quantities of these goods along the horizontal and vertical axes, respectively. Suppose the home country exports good x and imports good y while the foreign country does the converse. The slope of any ray from the origin expresses the relative price of x in terms of y. That is, it expresses the quantity of y exchanged per unit of x, or y price of x. Curve H is the home country's offer curve. Curve F is the foreigner's. Each curve shows alternative quantities of imports demanded and exports supplied at all price ratios or terms of trade.

As drawn, the curves display declining *elasticity,* or price responsiveness, throughout their length. They slope upward from left to right when the demand for imports in terms of exports is elastic—that is, when more exports are offered for imports at successively lower import prices. They cease to slope upward when import demand becomes unit-elastic. At such points, the quantity of exports offered for total imports remains unchanged as import prices fall. They bend backward (or downward) when import demand is inelastic. Along such segments, fewer exports are offered for total imports when import prices fall.

The author thanks Bob Hetzel, Peter Ireland, Tony Kuprianov, Jeff Lacker, and Dawn Spinozza for their helpful comments and suggestions. The views expressed are those of the author and not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

Figure 1 Offer-Curve Diagram



Curves *H* and *F* show alternative quantities of exports offered and imports demanded by each country at all different price ratios. Equilibrium occurs at intersection point *P*. The slope of ray 0*P* is the equilibrium price ratio or terms of trade. The coordinates of point *P* depict the equilibrium quantities of goods traded.

World trade equilibrium occurs at point *P*, where the offer curves intersect. At that point, the market-clearing price ratio, or terms of trade, given by the slope of the ray 0*P* equates each nation's import demand with the other's export supply. The supply of both commodities equals the demand for them, and the coordinates of point *P* show the resulting equilibrium volume of world trade.

The foregoing diagram has proved indispensable in illuminating the central ideas of trade theory. Generations of professors and their students have employed it to demonstrate how the strength and elasticity of each country's demand for the other's product determine the equilibrium volume and terms of world trade. Likewise, scores of textbooks use it to illustrate how tariffs, technological advances, resource discoveries, taste changes, and other such disturbances shift the offer curves and thereby alter world trade equilibrium.

That a simple geometrical diagram would prove so useful is hardly surprising. Other economic diagrams, including the Keynesian cross, Marshallian scissors, Hicksian IS-LM, Knightian circular flow, Vinerian cost envelope, Fisher-Haberler production possibility frontier, and expectations-augmented Phillips Curve, or zero long-run trade-off between inflation and unemployment, have proved equally indispensable. Indeed, as long ago as 1879, Alfred Marshall insisted that diagrams are absolutely essential to exact reasoning in economics

because they yield many of the same results as higher mathematics while being accessible to the mathematically untrained.

What is surprising is how little has been written on the doctrinal history of offer curves. Few systematic surveys of that topic exist. Textbooks scarcely do it justice. Even history-of-thought treatises spotlight at best only a handful of the chief contributions. Meade himself was largely silent on the diagram's history even though it was more than 100 years old when he published his *Geometry*.

The development of offer-curve analysis involves some of the leading names in classical and neoclassical economics. John Stuart Mill, Robert Torrens, Alfred Marshall, Francis Ysidro Edgeworth, and Abba Lerner all contributed to the diagram's development and policy applications. Mill invented reciprocal-demand schedules. He used them to determine precisely where, within the limits set by comparative-cost ratios, the terms of trade, or quantity of imports bought by a unit of exports, must fall. He used them also to estimate the impact of tariffs and technology shocks on the terms of trade. Torrens likewise employed such schedules to argue the merits of a policy of reciprocity in tariff erection and removal. Marshall translated Mill into geometry and examined the stability of offer-curve equilibria. Edgeworth combined offer curves with indifference maps to derive the theory of the optimum tariff. Lerner corrected Edgeworth's error of alleging an asymmetry between export and import taxes and also showed how the government's disposal of the tax receipts influences the position of the curves. Finally, Meade completed the analysis by deriving offer curves from price vectors and trade indifference maps—themselves derived from underlying production possibility frontiers and consumption indifference curves. His demonstration was crucial. It proved once and for all that domestic production as well as consumption conditions influence offer curves.

The paragraphs below attempt to trace this evolution and to identify specific contributions to it. Besides exhuming lost or forgotten insights, such an exercise serves as a partial corrective to the tendency of modern trade theory textbooks to overemphasize Meade's contribution at the expense of those of his predecessors. By resurrecting pathbreaking earlier work, the exercise dispels misconceptions concerning the origins of offer-curve analysis. It establishes that the diagram is not a twentieth-century innovation. In this connection, it reveals that at least two of Meade's predecessors instinctively grasped the concept of offer curves long before Marshall invented the diagram. It nevertheless indicates that the diagram played a crucial role in advancing the analysis. By crystallizing, condensing, and generalizing earlier insights into a powerful yet simple visual image, the diagram at once rendered them transparent and easy to comprehend. For evidence of the diagram's power to illuminate and enhance earlier work, one need only refer to Mill's and Torrens' laborious verbal and numerical examples. Those examples convey their full meaning only when translated into geometry. For this reason, the following paragraphs take the liberty of interpreting Mill and

Torrens with the aid of offer-curve diagrams unavailable to them when they wrote. Such anachronisms involve little distortion when, as is the case here, they correspond faithfully to the original work.

1. JOHN STUART MILL

Reciprocal-demand, or offer-curve, analysis originated to fill a gap in David Ricardo's theory of comparative advantage. Ricardo, in his 1817 volume *On the Principles of Political Economy and Taxation*, demonstrated (1) that comparative-cost ratios in each country determine pre-trade relative prices, (2) that international differences in such prices render trade advantageous, and (3) that countries therefore trade when their comparative-cost ratios differ, exporting their relatively cheap-to-produce goods and importing their relatively dear-to-produce ones. Ricardo also indicated that the post-trade terms of trade must fall somewhere between these limiting cost ratios. He did not, however, explain what determines the terms of trade or where it would tend to settle. He merely assumed it would fall roughly halfway between the cost ratios without explaining why.

Priority for identifying the relative strength of each country's demand for the other's product as the determinant of the terms of trade goes to James Pennington, Robert Torrens, and, above all, John Stuart Mill. Pennington in 1840 was the first to state the point in print. His account, however, was marred by the notion that volatile reciprocal demands cause the terms of trade to oscillate ceaselessly within the limiting cost ratios rather than to achieve a stable determinate value. Torrens was the first to coin the phrase "reciprocal demand" (see Viner [1937], p. 536). Because he used the concept to argue against unilateral tariff reduction, however, his analysis was condemned by his classical contemporaries, most of whom were free-traders. Of the three originators of the reciprocal-demand idea, John Stuart Mill exerted by far the greatest influence. His conception of reciprocal demand as a schedule or function of price enabled him to convey its importance more clearly, systematically, and convincingly than had Pennington and Torrens. In any case, it was from Mill rather than from the latter two writers that later economists took the idea.

Reciprocal-Demand Theory

Mill stated the idea first in his essay "On the Laws of Interchange Between Nations," which he wrote in 1829–30 but did not publish until 1844 in response to Torrens' (1841–42) *The Budget*. He presented it again in Chapter 18 of Book III of his 1848 *Principles of Political Economy*. His statement is as modern as the latest trade textbook.

First comes the notion of comparative-cost ratios as limiting values for the terms of trade.[1] Next comes the argument that reciprocal-demand schedules express an inverse relationship between import relative price and quantity demanded.[2] There follows the idea that to each quantity of imports demanded along a reciprocal-demand schedule there corresponds an associated amount of exports supplied. This amount equals the product of import price and quantity. It expresses the condition that a trading country's export receipts constitute its means of purchasing imports of the same value.[3] In other words, reciprocal-demand schedules are at once demand-and-supply curves expressing import demand in terms of export supply.[4]

Finally comes Mill's requirement that reciprocal-demand schedules intersect at the equilibrium volume and terms of trade. The latter variable, the equilibrium price ratio, clears the world goods market such that each country's demand for imports equals the other's export supply. Mill referred to this equilibrium condition as the Law of International Values.[5]

Having shown how reciprocal-demand schedules intersect to determine world trade equilibrium, Mill examined the stability of that equilibrium. In language familiar to any modern economist, he argued that a displacement of the terms of trade from its equilibrium value would invoke an excess world demand for one good and corresponding world excess supply of the other. These excess demands and supplies would exert corrective pressure on the terms of trade until it returned to equilibrium.

Clearly, Mill had put his stamp on the diagram just as surely as if he had drawn it himself. That much is evident from how easily his statements map into offer-curve space (see Figure 2). Closest to the axes lie the comparative-cost (*CC*) lines. Their slopes represent the production substitution, opportunity cost, and domestic price ratios of the two goods in each country in the absence of trade. The lines are drawn straight to correspond to the Ricardian or classical

[1] "The limits within which the [relative price or exchange ratio between importables and exportables] is confined, are the ratio between their costs of production in the one country, and the ratio between their costs of production in the other" ([1844] 1968, p. 12).

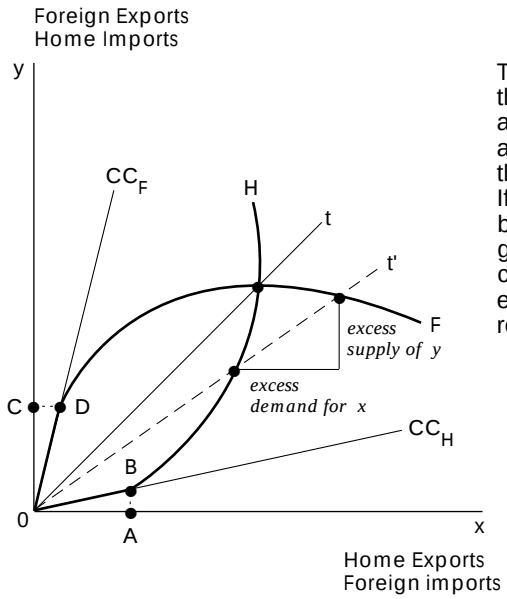
[2] "The higher the price, the fewer will be the purchasers, and the smaller the quantity sold. The lower the price, the greater will in general be the number of purchasers, and the greater the quantity disposed of" ([1844] 1968, p. 9).

[3] Let X and M denote export and import quantities and p_x and p_m their money prices. Then, at each point on a reciprocal-demand schedule, the value of exports supplied p_xX equals the value of imports demanded p_mM, or $p_xX = p_mM$. Dividing both sides of this equation by p_x yields $X = (p_m/p_x)M$, which says that the amount of exports offered equals the product of import relative price and quantity.

[4] "The supply brought by the one constitutes his demand for what is brought by the other. So that supply and demand are but another expression for reciprocal demand" ([1848] 1909, p. 593).

[5] "The produce of a country exchanges for the produce of other countries, at such values as are required in order that the whole of her exports may exactly pay for the whole of her imports. This law of International Values is but an extension of the more general law of Value, which we called the Equation of Supply and Demand" ([1848] 1909, p. 592).

Figure 2 Mill's Reciprocal-Demand Schedules



The schedules *H* and *F* follow the comparative-cost lines over a range in which the countries are indifferent to trade. Then they bend toward equilibrium. If the terms-of-trade ray lies below the market-clearing ray going through the point of offer-curve intersection, the resulting excess demands and supplies restore it to equilibrium.

assumption of constant marginal and average costs. That the home country's line is the flatter of the two indicates that it (the home country) possesses a comparative-cost advantage in producing the good measured along the horizontal axis. Conversely, the steep slope of the foreigner's cost line signifies his comparative-cost advantage in producing the good measured along the vertical axis.

As for the offer curves, they follow the cost lines over a range in which the countries are indifferent to trade. Thus if the home country faces world terms of trade equal to its domestic opportunity cost ratio $AB/0A$, it cares not whether it obtains AB units of good y through domestic production or through foreign trade. Either way, the cost is the same, namely, $0A$ units of good x. Likewise, when the terms-of-trade ray coincides with the foreigner's cost line, he is equally willing to obtain CD units of good x through trade or domestic production. In each case, he sacrifices $0C$ units of good y.

At a certain point, however, the offer curves depart from the cost lines. Thus, at point B, the home country's curve begins to bend away from its cost schedule. Precisely at this point, the home country specializes completely in the production of its exportable, the excess of which it trades for importables to reach its desired consumption bundle—the same bundle it would consume under

self-sufficiency. Beyond this point, however, the offer curve bends upward in response to better terms of trade. The resulting fall in the export price of imports has a twofold effect. It increases the quantity of imports demanded. And, by inducing the country to consume fewer exportables, it makes more of those goods available for sale abroad and so raises the quantity of exports offered. This latter step is necessary since the country already is at its specialization point and can produce no more goods for export. A similar analysis holds for the foreign offer curve, which at point *D* bends away from its cost line toward equilibrium.

Equilibrium occurs where the curves intersect. Running through that intersection point is a ray from the origin whose slope represents the market-clearing price ratio or terms of trade. A self-correcting mechanism ensures this equilibrium price ratio will prevail. Should a disequilibrium terms of trade such as that indicated by the slope of the lower dashed ray occur, the result would be an excess demand for the home country's exports and an excess supply of its imports. The resulting rise in export prices and fall in import prices would restore the equilibrium terms of trade.

Applications

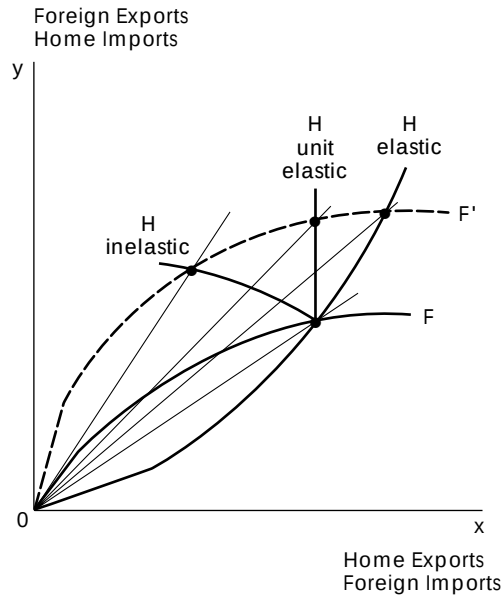
Having derived his reciprocal-demand apparatus, Mill put it to work in analyzing a variety of cases. He showed that where the terms of trade settle between the autarkic cost lines depends on the relative strength and elasticities of the reciprocal demands. The greater and more elastic one country's demand relative to the other's, the more the terms of trade would move against the first country and in favor of the second. And in the case of a large country trading with a small one, he showed that the latter's offer curve might cut the former's in its linear segment. If so, the terms of trade would coincide with the large country's cost ratio. All gains from trade would go to the small country, and the large one would be incompletely specialized in production.

Technological Improvements and the Terms of Trade

He likewise applied his reciprocal-demand technique to predict the terms-of-trade effects of a cost-reducing improvement in the foreign country's export sector (see Figure 3). To do so, he distinguished between elastic, unit-elastic, and inelastic home import demands. Such elasticities result in greater, unchanged, and smaller outlays of exports as import prices fall. Accordingly, they give rise to upward-sloping, vertical, and backward-bending home offer curves, respectively.

Mill concluded that the improvement would, by raising the supply of exports relative to the demand for them, turn the terms of trade against the foreign country by an amount that depended on the home country's import-demand elasticity. The improvement lowers the foreigner's opportunity cost of

Figure 3 Technological Advance and Trade Taxes



Technological advance shifts the foreigner's offer curve upward. The extent of his terms-of-trade deterioration depends on the elasticity of the home country's curve. Conversely, the foreigner's trade taxes shift his curve downward. His terms of trade improve by an amount that depends on the elasticity of the home country's curve.

producing exports. It thus enables him to offer more for any given quantity of imports. In so doing, it shifts up his offer curve equiproportionally to the cost reduction.[6] The resulting counterclockwise rotation of the equilibrium price ray constitutes the terms-of-trade deterioration. The deterioration is proportionally greater than, equal to, or less than the cost reduction as the home offer curve is inelastic (backward-bending), unitary-elastic (vertical), or elastic (upward-sloping). Mill's inelastic case anticipated the modern concept of *immiserising growth* in which the adverse terms-of-trade effects of improved productivity swamp the beneficial output effects and so make the country worse off than before.

Trade Taxes and the Terms of Trade

Mill also employed his reciprocal-demand apparatus to examine the terms-of-trade effects of a tax on exports or imports. Despite his aversion to all forms of trade restriction, he demonstrated that such taxes could improve the levying

[6] Mill admitted that this result might not hold exactly if the cost reduction exerted an income effect on the foreigner's own demand for his exportable good.

country's terms of trade in proportion to the elasticity of the other's reciprocal demand.

Let the foreign government levy the tax and consume the proceeds (see Figure 3 again). If the tax is on imports, it reduces the demand for them. If it is on exports, it reduces their supply. Either tax, therefore, causes an equiproportionate downward shift in the foreigner's offer curve and thus improves his terms of trade. The improvement is in greater, equal, or lesser proportion than the tax depending on whether the home country's offer curve is backward-bending (inelastic), vertical (unit-elastic), or upward-sloping (elastic).

Theoretically, then, trade taxes could improve the terms of trade and thus national welfare. Nevertheless, Mill opposed them on practical and moral grounds. In his view, they invite retaliatory duties abroad that nullify the initial terms-of-trade improvement. Worse still, they bring costly reductions in the volume of world trade. Even in the absence of retaliation, they are unjust because they benefit the levying country at the expense of other countries. Since the rest of the world's loss exceeds the dutying country's gain, such taxes are inimical to global welfare and cannot be justified from a cosmopolitan standpoint.

Mill's Failure

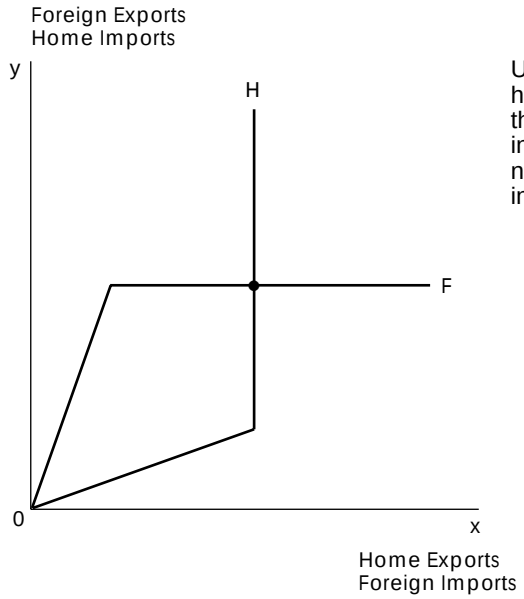
The bulk of Mill's analysis remains as valid today as when he wrote it. Still, he was not completely successful. He failed to resolve the problems that arise when offer curves exhibit (1) multiple equilibria and (2) indeterminacy of equilibrium. The first problem arises when the curves intersect more than once; the second when they coincide over certain ranges. Both phenomena require for their occurrence inelastic offer curves. Unfortunately, however, Mill chose to analyze them under the special assumption that the curves are unit-elastic. Neither Alfred Marshall ([1879] 1975, pp. 148–49) nor Francis Edgeworth (1894a, pp. 609–14) let this slip pass unnoticed. They pointed out that unit-elastic curves intersect only once and cannot coincide (see Figure 4). Accordingly, they concluded that Mill's choice of unit-elastic curves was useless in resolving questions of indeterminacy and multiple equilibria.

Mill's Paradox

Mill was successful, however, in using his unit-elastic schedules to demonstrate what Akira Takayama (1972, pp. 144–45) calls "Mill's paradox." That paradox states that a country's gains from trade decline as its resource endowment expands.

Let the offer curves be unit-elastic beyond the production-specialization points on the comparative-cost lines (see Figure 5). Suppose that prior to resource expansion the foreign curve initially cuts the home curve at its kink. The result is that the terms of trade coincide with the home country's cost ratio and the foreigner reaps all the gains from trade.

Figure 4 Unit-Elastic Reciprocal-Demand Curves

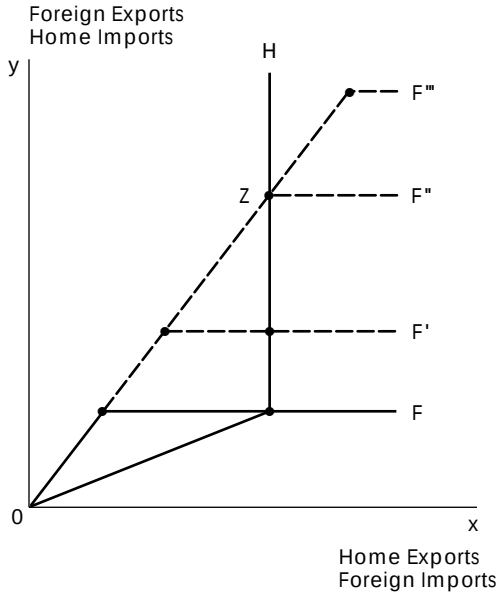


Mill then assumes that resource expansion occurs in the foreign country. Such expansion, provided it raises the output of exportables more than it raises the foreigner's own demand for them, shifts upward his production (export-capacity) point and with it his offer curve. The resulting growth-augmented curve cuts the unit-elastic segment of the home country's curve, thus yielding a terms-of-trade deterioration for the foreigner. In the limit, growth continues until the terms of trade coincide with the foreigner's cost ratio and all trade gains accrue to the home country. Here is the rationale for Mill's statement that "the richest countries, ceteris paribus, gain the least by a given amount of foreign commerce" ([1848] 1909, p. 604).

Assessment

Overall, Mill's analysis must be judged one of the greatest contributions in the history of economics. It generalized classical value theory by shifting the emphasis from cost of production to equilibrium of demand and supply. True, Mill's predecessors occasionally acknowledged demand as a determinant of price. But they did so only for the singular case of nonreproducible goods in absolutely fixed supply. Mill now extended that analysis to cover labor-produced goods as well. He showed that even if cost determines the autarkic

Figure 5 Mill's Paradox



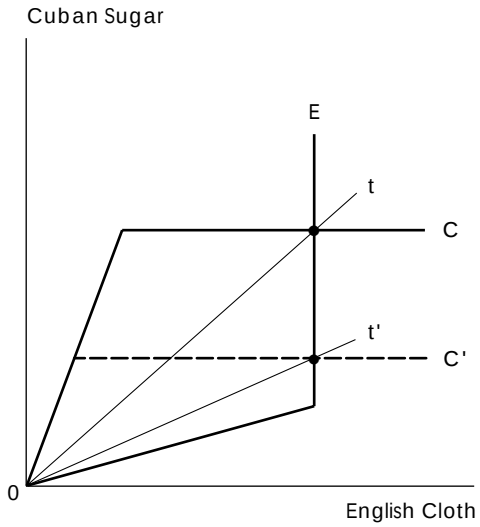
Growth of the foreigner's resource endowment shifts his offer curve upward. His terms of trade deteriorate until they eventually coincide with his cost ratio. Beyond point Z he ceases to specialize.

value of such goods, as Ricardo claimed, another principle, namely that of reciprocal demand, determines their international value. By distinguishing between cost-determined domestic prices and international prices determined jointly by supply and demand, he identified both blades of the Marshallian scissors. True, it remained for the neoclassical school to elaborate his insight into a full microeconomic theory of price determination. But Mill clearly pointed the way.

2. ROBERT TORRENS

Mill had shown that with unit-elastic home demand for imports, the foreigner's trade tax improves his terms of trade equiproportionally with the tax. Two years before Mill published his analysis, Robert Torrens (1841–42) independently reached this same conclusion in a numerical example presented in his Postscript to Letter IX of *The Budget*. There he showed that a 100 percent import tariff, through its effect on reciprocal demands, improves the levying country's terms of trade by the same 100 percent. His example has the foreign country, Cuba, imposing the tariff on imports of English cloth. His assumption of unit-elastic reciprocal demands implies that Cuba's offer curve is horizontal and England's curve is vertical in the relevant range (see Figure 6).

Figure 6 Torrens' Tariff Model

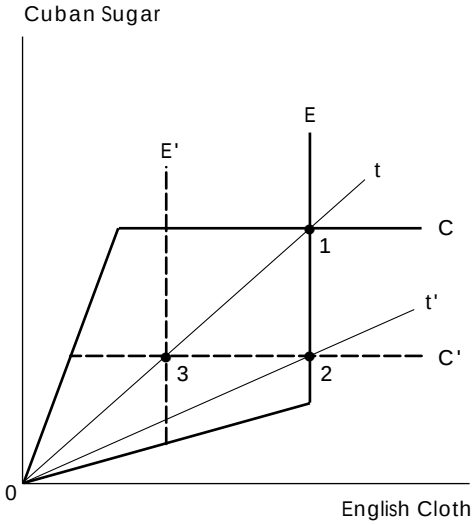


Cuba's 100 percent tariff shifts her offer curve down from *C* to *C'*. Her terms of trade improve equiproportionally such that the slope of ray 0*t'* is half that of ray 0*t*.

Start from the free-trade equilibrium. Cuba's imposition of the 100 percent tariff shifts her effective offer curve down to half its initial level. At the original terms of trade, there occurs an excess world demand for Cuba's export good, sugar, and a corresponding excess world supply of her import good, English cloth. To eliminate these excess demands and supplies, Cuba's terms of trade must improve—and England's deteriorate—by 100 percent. In the new, tariff-ridden equilibrium, Cuba imports the same initial amount of cloth at the cost of only half the initial amount of sugar given up. England, on the other hand, receives only half the initial amount of sugar at the cost of the same amount of cloth sacrificed. Torrens' conclusion: Foreign governments can, by means of tariffs, manipulate reciprocal demands to their advantage and thereby worsen the other country's terms of trade.

Having shown how the home country might lose from the foreigner's tariff, Torrens next used his reciprocal-demand schedules to argue for reciprocity in tariff removal (see Figure 7). He pointed out that the home country's unilateral abolition of tariffs would, in the face of their existence abroad, only worsen her terms of trade. He likewise noted that the home country's retaliatory duties would cancel the unfavorable terms-of-trade effects of foreign levies. Finally, he observed that the simultaneous imposition or removal of duties by all countries tends to leave the terms of trade unchanged. Like today's proponents of "a level playing field," he proposed that England counter foreign tariffs with

Figure 7 Tariff-Policy Reciprocity



Cuba's unilateral tariff improves her terms of trade. Equilibrium moves from point 1 to point 2. England's reciprocal duty shifts her offer curve leftward and corrects her terms-of-trade deterioration (point 2 to point 3). Simultaneous tariff removal by both parties restores initial equilibrium without affecting the terms of trade (point 3 to point 1).

equal duties of her own, that she trade freely only with countries admitting her goods duty-free, and that she drop her tariffs only insofar as her trading partners abolish theirs.

Torrens' analysis was unsympathetically received by his classical contemporaries who feared it would undermine the case for free trade (see O'Brien [1975], pp. 194–97). They noted that tariff removal would hardly worsen England's terms of trade to the extent Torrens claimed if reciprocal-demand elasticities were, as they believed, greater than one.[7] Furthermore, they contended that any adverse terms-of-trade effects of moving toward freer trade would be more than offset by gains in productivity and competitiveness due to enhanced international specialization and division of labor. Finally, they noted that the gist of Torrens' analysis implied that England should levy not equal but higher retaliatory duties than those levied abroad to improve her terms of trade. They saw such action as intensifying the danger of a trade war with all parties

[7] To Torrens' critics, high elasticities stemmed from the availability of numerous alternative goods and markets in the world economy. Such availability meant that a country could avoid tariffs levied by its trade partner. The country could divert production from taxed to nontaxed exports. Or it could export to third, tariff-free countries, competition from which would limit the tariff-imposing country's power to manipulate the terms of trade. In short, access to multiple export outlets and import sources rendered reciprocal demands extremely elastic with respect to price changes emanating from any single source.

losing. In response to the criticism that his analysis aided protectionists, Torrens rather disingenuously protested that he was a free trader merely applying the logic of the classical model.

3. ALFRED MARSHALL

Without actually drawing the offer-curve diagram, Mill and Torrens had described the workings of its principal components. Indeed, Mill's account, in the words of Joseph A. Schumpeter (1954), "reads almost like a somewhat clumsy instruction for choosing these curves rather than others" (p. 609). It was Alfred Marshall who took the crucial step of translating Mill's instructions into geometry and thus invented the diagram that bears his name.

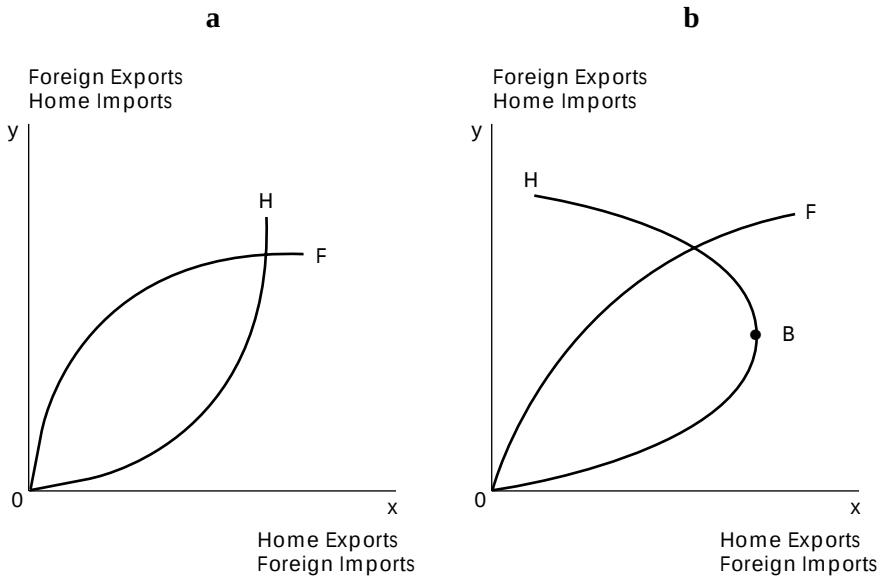
That Marshall was the first to draw the diagram is beyond dispute. What is disputed is the originality of his contribution. Did his trade diagrams do no more than merely "polish and develop Mill's meaning," as Schumpeter (1954, p. 609) claimed? Or were they "of such a character in their grasp, comprehensiveness and scientific accuracy" as to put them "far beyond the 'bright ideas' of his predecessors," as John Maynard Keynes ([1925] 1956, p. 24) thought? Marshall himself disclaimed originality by stressing the Millian pedigree of his diagrams. "As to International Trade curves," he wrote, "mine were set to a definite tune, that called by Mill" (Pigou [1925] 1956, p. 451). He dismissed his curves as nothing more than "a diagrammatic treatment of Mill's problem of international values" (Pigou [1925] 1956, p. 416).

Of the fourteen diagrams Marshall presented in his *Pure Theory of Foreign Trade* (1879), at least five appear to confirm Schumpeter's and Marshall's judgments. For they merely elaborate in elegant and compact geometry what Mill had already expressed in words and numerical examples. Certainly Mill would have found unexceptional the curves in Figure 8a just as he would Marshall's explanation of their convex (bowed in toward their respective axes) shapes and their positive slopes. Their convexity, Marshall held, captures the inverse demand relationship between import price and quantity. And their positive slopes indicate the normal case of elastic demands in which import-sales proceeds—and thus the quantity of exports produced with the aid of those proceeds—rise with import-quantity demanded.[8]

Nor would Mill have been surprised by Figure 8b. There Marshall depicts a case of inelastic import demand as manifested in a backward-bending offer curve. Mill would have agreed with Marshall that beyond point *B*, import-sales proceeds, and so the export volume they finance, must fall as import-quantity demanded rises. He had said much the same thing himself.

[8] Marshall always assumed that the price-times-quantity sales receipts of importers pay for the cost of exports. Import receipts finance export production.

Figure 8 Marshall's Normal Class and Exceptional Class I Curves

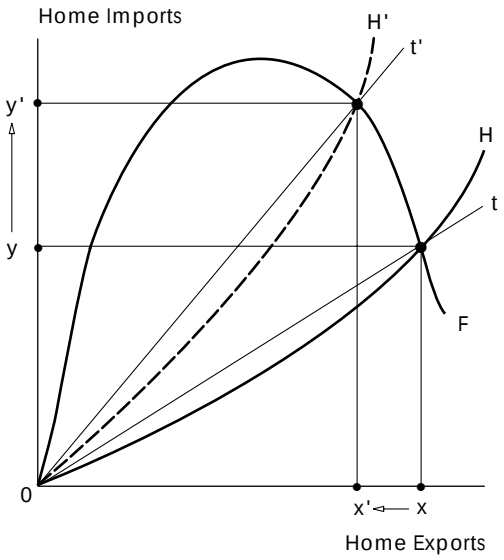


In panel a, both curves are elastic. Import and export quantities rise as the relative price of imports falls. In panel b, the home curve becomes inelastic beyond point *B*. Import quantity rises but export quantity falls with decreases in the relative price of imports.

Finally, Mill would have found Marshall's diagrammatic treatment of trade taxes totally unsurprising. Marshall showed that when both offer curves are elastic (provided the foreigner's is not infinitely so), tariffs and export taxes always improve the levying country's terms of trade. He also showed that when the levying country's curve cuts the foreigner's curve in its inelastic range, a trade tax yields a twofold gain (see Figure 9). The taxer's terms of trade improve. And, by obtaining a larger quantity of imports at the sacrifice of a smaller quantity of exports, the taxer has more of both goods to consume at home. A country lucky enough to face an inelastic foreign offer curve, said Marshall, has nothing to lose and everything to gain by exploiting it.[9] But Mill had already arrived at these conclusions. Thus Marshall's diagrammatic tax analysis goes little beyond Mill's work on that subject.

[9] Marshall assigned a low probability to this case. He thought that (1) international competition, (2) countries' ability to shift production from taxed to nontaxed exports, and (3) the option of trading with third, free-trade nations rendered offer curves highly elastic. Levying countries were left with little scope for tariff-induced improvements in the terms of trade.

Figure 9 Exploiting an Inelastic Foreign Offer Curve



The home country imposes a trade tax that shifts her offer curve counterclockwise. Her terms of trade improve. And she obtains more imports at the cost of fewer exports given up and so has more of both goods to consume at home.

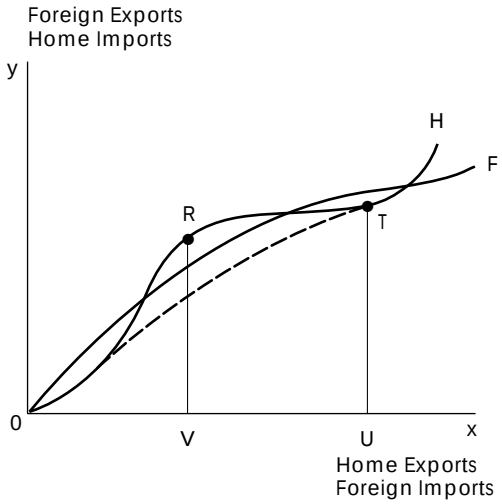
Scale Economies and Offer Curves

Figures 10 through 13, on the other hand, go far beyond anything Mill or Torrens had to offer. Figure 10 constitutes what John Chipman (1965) calls "the first fairly rigorous approach to the treatment of scale economies in international trade" (p. 738). It depicts Marshall's Exceptional Class II, in which economies in the production of exportables render the offer curve nonconvex and subject to irreversible downward shifts.

Let trade expansion move the home country from point *R* to point *T* on its offer curve. The resulting increased export production invokes scale economies associated with enhanced specialization and division of labor, with improved know-how (learning by doing), and with use of advanced technology and large machines. These economies in turn enable the larger quantity of exports to be produced at lower unit cost than the original quantity. Since unit-cost reductions pass through into product prices, it follows that scale economies cause export prices to fall from *RV*/0*V* to *TU*/0*U*. Such economies account for the inflection points on the offer curve.

Moreover, once the scale economies are put in place, they and their associated cost reductions cannot be reversed even if output drops back to its original level. To capture such irreversible path-dependent effects, the offer curve shifts downward toward the export axis (see dashed line). In short, scale economies

Figure 10 Marshall on Scale Economies



Because of irreversible scale economies, movement from point *R* to point *T* causes a reduction in the unit cost of producing exports and a downward shift in the offer curve.

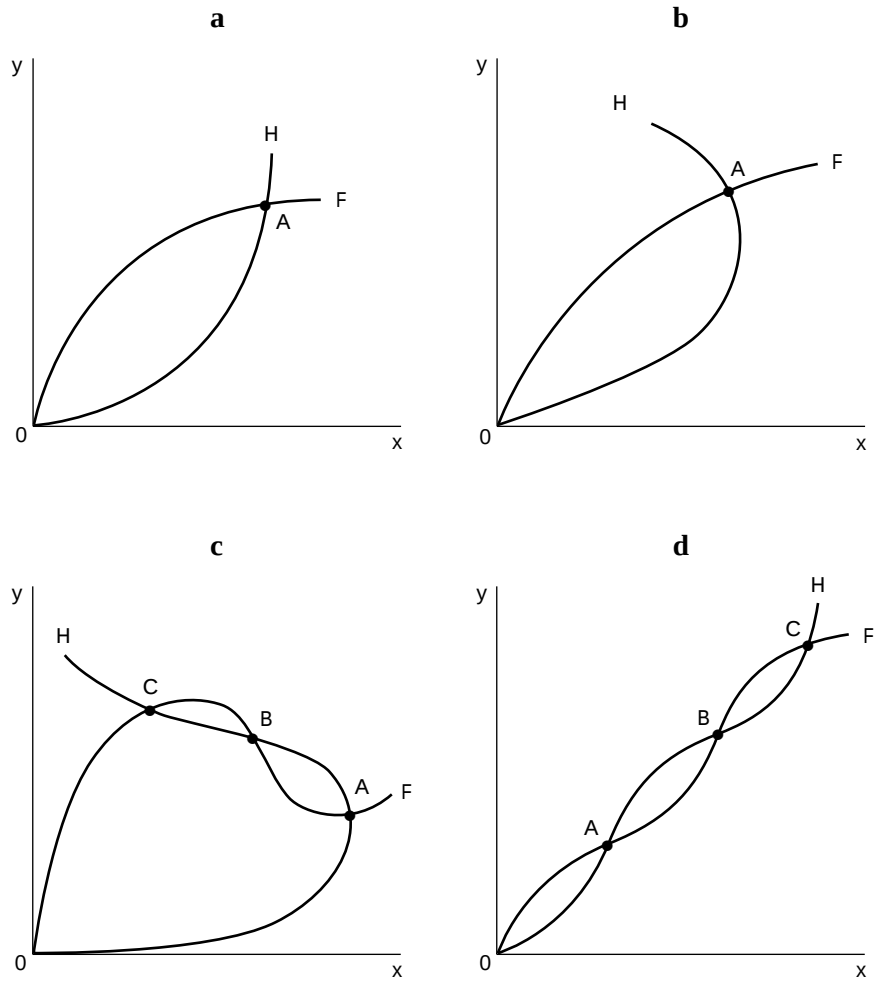
constitute a form of technical progress that shifts the offer curve simultaneously with movements along it. Here was a novel element in offer-curve analysis.

Uniqueness and Stability of Equilibrium

Further confirming Marshall's originality was his analysis of uniqueness and stability of offer-curve equilibrium. Indeed, Murray Kemp (1964) calls this analysis one of "the most remarkable contributions ever made to theoretical economics" (p. 66). Regarding uniqueness, Marshall noted that there can be but one equilibrium when both curves are positively sloped and possess no inflexion points (see Figure 11a). Equilibrium is likewise singular when one curve is elastic and the other inelastic in a certain range (see Figure 11b). If, however, both curves are inelastic (Figure 11c), or if at least one contains inflexion points (Figure 11d), multiple equilibria may result. Such equilibria, according to Marshall, are always odd in number. Moreover, they are alternately stable and unstable, with the stable equilibria flanking the unstable ones (see Figures 11c and d in which stable equilibrium points *A* and *C* flank the unstable point *B*).

As for stability of equilibrium, Marshall analyzed it with phase diagrams superimposed on his offer curves. His phase diagrams—the first ever used in print by an economist—treat points off the curves as disequilibrium phenomena produced by random real shocks such as wars, harvest failures, and the like. For any given disequilibrium trading point, Marshall sketched the dynamic

Figure 11 Single and Multiple Equilibria

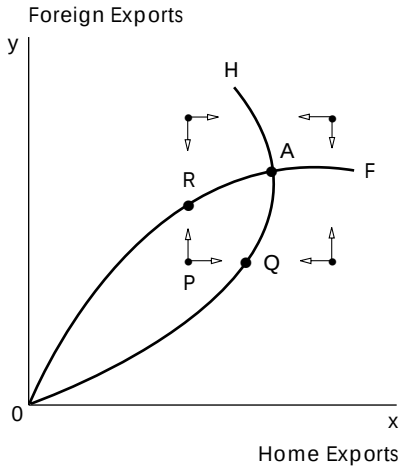


Panels a and b display cases of unique (singular) equilibrium. Panels c and d display multiple equilibria, with the stable equilibria (*A* and *C*) bounding or bracketing the unstable one (*B*).

adjustment mechanism that moves the point toward equilibrium. To him, the propelling force consisted of the profitability of expanding production of exports when they are in short supply.

Consider trading points to the left of the home country's curve and below the foreigner's curve. Such points represent shortfalls of actual quantities of

Figure 12 Marshall's Disequilibrium Mechanism



The shortfall of actual below desired exports at trading point *P* stimulates production of exports and so moves that point toward equilibrium.

exports below quantities the countries are willing to offer (see Figure 12). These shortfalls render exports extraordinarily profitable and induce competitive producers to produce more. The resulting export expansion moves the trading point toward the curves just as—to use Marshall's analogy—the force of magnetic attraction moves metal filings toward a rigid wire. The arrows point rightward and upward to indicate the trading point's movement.

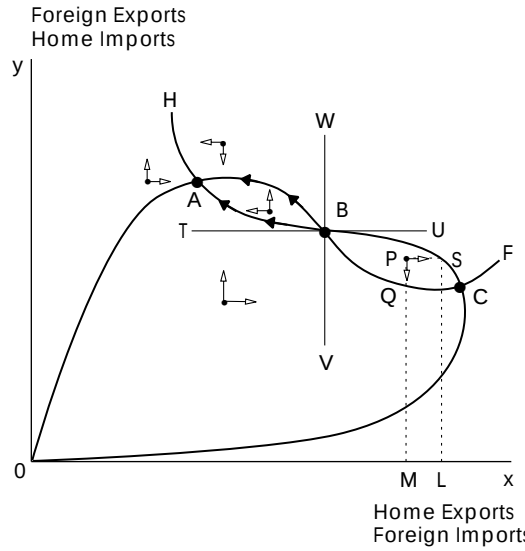
Conversely, disequilibrium trading points to the right of the home country's curve and above the foreigner's curve spell surpluses of actual over desired exports. The resulting losses bring declines in export production as shown by the leftward and downward direction of the arrows.

In all cases, the directional arrows indicate whether trading points will move away from or toward their neighboring equilibria (see Figure 13). On the basis of such analysis, Marshall concluded that every equilibrium intersection is stable except those in which (1) both curves slope in the same direction and (2) the foreign curve is more nearly vertical than the domestic one (see Amano [1968], pp. 327–28).

Surplus Analysis of Gains from Trade

Further proof of Marshall's originality is his diagrammatic treatment of the gains from trade. In a straightforward application of his concept of consumer surplus, he expressed such gains as the excess of the maximum prices a country would be willing to pay for successive units of imports over the market price, or terms of trade, it actually pays. Accordingly, he devised a technique

Figure 13 Marshall's Phase Diagram



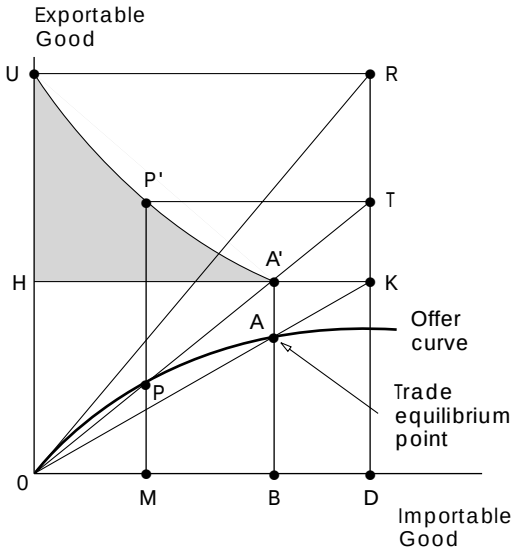
for projecting from the offer curve a series of unit surpluses into a triangle resembling the area lying between an ordinary Marshallian demand curve and the price line (see Figure 14). Expressed in terms of export quantities, the resulting triangular area UHA' sums the excesses of the maximum unit prices over the prevailing terms of trade shown by the slope of the ray through the trade equilibrium point *A*. This was his measure of the net benefit a country derives from trade.[10]

Assessment

It should be obvious by now that Keynes was right. Marshall's diagrams were more than a mere refinement of Mill's analysis. They were a major innovation and a powerful aid to theorizing. The mystery is why Marshall himself refused to acknowledge as much. Perhaps a desire to stress the intellectual continuity of trade theory led him to disguise his contribution modestly as part of the accumulated wisdom of his classical predecessors. Or perhaps his reluctance to claim originality for his diagrams stemmed from a puritanical sense of guilt over the pleasure he derived from them. Jacob Viner (1941) writes that mathematics,

[10] Actually, Marshall divided the surplus triangle by the distance $0D$ to correct for the arbitrary choice of point *D*. That point fixes the location of the vertical line used in projecting unit price surpluses onto the quasi-demand curve $UP'A'$.

Figure 14 Surplus Measure of the Gain from Trade



The maximum price the country is willing to pay for successive units of imports ranges from the slope of the top ray to the slope of the bottom one. A mapping of the excess of these unit prices over the equilibrium terms of trade forms the surplus triangle *UHA'*. The net gain from trade is measured by the triangle *UHA'* divided by the distance 0*D*, which corrects for the arbitrary choice of the location of the projection line *DR*.

especially geometry, yielded Marshall "so much intellectual and aesthetic delight that it for that reason alone became somewhat suspect to him as a worthwhile occupation. Mathematics, and especially graphs, were Marshall's fleshpots, and if he frequently succumbed to their lure it was not without struggle with his conscience. . . . [W]hen he did succumb he . . . warned his readers not to take his mathematical adventures too seriously" (p. 231). Keynes agreed. He pointed out that when Marshall's "intellect chased diagrams and Foreign Trade and Money, there was an evangelical moralizer of an imp somewhere inside him, that was so ill-advised as to disapprove" ([1925] 1956, p. 37). But disapprove Marshall did and in so doing disclaimed originality for his invention.

4. FRANCIS YSIDRO EDGEWORTH

As sophisticated as they were, Marshall's offer curves lacked clear grounding in the underlying utility functions. Credit for establishing these foundations and for introducing utility considerations into the diagram goes to F. Y. Edgeworth, Marshall's colleague and the inventor of the indifference map and the contract curve.

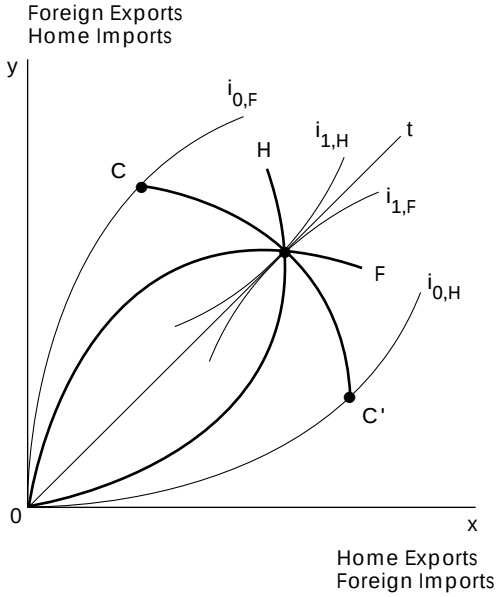
Edgeworth's earliest work on the diagram appears on pages 113–14 of his *Mathematical Psychics* (1881). There he derived offer curves for two representative price-taking individuals. Each possesses (1) a fixed endowment of

goods and (2) a utility function described by a consumption indifference map. Like trade theorists today, Edgeworth defined each individual's offer curve as the locus of points of tangency of indifference curves and the price ray as it pivots about the origin (see Figure 15). Each point represents an outcome of constrained utility maximization in which the price ratio, or slope of the price ray, equals the ratio of marginal utilities, or slope of the indifference curves.

In the same diagram, he demonstrated that the offer curves must intersect on the contract curve, or locus of points at which one trader's indifference curves are tangent to the other's. Along the contract curve, neither trader's utility can be increased without reducing the other's. In demonstrating that offer curves intersect on the contract curve, Edgeworth proved that price-taking equilibrium is efficient in the sense that both traders together cannot be made better off by another outcome. Moreover, since the equilibrium outcome lies between the two indifference curves passing through the origin, or endowment point, where no trade occurs, he also proved that free trade leaves each party at least as well off as no trade.

He next extended this latter insight to measure a single country's gain from trade. In so doing he provided an alternative to Marshall's measure of the gain. In his 1889 *Journal of the Royal Statistical Society* article, "On the

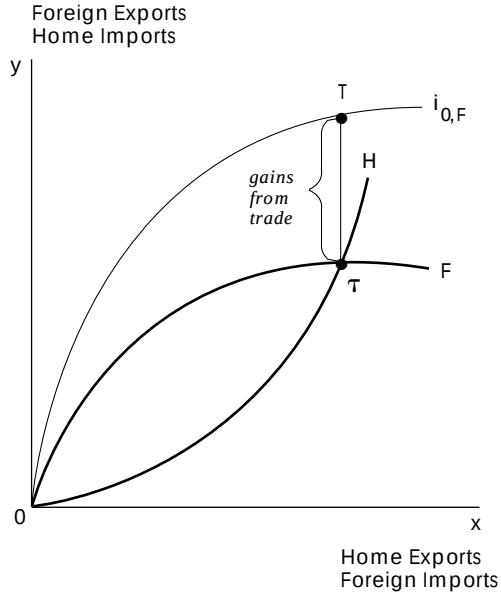
Figure 15 Traders' Price-Taking Equilibrium



Offer curves intersect on the contract curve *CC'*. At free-trade equilibrium, both traders occupy indifference curves superior to those going through the origin or endowment point where no trade occurs.

Application of Mathematics to Political Economy," he drew a community trade indifference curve passing through the origin (see Figure 16). This curve, of course, shows all combinations of exports and imports that leave the country no better off than if it refrained from trade. The vertical distance between this "no gain from trade" curve and the country's offer curve—or rather the indifference contour going through it—at free-trade equilibrium measures the utility gain from trade.

Figure 16 Utility Gain from Trade

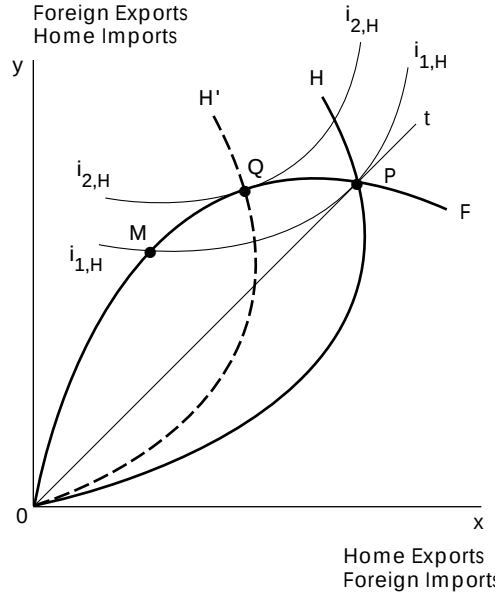


The vertical distance between the trade equilibrium point and the foreigner's "no gain from trade" indifference curve–the particular curve passing through the origin–constitutes the foreigner's utility gain from trade.

Optimum-Tariff Analysis

The foregoing pathbreaking innovations were but a prelude to Edgeworth's crowning achievement—his geometrical demonstration of the optimum-tariff argument. An optimum tariff, of course, maximizes the excess of the gains from terms-of-trade improvement over the loss from lower trade volume and reduced international specialization and division of labor. The idea itself goes back to Mill and Torrens. They had argued that a large country caring only for its own welfare and facing an imperfectly elastic foreign demand schedule could exploit its monopoly power in world markets through such a tariff. But rigorous diagrammatic illustration of the argument was lacking until Edgeworth provided it in his 1894 *Economic Journal* article, "The Theory of International Values, II."

Figure 17 The Optimum Tariff



The optimum tariff puts the levying country on the point where its highest attainable trade indifference curve is tangent to the foreigner's offer curve.

His demonstration begins with the home country's trade indifference curve passing through the free-trade point where the offer curves intersect (see Figure 17). That particular indifference curve indicates the level of welfare or satisfaction the home country enjoys under free trade. It provides a benchmark against which to compare alternative welfare levels produced by different degrees of trade restriction.

It also, together with the foreign offer curve, specifies the range of tariff rates beneficial to the home country. In this connection, Edgeworth noted that the same indifference curve that passes through the free-trade point *P* also cuts the foreign offer curve at point *M*. That latter point therefore yields the same level of welfare as free trade. Since all points on the foreign offer curve between these two extremes lie on higher indifference curves, it follows that any movement to a position between points *P* and *M* will result in the home country being better off than under free trade. In other words, points *P* and *M* mark the range of tariff-induced terms-of-trade improvement beneficial to the home country. Somewhere within this range, benefit is at a maximum.

Edgeworth identified this maximum with point *Q*, the point where the home country reaches its highest possible trade indifference curve given the foreign offer curve. The optimum tariff, he argued, is that which distorts the home

country's offer curve such that it intersects the foreign offer curve at this point of tangency with the highest attainable indifference curve. Here is the famous tangency solution to the determination of the optimum tariff.

Edgeworth then showed that if the tariff is too high, it reduces rather than increases welfare. Suppose the country progressively raises its tariff from the zero rate corresponding to point *P* to positive rates corresponding to points *Q* and *M*. As it does so, it finds that its welfare first rises, then reaches a maximum, and finally starts to fall. If it persists in raising the rate beyond that corresponding to point *M*, it will discover that its welfare has fallen below the level attained at the free-trade position *P*. It follows that the tariff must not be too large if the nation is to benefit.

Finally, Edgeworth noted some pitfalls to the practical application of his diagram. First, the optimum tariff, though precisely identified in theory, cannot be ascertained with any accuracy in practice. Second, protectionists will exert strong political pressure on policymakers to raise tariffs far beyond the optimum point, thereby reducing welfare. Third, retaliation by foreign countries may erase any gains generated by the tariff. Fourth, viewed from a global standpoint, tariffs are harmful since other countries lose more than the levying country gains. For these reasons, free trade remains the best and most practical policy for a country to pursue.

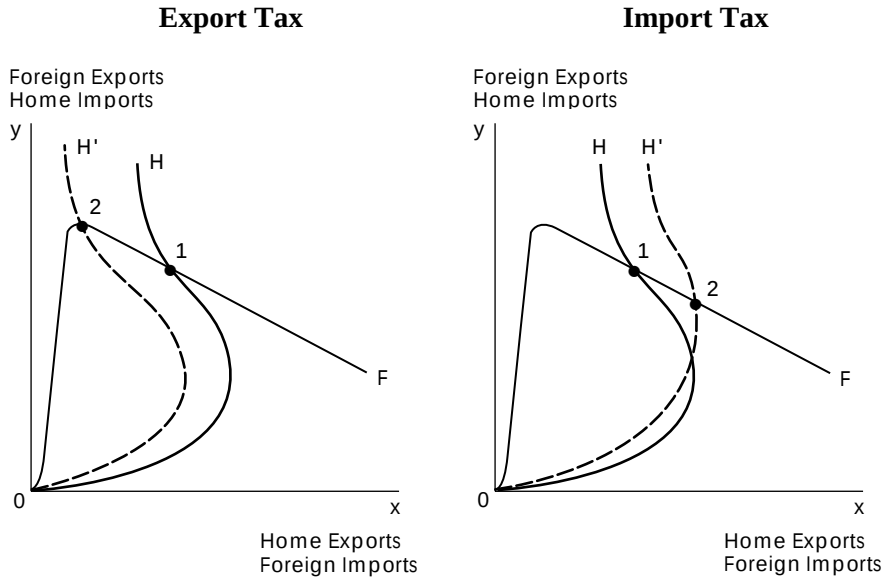
Alleged Asymmetry of Export and Import Taxes

No scholar is infallible, not even one of Edgeworth's stature. In the very same *Economic Journal* article containing his optimum-tariff demonstration, Edgeworth (1894b) committed a celebrated error. He rejected the standard proposition that export and import taxes are equivalent in the sense of having identical real effects. Other leading classical and neoclassical theorists, including Mill, Marshall, A. C. Pigou, and C. F. Bastable, took such equivalence for granted. But Edgeworth alleged that the two taxes shift the dutying country's offer curve differently and therefore have disparate real effects.

According to Edgeworth, export taxes shift the curve horizontally to the left. But import taxes shift it vertically upward such that it lies everywhere above the original curve. The result is that the tax-ridden curves intersect the foreign offer curve at different points, especially when both foreign and domestic curves are in their inelastic ranges (see Figure 18). In such cases, the export tax-ridden equilibrium lies to the northwest of the free-trade point so that the levying country is on a higher indifference curve with better terms of trade. By contrast, the import tax-ridden equilibrium lies to the southeast of the free-trade point, putting the country on a lower indifference curve with worsened terms of trade. Such was Edgeworth's allegation.

It took 42 years to identify and correct Edgeworth's error. Abba Lerner finally did so in his classic 1936 paper, "The Symmetry Between Import and

Figure 18 Alleged Asymmetry of Export and Import Taxes



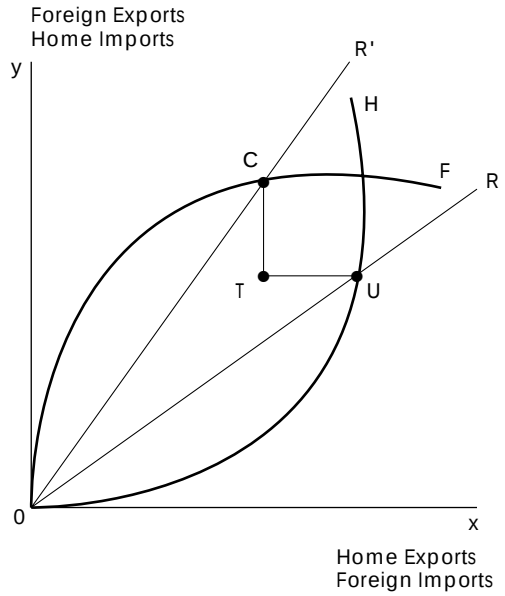
Edgeworth alleged that an export tax shifts the home country's offer curve horizontally to the left whereas an import tax shifts it vertically. As a result, the export tax moves the home country to a superior position (better terms of trade and a higher indifference curve), whereas the import tax moves her to an inferior position.

Export Taxes." There Lerner argued, contrary to Edgeworth, that export and import taxes indeed affect the offer curve identically. They thus have symmetrical effects on the volume and the terms of trade. Differential effects stem not from the taxes per se. Rather they stem from how the government disposes of the revenue. The greater the proportion spent on the levying country's export good, the greater the improvement in its welfare and terms of trade. Conversely, the greater the proportion spent on imports, the smaller the improvement. Edgeworth's first result obtains when all the proceeds are spent on export goods; his second when all are spent on imports. His error lay in confusing these expenditure effects for tax effects. What he saw as differential results of trade taxes were really outcomes of how the government spent the revenue.

5. ABBA LERNER

Lerner (1936) established the foregoing results by means of the ingenious device of a geometrical *pencil*, or wedge, superimposed on the offer curves (see Figure 19). Consisting of two price radiants, the pencil expresses the tax-induced divergence between world and domestic relative prices. Its width shows the rate of the tax. Its location on the offer curves depicts the government's apportionment of the proceeds between exportable goods and imports. And its position around the free-trade price ray shows how tax imposition and disposal affects the terms of trade and domestic relative prices. Finally, the pencil embodies the symmetry notion that export taxes are equivalent to import taxes of

Figure 19 Trade-Tax Pencil



The pencil *R'* 0*R* represents the tax wedge between world and domestic relative prices. The pencil's dimensions are the same whether the tax is levied on imports or exports. The arms of right angle *CTU* correspond to the proportions of the tax revenue that the government spends on the two goods.

the same percentage rate. Since both taxes produce the same divergence between world and domestic prices in a two-good model, the pencil's dimensions are the same measured in either tax.[11] What matters is not which good is taxed but how the government disposes of the tax proceeds.

[11] A tax on imports renders them dearer at home than abroad. By contrast, a tax on exports raises their foreign price, thus making them cheaper at home than abroad. But a fall in the relative price of exports is equivalent to a rise in the relative price of imports in Lerner's two-good model. Hence, an export tax raises the domestic real price of imports above the world price just as does an import tax. The two taxes are symmetrical.

Lerner's demonstration of this point was at once seminal and definitive. Inserting the pencil into the offer curves, he obtained a right angle *CTU* connecting the points of entry of the pencil's radiants. This right angle has a special meaning. Its vertical and horizontal arms measure the world excess supplies of the two goods resulting from the tax. Market-clearing equilibrium requires that the government eliminate these excess supplies by consuming them in the proportion in which they occur. That is, world market equilibrium obtains when the ratio of the lengths of the right angle's arms matches the ratio in which the government consumes the two goods.

Lerner took this latter ratio as given and known. Then he found equilibrium by pivoting the pencil about the origin until the matching right angle appeared. For example, suppose the government spends the tax proceeds equally on exports and imports. Then, following Lerner, swing the pencil until it yields a right angle whose arms are of equal length (see lines *CT* and *TU* in all panels of Figure 20). Alternatively, suppose the government spends all the proceeds on exportables such that the right angle reduces to a horizontal line. Then rotate the pencil counterclockwise until it yields a flat line traversing the pencil and meeting the offer curves (see lines C_1U_1 in all panels). Finally, suppose the government spends all the proceeds on imports so that the right angle reduces to a vertical line. Then swivel the pencil clockwise until it yields a vertical line between the offer curves (see lines C_2U_2 in all panels). In each case, Lerner examined the resulting location of the pencil's radiants relative to the free-trade price ray passing through the point of offer-curve intersection. These indicate how the disposition of the tax affects the terms of trade and domestic price ratio, respectively. Radiants to the left of the free-trade ray represent a fall and those to the right a rise in post-tax relative prices.

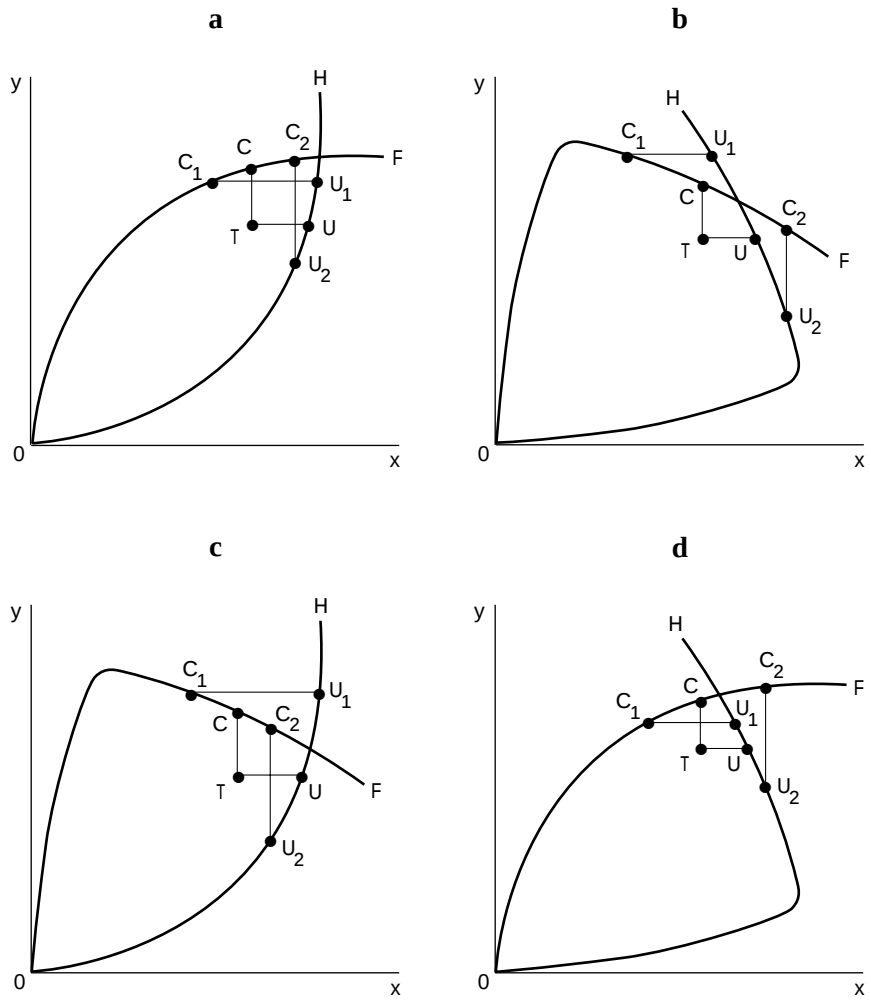
Standard Tariff Propositions

Employing this technique, Lerner derived four key propositions of standard tariff theory. His derivation marks a turning point in the diagram's history. Before him, the diagram was largely regarded as an esoteric tool employed by a select circle of economists. After him, it was seen as a conventional instrument and widely used. His work, more than any other, convinced the economics profession of the diagram's power and versatility as an analytical tool.

His standard propositions are as follows. First, provided both offer curves are elastic but not infinitely so, a trade tax, no matter how spent, improves the levying country's terms of trade and raises the domestic relative price of its imports (see Figure 20a). In other words, the radiants of the pencil encompass the free-trade price ray.

Second, regardless of elasticities, a tariff improves the terms of trade more (or worsens it less) the larger the fraction of the tax spent on the country's exportable good. Thus, the pencil's upper radiant lies more to the left (or less

Figure 20 Lerner's Trade-Tax Analysis



The pairs of *C*-*U* points show where the upper and lower radiants of the pencil enter the offer curves. Pairs C_1-U_1, *C*-*U*, and C_2-U_2 correspond to tax proceeds allocated all, some, and none, respectively, to exportables. Taxes and their disposal improve the terms of trade when radiants through the *C* points lie to the left of the free-trade ray (not shown) that passes through the point of offer-curve intersection. Taxes and their disposal raise the domestic relative price of imports when radiants through the *U* points lie to the right of the free-trade ray.

to the right) of the free-trade ray as it passes through points C_1, C, and C_2, representing export expenditure shares of one, half, and none, respectively (see Figure 20, all panels).

Third, suppose the taxing country spends all the tax on imports and possesses an inelastic import demand, or backward-bending offer curve. In this case, a trade tax actually worsens the terms of trade. Geometrically, the radiant passing through point C_2 lies to the right of the free-trade ray (see Figure 20b and d). Indeed, one can relax the assumption that all the tax is allocated to imports. Lerner's result holds as long as the taxer's import-demand elasticity is less than the fraction of the proceeds spent on imports.[12]

Fourth, assume the home country faces an inelastic foreign offer curve and spends its tariff proceeds largely on its exportable good. In this case, a tariff may improve the terms of trade by more than the tariff such that the domestic price of imports falls below its free-trade level. Geometrically, the radiant passing through point U_1 lies to the left of the free-trade ray (see Figure 20b and c). If so, the tariff achieves the opposite of its intended purpose. By lowering the domestic price of imports, it harms rather than protects domestic import-competing industries and the relatively scarce factors they employ intensively. Today, textbooks attribute this paradoxical result to Lloyd Metzler. He proved, in 1949, that it holds when the dutying country's marginal propensity to spend the tariff proceeds on its own export good exceeds the foreigner's elasticity of demand for that good.[13] But it was Lerner, not Metzler, who first established this proposition.

6. CONTROVERSIES IN THE 1920s AND 1930s

Offer curves also constituted the focus of Frank Graham's (1923, 1932) celebrated critique of Marshall's work. Graham's critique raised issues not fully resolved until the 1950s.

The first issue concerned the effects of demand-induced shifts in the home country's curve. In Appendix J of his 1923 *Money, Credit and Commerce*, Marshall analyzed such shifts stemming from autonomous increases in import demand. He argued that the resulting extent of terms-of-trade deterioration would vary directly with the home import-demand elasticity and inversely with

[12] A high import-expenditure fraction augments the demand for imports and tends to raise their relative price. But a low elasticity spells little offsetting fall in import-quantity demanded in response to the higher price. The resulting excess demand for imports raises their price and causes the terms of trade to deteriorate.

[13] The government's high propensity to spend on its exportable puts upward pressure on the domestic (and world) price of that good. But the low foreign demand elasticity militates against offsetting falls in quantity demanded abroad in response to the higher price. The net result is an excess demand for exportables. This excess demand raises the relative price of exports and lowers its inverse, the relative price of imports.

the foreign one. Similarly, he thought the accompanying degree of expansion in trade volume would vary directly with both elasticities.

Graham, however, disagreed. He thought that the extent of terms-of-trade deterioration would vary inversely with both elasticities. He also thought trade-volume expansion would vary directly with the foreign elasticity but inversely with the home elasticity.

How could two leading economists differ over something as elementary as the effects of shifts of offer curves? Murray Kemp supplied the answer in 1956. The disagreement stemmed from ambiguity of the phrase "increase in reciprocal demand." More precisely, it stemmed from Marshall's failure to state explicitly the type of shift postulated. It turns out that he implicitly posited horizontal shifts due to increases in the quantity of exports offered against a given quantity of imports. By contrast, Graham posited equiproportional or radial shifts due to increases in the quantity of exports offered at given terms of trade. Both were right in terms of their own implicit definitions. Still, the controversy taught the economics profession a lesson. Elasticity affects the extent to which demand shifts alter price and quantity. Exactly how it does so depends on the precise definition of such shifts (see Bhagwati and Johnson [1960], p. 78).

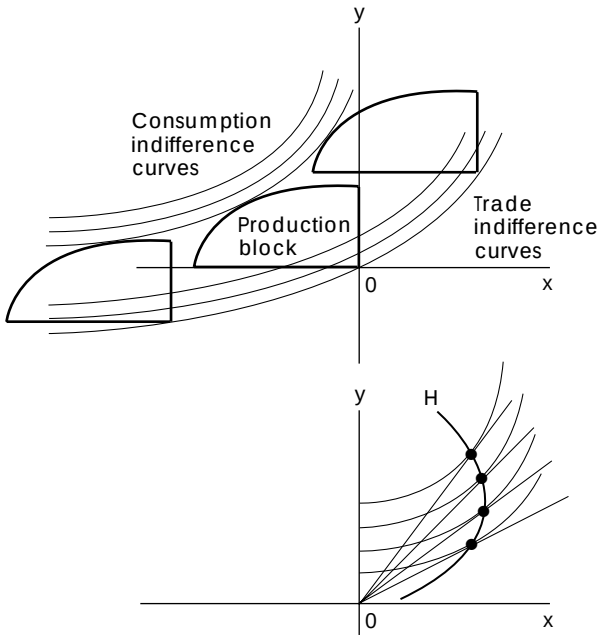
The second issue concerned the link between offer curves and the underlying production conditions. Graham accused Marshall, Mill, and their followers of neglecting these conditions and overemphasizing demand. But this accusation was hardly fair since Marshall and the others always viewed the offer curve as embodying an exhaustive classification of all its determinants, supply as well as demand. True, Edgeworth initially derived Marshallian curves for a pure exchange economy involving no production. But he later explicitly acknowledged underlying changes in production in his famous analogy comparing Marshall's offer curves to the hands of a clock driven by the workings of a complex but hidden machinery.

The full revelation of the machinery, however, had to wait for the famous demonstrations of Leontief (1933) and Meade (1952). Both derived offer curves from production transformation frontiers (expressing supply conditions) and consumption indifference curves (expressing demand conditions).

Leontief's derivation was the simplest. He superimposed a trading country's consumption indifference curves directly on its transformation curve. He then assumed alternative international price ratios represented by negatively sloped lines. Tangency of such lines with the transformation and indifference curves gave him the quantities of the two goods produced and consumed at each price ratio. The excess of production over consumption of the one good and of consumption over production of the other at each price ratio constituted export-import bundles lying on the offer curve.

Meade, on the other hand, derived offer curves in two stages (see Figure 21). First, he slid a transformation curve, or production block, along a succession of consumption indifference curves. The origin of the block traced out a set

Figure 21 Meade's Derivation of the Offer Curve



Meade slides the production block along each consumption indifference curve. The block's origin traces out a set of trade indifference curves. He then derives the offer curve as the locus of the points of tangency of the trade indifference curves and the price line as it swivels about the origin.

of trade indifference curves, each curve showing alternative export-import bundles that yield the same level of collective satisfaction. From these trade indifference curves he derived offer curves just as Edgeworth had done. He found the locus of points of tangency of trade indifference curves and alternative price rays emanating from the origin. This locus constituted the offer curve. In deriving the curves from the underlying production conditions, Leontief and Meade vindicated Marshall and exonerated him from Graham's charge.

7. CONCLUSION

Historically, the offer-curve apparatus has been put to two uses. Modern analysts employ it as a pedagogical or expository device to illustrate established truths. By contrast, the concept's originators applied it as an analytical tool to derive new propositions and postulates. They used it to generate key theorems on the gains from trade, on the efficiency of free-trade equilibrium, on the effects of tariffs and technological change on the terms of trade, and on the specification of the optimum rate of a tariff. That they were able to do so using nothing more sophisticated than numerical examples and geometrical diagrams shows what keen minds can accomplish with the simplest of analytical tools.

Their successive accomplishments typify the workings of normal science wherein the drive to perfect an existing paradigm propels advances in theory. In their case, the paradigm consisted of the Mill-Marshall model of terms-of-trade determination. Perfecting it meant (1) making it more precise, (2) generalizing it to cover the widest possible range of cases, and (3) purging it of errors and inconsistencies.

Offer-curve pioneers were more than up to these tasks. Thus Mill's and Torrens' concept of reciprocal demand expunged terms-of-trade indeterminacy from Ricardo's analysis. Mill generalized Torrens' unit-elastic demand schedules to include elasticities ranging from zero to infinity. Marshall generalized Mill's model to cover cases of (1) multiple as well as singular equilibrium and (2) nonconstant as well as constant costs. Edgeworth's invention of indifference maps and the contract curve lent precision to Marshall's concept of the gains from trade. Lerner's innovation, the tax pencil, helped correct Edgeworth's error regarding the symmetry of trade taxes. Finally, Leontief and Meade extended the entire apparatus to include production as well as preference functions.

The result was that offer curves became a fixture of trade theory and a commonplace of textbooks. The survival of the concept testifies to its continued usefulness. Modern students owe the originators of this tool a debt of gratitude. Even today, if one understands the diagram, one comprehends how various disturbances—technology shocks, resource discoveries, taste shifts, erection and removal of trade barriers, and the like—affect the volume and terms of world trade.

REFERENCES

Amano, Akihiro. "Stability Conditions in the Pure Theory of International Trade: A Rehabilitation of the Marshallian Approach," *Quarterly Journal of Economics,* vol. 82 (May 1968), pp. 326–39.

Bhagwati, Jagdish, and Harry G. Johnson. "Notes on Some Controversies in the Theory of International Trade," *Economic Journal,* vol. 70 (March 1960), pp. 74–93.

Chipman, John S. "A Survey of the Theory of International Trade: Part 2, The Neo-classical Theory," *Econometrica,* vol. 33 (October 1965), pp. 685–760.

Edgeworth, Francis Y. "On the Application of Mathematics to Political Economy," *Journal of the Royal Statistical Society,* vol. 52 (December 1889), pp. 538–76, reprinted in Francis Y. Edgeworth, *Papers Relating to Political Economy,* Vol. 2. London: Macmillan (for Royal Economic Society), 1925.

_____. "The Theory of International Values, III," *Economic Journal,* vol. 4 (December 1894a), pp. 606–38.

_____. "The Theory of International Values, II," *Economic Journal,* vol. 4 (September 1894b), pp. 424–43.

_____. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences.* London: Kegan Paul, 1881.

Graham, Frank D. "The Theory of International Values Re-examined," *Quarterly Journal of Economics,* vol. 28 (November 1923), pp. 54–86, reprinted in Howard S. Ellis and Lloyd A. Metzler, eds., *Readings in the Theory of International Trade*. Philadelphia: The Blakiston Co., 1949, pp. 301–30.

_____. "The Theory of International Values," *Quarterly Journal of Economics,* vol. 46 (August 1932), pp. 581–616.

Kemp, Murray C. *The Pure Theory of International Trade.* Englewood Cliffs, N.J.: Prentice-Hall, 1964.

_____. "The Relation Between Changes in International Demand and the Terms of Trade," *Econometrica,* vol. 24 (January 1956), pp. 41–46.

Keynes, John Maynard. "Alfred Marshall, 1842–1924," in Arthur Cecil Pigou, ed., *Memorials of Alfred Marshall*. London: Macmillan, 1925, reprinted, New York: Kelley and Millman, 1956.

Leontief, Wassily W. "The Use of Indifference Curves in the Analysis of Foreign Trade," *Quarterly Journal of Economics,* vol. 47 (May 1933), pp. 493–503, reprinted in Howard S. Ellis and Lloyd A. Metzler, eds., AEA *Readings in the Theory of International Trade*. Philadelphia: The Blakiston Co., 1949, pp. 229–38.

Lerner, Abba P. "The Symmetry Between Import and Export Taxes," *Economica,* vol. 3 (August 1936), pp. 306–13, reprinted in Richard E. Caves and Harry G. Johnson, eds., AEA *Readings in International Economics.* Homewood, Ill.: Richard D. Irwin, 1968.

Marshall, Alfred. *The Pure Theory of Foreign Trade.* Privately printed, 1879, reprinted in amplified form in John K. Whitaker, ed., *The Early Economic Writings of Alfred Marshall 1867–1890,* Vol. 2. New York: Free Press, 1975.

_____. *Money, Credit and Commerce*. London: Macmillan, 1923, reprinted, New York: Augustus M. Kelley, 1963.

Meade, James E. *A Geometry of International Trade*. London: George Allen and Unwin, Ltd., 1952.

Metzler, Lloyd A. "Tariffs, the Terms of Trade, and the Distribution of National Income," *Journal of Political Economy,* vol. 57 (February 1949), reprinted in Richard E. Caves and Harry G. Johnson, eds., AEA *Readings in International Economics.* Homewood, Ill.: Richard D. Irwin, 1968.

Mill, John Stuart. "On the Laws of Interchange Between Nations: and the Distribution of the Gains of Commerce Among Countries of the Commercial World," in *Essays on Some Unsettled Questions of Political Economy.* 1844. Reprinted, New York: Kelley, 1968.

_____. *Principles of Political Economy, with Some of Their Applications to Social Policy*. 1848. Edited with an introduction by W. J. Ashley, London: Longmans, Green and Co., 1909.

O'Brien, Denis P. *The Classical Economists.* London: Oxford University Press, 1975.

Pennington, James. "Letter to Kirkman Finlay Esq." 1840. Reprinted in R. S. Sayers, ed., *Economic Writings of James Pennington*. London: London School of Economics, 1963.

Pigou, Arthur Cecil, ed. *Memorials of Alfred Marshall*. London: Macmillan, 1925, reprinted, New York: Kelley and Millman, 1956.

Ricardo, David. *On the Principles of Political Economy and Taxation*. London: Murray, 1817, reprinted in Piero Sraffa, ed., *The Works and Correspondence of David Ricardo,* Vol. 1. Cambridge: Cambridge University Press (for Royal Economic Society), 1951.

Schumpeter, Joseph A. *History of Economic Analysis*. London: George Allen and Unwin, 1954.

Takayama, Akira. *International Trade*. New York: Holt, Rinehart & Winston, 1972.

Torrens, Robert. Postscript to Letter IX of *The Budget: A Series of Letters on Financial, Commercial, and Colonial Policy*. London: Smith, Elder & Co., 1841–42.

Viner, Jacob. "Marshall's Economics, In Relation to the Man and to His Times," *American Economic Review,* vol. 31 (June 1941), pp. 223–35.

_____. *Studies in the Theory of International Trade*. New York: Harper, 1937.

An Index of Leading Indicators for Inflation

Roy H. Webb and Tazewell S. Rowe

Macroeconomic forecasts attempt to provide useful information on aggregate economic conditions. A good forecast provides a user with specific information that allows him or her to make better decisions. A forecast, whether explicit or implicit, underlies a wide range of choices, such as consumer decisions on whether to spend or save, business decisions on investments in plant and equipment, and central bank actions affecting reserve supply.

No single approach to macroeconomic forecasting has dominated the others. Different users may require different types of information, leading to different forecasting methods. For example, researchers have proposed substantially different strategies for predicting the timing of an event, such as a recession, versus predicting the magnitude of a related statistic, such as the rate of real GDP growth. Probably most important, even the best forecasts lack precision. Macroeconomic forecasts usually have high average errors, but even the average size of errors can change substantially over time. It can therefore be difficult to distinguish a good forecasting method from a mediocre one.

One approach to forecasting is to construct a theoretical model, use it to identify the shocks affecting economic activity, and then use it for forecasting. But forecasters of inflation must confront the difficulty in modeling the interaction of real and nominal variables. No consensus has emerged among economists on the best way to model that interaction. The large macroeconomic models designed specifically for forecasting typically incorporate such ingredients as a Phillips Curve relationship between wage inflation and unemployment, and a backward-looking method for modeling how individuals form expectations. Many macroeconomists, however, do not believe that such relationships

The authors gratefully acknowledge helpful comments from Mary Finn, Robert Hetzel, Thomas Humphrey, Peter Ireland, Stephen McNees, and Michael Niemira. The views and opinions expressed in this article are solely those of the authors and should not be attributed to the Federal Reserve Bank of Richmond or the Federal Reserve System.

accurately reflect actual behavior. In the 1970s those models had large errors when predicting inflation, which is consistent with the critics' concerns.[1]

Another approach to forecasting involves using an explicit statistical model that requires little economic theory. A prime example of this "atheoretical" approach is the vector autoregressive (VAR) model. While that strategy has produced relatively accurate forecasts of real variables, it has also produced inflation forecasts that not only failed to be more accurate than the large models, but also were worse than a naive no-change forecasting method.[2]

This article takes a different approach to forecasting inflation. Like VAR models, it uses little explicit theory. Unlike the standard theoretical and atheoretical models, however, its primary contribution is not to predict the magnitude of future inflation, but rather to help recognize and predict major swings in inflation, based on an index of leading indicators for inflation (ILII). The article first presents background information on leading indicators, followed by a detailed account of the ILII's construction. The index's performance is then evaluated. Finally, that performance is related to the business cycle and the strategy of monetary policy.

1. WORK BY OTHER AUTHORS ON LEADING INDICATORS

The study of leading indicators of cyclical change was an important part of the pathbreaking studies of business cycles conducted by scholars associated with the National Bureau of Economic Research (NBER). This classic NBER approach is well represented by Burns and Mitchell (1946) and Moore (1961). That work has inspired more recent work such as that by Stock and Watson (1989).

The performance of traditional leading indicators has been mixed. The same, of course, can be said about every macroeconomic forecasting method. One problem is that the best-known index, the Commerce Department's Composite Index of Leading Indicators (CLI), does not have a precise meaning defined by economic or statistical theory. Any evaluation of that index must therefore begin with two key considerations: the objective of the CLI and a method for defining signals. The objective of predicting cyclical turning points is usually taken for granted, and perhaps the most common definition is that two or three successive declines signal an imminent recession.

Diebold and Rudebusch (1991) evaluated the three-decline rule and also a newer technique proposed by Neftci (1982) for using the index of leading

[1] Lucas and Sargent (1979) give a forceful statement of that view.

[2] McNees (1986) documents the poor performance of Robert Litterman's VAR inflation forecasts versus other forecasters. Webb (1995) documents the poor performance of many VAR forecasts of inflation in comparison to the naive no-change forecast.

indicators to predict cyclical changes. When using originally released data and the Neftci approach, they found at best a slight improvement over a simple rule of always predicting a constant probability of a turning point. They found no improvement for the three-decline rule when compared to the simple prediction. Their negative judgment was seconded by Koenig and Emery (1991). Niemira and Fredman (1991) found a more positive value for the index, possibly because they used revised values of the CLI instead of originally released data. Zarnowitz (1992a) presented another positive view of the leading index. Instead of using the usual three-decline rule, he used a multi-step rule that yielded a more complicated signal[3] of an approaching cyclical turning point. Despite their advocacy, this rule has not been widely used, although it continued to work well after they proposed it.

Responding to the lack of specific meaning of the Commerce Department's leading index, Stock and Watson (1989) proposed an index of leading indicators that has a well-defined meaning in a particular statistical model. First, they defined a coincident index as an estimate of the unobserved state of the economy, that is, as a measure that summarizes the economy's position in relation to the business cycle. They then constructed a leading index by predicting the value of the coincident index six months ahead. They were then able to calculate a recession index as the probability that the coincident index would decline over the next six months. In its first post-sample test, their index failed to predict or recognize the 1990 recession (Stock and Watson 1993).

A few authors have constructed leading indicators for inflation. Roth (1991) gives an initial assessment of their performance. Most prominent is a leading series constructed by Geoffrey Moore and his associates at the Center for International Business Cycle Research (CIBCR).[4] That series now includes seven constituent series, including a commodity price measure, the growth rate of total debt, and the ratio of employment to population. Roth found that the Moore index anticipated turning points in CPI inflation "quite well."

All of the leading indicator indexes mentioned above share an important characteristic: they are constructed as a weighted average of a fixed set of indicators. The weights and components, however, are subject to change at irregular intervals according to criteria that have not been specified in advance. An index can therefore be constructed to do well in a particular period under study, but when the economy changes, the index will need revision. Users are thus faced with the necessity of deciding whether a signal may have been produced by

[3] The complexity of the signal results from it having three parts at peaks and troughs. The first indication of a peak, labeled P1, is a long-leading signal that has produced several false positives. A first confirmation, labeled P2, has had only one false positive, in 1951, and has correctly anticipated or confirmed the eight peaks since then. The median P2 signal arrives two months following the peak. In addition, there is a second confirming signal labeled P3 that has had no false positives.

[4] See, for example, Klein (1986).

an out-of-date index that will be substantially revised in the near future. This is a particular problem for a leading indicator of inflation, since changes in monetary regimes may well change previous historical relationships.[5]

Sims (1989) proposed a solution for the problem of adapting an index to a changing economic environment. In his comments on the work of Stock and Watson (1989), he advocated using a model with time-varying coefficients, rather than the fixed coefficients they actually employed. In addition, he proposed performing their variable selection process annually. (Stock and Watson examined 280 series in order to select the 7 in their leading index.) Sims argued that because of abnormal events in the 1970s, Stock and Watson's index overemphasized interest rates, which affected estimates for the whole sample period. That large emphasis on interest rates did lead to the failure of their index to capture the 1990 recession, which in turn led them to propose an alternative leading index that omits the financial variables.

The index of leading indicators for inflation that we propose incorporates one of Sims's suggestions. Instead of relying on a fixed set of series that will probably be changed at an unspecified future date, we propose a strategy for each month selecting seven indicators from a much larger set of candidates. The following section explains that strategy in detail.[6]

2. CREATING AN INDEX OF LEADING INDICATORS OF INFLATION

To create an index of leading indicators of inflation (ILII), we initially specified a set of time series that might be included in the ILII. Potential indicators had to meet two criteria. First, each series had to be related to inflation in some plausible manner since we did not want to include any series that had a completely spurious correlation with inflation. Second, in order to construct an index that would be available promptly, we studied only potential indicators that would be available prior to the release of the monthly CPI figures. A notable example of a series that failed to meet the latter requirement is the capacity utilization rate.

Table A1 in the appendix lists the potential indicators used below. Series can be grouped into several broad categories, including money supply data, interest rates (studied as a leading indicator of inflation, for example, by Dasgupta and Lahiri [1991]), commodity prices (for example, see Boughton and

[5] For example, Webb (1995) found that two changes in the monetary regime account for the poor forecasting record for inflation rates of VAR models using postwar U.S. data.

[6] Another strategy for handling a changing relationship between indicators and inflation is sketched by Niemira and Klein (1994, pp. 383–88). Their prediction of inflation from seven leading indicator series is based on a neural network method, which was designed to be able to adapt over time to certain economic changes.

Branson [1991]), and labor market measures. Note that in some cases, one series is simply a transformation of another series, such as an interest rate and its difference over six months. Those cases resulted if we were unsure as to whether to remove a trend or how best to transform a nonstationary variable to a stationary one. There are 30 potential indicators, including different transformations of the same variable.

The second step was to create a strategy to select seven series for the index.[7] Rather than following the traditional approach and using a single set of inflation indicators for the entire sample, we developed a method for creating an index for which components could change frequently. The strategy was designed to use only information that would have been available to a "real time" user; that is, the index for January 1966 would be based only on data released by the middle of that month.

For each month from January 1958 to December 1994, the strategy was to select the seven candidate series that had the largest correlation coefficients with inflation. We measured inflation by the percentage change in the monthly level of the core CPI—that is, the CPI excluding food and energy prices—from its value 12 months earlier. We used the core CPI in order to focus on sustained inflation trends; the core CPI removes transitory changes in the CPI caused by movements in volatile food and energy prices.[8] In order to reflect current economic conditions, each correlation coefficient was calculated over the most recent 48-month period rather than using a longer sample. And to examine correlations with *future* inflation, we lagged each candidate series 12 months. For example, in January 1995 the latest inflation reading would be calculated from December 1993 to December 1994, and the latest observation of a candidate series *before* that inflation occurred would be December 1993. A correlation coefficient dated January 1995 would thus be computed between (1) inflation rates calculated using price levels from December 1989 to December 1994 and (2) a candidate series from December 1989 to December 1993.

At each date the seven selected series were then combined into a leading indicator index. First, each series was adjusted for differences in levels and volatility by subtracting the mean (computed over the previous 48 months) and dividing by its standard deviation (also calculated over the previous 48 months). To avoid undue influences from highly unusual events, such as the government's freeing the price of gold, each observation had a maximum absolute value of three (larger values were accordingly reduced). Unlike the procedure for

[7] Why seven? That seems to be a popular number that works reasonably well. Stock and Watson (1989) include seven series in their index of leading indicators for predicting the real economy. The CICBR index of leading indicators for predicting inflation has seven components, as does the index for predicting inflation described in Niemira and Klein (1994).

[8] Official data on the CPI excluding food and energy prices only extend back to 1959. For earlier data, we used the nonfood CPI.

producing the CLI for most of its history, the strategy employed here was to use equal weights for the series. Our index was simply the average of the seven transformed series.[9]

The graph of the resulting series, along with the 12-month change in the core CPI, is presented in Figure 1. The inflation series is dated so that an entry at date t is the percentage change in the core CPI from t to $t+12$. Table A2 in the appendix shows how often the various series enter the index, and Table A3 contains the composition of the index at turning points of the inflation cycle.

3. PERFORMANCE OF THE INDEX OF LEADING INDICATORS OF INFLATION

Ex Post Qualitative Evaluation

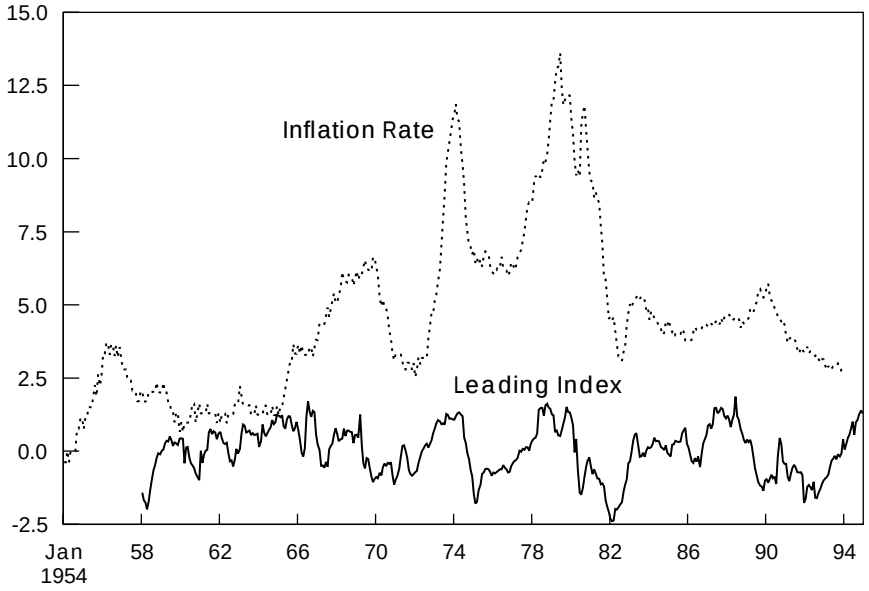
We experimented with the possibility of following Stock and Watson (1989) and constructing an index with an explicit statistical meaning, namely, the forecast of the core CPI from a bivariate VAR of the core CPI and the ILII. When such models are estimated for U.S. data over the last 30 years, however, the estimated coefficient on the first lagged value of the inflation rate in the price equation is always relatively large and tends to overshadow other terms. If one then uses that estimated equation for forecasting, it therefore tends to place such a large weight on recent inflation that the resulting forecasts are lagging indicators around turning points. Since the goal of the ILII is to promptly recognize or predict sustained and substantial changes in the inflation rate, the additional lag introduced by stating the index as a VAR forecast is unacceptable.

The ILII therefore needs a well-defined signal before its performance can be assessed. Figure 1 indicates that the ILII tends to promptly recognize substantial changes in inflation, although inevitably there are a lot of small fluctuations in the graph. In order to filter out small changes, we reduced to zero those values that had absolute values of less than one, thereby producing the series shown in Figure 2. A signal of rising inflation is thus a value greater than or equal to one, and a signal of declining inflation is a value less than or equal to minus one. An observation of at least one in absolute value will be referred to as a main signal.

The interpretation of observations with absolute values less than one is less obvious. We adopt the rule that a main signal is valid for up to 11 months if followed by absolute values less than one. Twelve months of such small readings, however, can be an early signal of a turning point. We define it as a signal if the ILII in the twelfth month is positive for a signal of rising inflation

[9] Although the weights on individual series are equal, it is possible for several closely related series to be included. The effective weight on commodity prices, for example, could be quite high. In Table A3, note that the index in November 1983 contained six commodity price series.

Figure 1 Core CPI and Leading Indicator Index

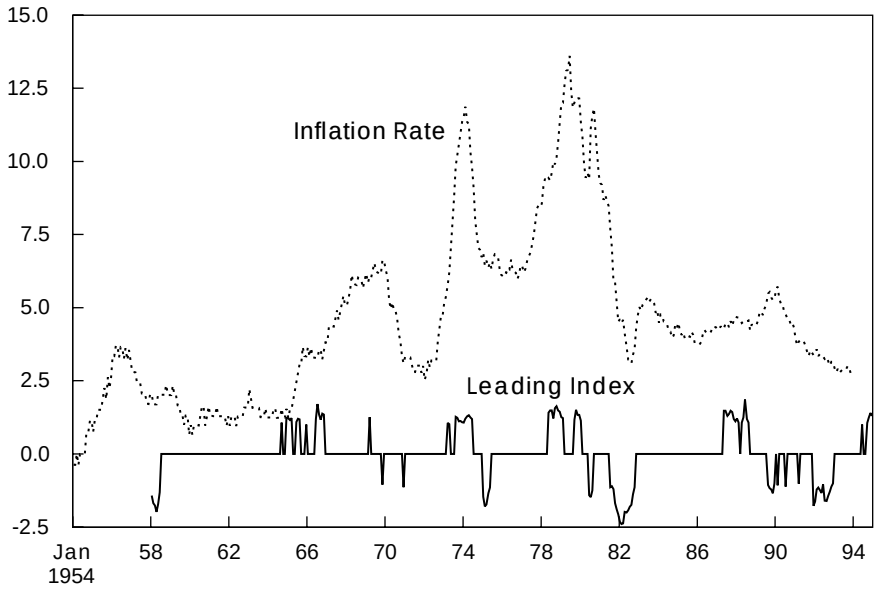


Note: The inflation rate is from t to $t + 12$.

(that is, in the neighborhood of a trough) or if the ILII is negative for a signal of falling inflation. An early signal remains valid until a main signal is received. As can be seen in Figure 2 or Table 1, in several instances there is no early signal of a turning point.

There is also no official dating of periods of substantial and sustained changes in the rate of inflation. Inspecting Figure 1 yields the following dates: peaks in March 1956, November 1969, February 1974, June 1979, and February 1990, and troughs in January 1962, January 1972, October 1976, August 1982, and possibly in December 1993. There is, of course, room for disagreement about particular dates. The hardest call was whether to define another peak and trough in the mid-1980s. Other authors, looking at slightly different data, have taken opposing sides. Roth (1989) argued that "[t]he eleven-month upturn in inflation beginning in March 1983 is most likely a statistical artifact" (p. 283). Moore (1991), however, found a peak in early 1984 and a trough in 1986. By looking at the 12-month forward rate of change of the core CPI, one can see that inflation reached a local minimum in August 1982 at 3.1 percent. It then rose to 5.3 percent in July 1983, which is a substantial change. However, it then fell to 4.2 percent by July 1984 and to 3.8 percent in November 1985. Thus the bulk of the increase was not sustained but rather was fairly quickly

Figure 2 Core CPI and Filtered Leading Indicator Index



Note: The inflation rate is from t to $t + 12$.

reversed. Accordingly, the 11-month upswing is not counted as a substantial and sustained increase in the rate of inflation.

Table 1 contains the resulting inflation signals from the ILII and compares them with peaks and troughs. The index is helpful in recognizing major changes in inflation rates, but often does not anticipate turning points; the median time for receiving the first signal is five months after the turning point. The 1990 peak is the only one that is clearly predicted, although the turn is recognized immediately at the November 1969 peak. If December 1993 or a nearby date turns out to be a trough, the ILII will have given a prompt signal; it is possible, however, that it will turn out to be a false signal. The worst performance is the 1982 trough, which is only recognized after 15 months. Other turning points are recognized within a year. Importantly, *no* false signals are generated[10] and no turning points are missed. In addition, although the ILII appears to recognize, not anticipate, the dates of major changes in the rate of inflation, Table 2 presents evidence that it does anticipate the bulk of the change in the inflation rate. The change in the inflation rate before a signal is no greater than 0.5

[10] A false signal would be one that is later reversed before a predicted peak or trough occurs.

Table 1 Turning-Point Signals from the Index of Leading Indicators of Inflation

Date of Turning Point	Early Signal	Main Signal	Lead (+) or Lag (−) from First Signal	Recognition Rule
Peak, March 1956 (April 1957)		Down, January 1958	na	November 1958
Trough, January 1962 (January 1963)	Up, June 1959	Up, September 1964	+31	January 1964
Peak, November 1969 (April 1970)		Down, November 1969	0	August 1971
Trough, January 1972 (January 1973)	Up, June 1972	Up, March 1973	−5	November 1973
Peak, February 1974 (October 1974)		Down, January 1975	−11	June 1975
Trough, October 1976 (December 1976)	Up, May 1977	Up, May 1978	−7	October 1978
Peak, June 1979 (April 1980)		Down, June 1980	−12	January 1981
Trough, August 1982 (March 1983)	Up, November 1983	Up, May 1987	−15	March 1984
Peak, February 1990 (August 1990)		Down, August 1989	6	February 1992
Trough? December 1993	Up, December 1993	Up, June 1994	0?	

Notes: The leading indicator series begins in January 1958 and ends in December 1994. Peaks and troughs correspond to the 12-month percentage change in the CPI excluding food and energy, as shown in Figures 1 and 2. (The dates in parentheses are peaks and troughs in the "6-month smoothed inflation rate" discussed in the text and footnote 11.) The inflation series begins in January 1954 and ends in December 1993 and represents the inflation rate from each date to 12 months ahead. The last trough, identified by the question mark, is tentative. The rightmost column presents the first date that a two percentage point change in the inflation rate from a previous turning point could have been observed.

Table 2 Inflation Rates Before and After Lagging Signals of Turning Points

Turning-Point Date	Inflation Rate, Previous Turning Point to First Signal	Inflation Rate at Next Turning Point	Anticipated Change	Unanticipated Change
January 1972	2.8	11.9	9.1	0.3
February 1974	11.8	6.0	5.8	0.1
October 1976	6.2	13.6	7.4	0.2
June 1979	13.6	3.1	10.5	0.0
August 1982	3.6	5.7	2.1	0.5

Notes: The table presents turning-point dates for which the first signal from the leading index occurred after a turning point. The first column lists turning-point dates at which a lagging signal of a turning point was given. The second column gives the annualized rate of change in the core CPI from the turning-point date to the date of the first signal given by the ILII. The third column gives the rate of inflation at the next turning point. The fourth column represents the change in inflation after a signal is received, calculated as the difference between column three and column two. The last column represents the change in inflation before a signal is received and is calculated as the difference between the inflation rate at the previous turning point and the value listed in column two.

percent, whereas the change after the signal is received ranges from 2.1 to 10.5 percent.

Although the format of Table 1 and Figure 2 may at first glance resemble those used by others who have evaluated leading indicators, such as Klein (1986), Moore (1991), and Roth (1991), there is a key difference. The other authors compare the value of a leading indicator with inflation calculated as the contemporaneous value's change from *lagged* values. Thus they are comparing an indicator with lagging inflation, a comparison that may not be relevant for actual use of a leading index. Our analysis compares the leading indicator with future inflation. The difference can be seen in Table 1, in which inflation is also calculated in the manner used by the other authors, and the resulting dates of turning points are displayed in parentheses. From those dates it would appear that the index has more predictive power than originally indicated, even though the ILII is unchanged. What has changed is the method of calculating inflation, which shifts the dates of turning points forward by a little over eight months, on average.[11]

Recognizing major swings in inflation is not always a simple exercise, as Cullison (1988) demonstrates. An example is 1972: inflation's low point was

[11] The alternative method of calculating inflation is referred to as the "6-month smoothed annual rate." It is calculated as the ratio of the current month's price index to the average index of the preceding 12 months and is converted to an annual rate by raising the ratio to the 12/6.5 power.

in January, and the ILII gives an early signal in June, lagging the change by five months. The following commentary on a well-regarded model's forecasts is recorded by Cullison:

> April, 1972: "The rate of price increase is expected to slow. . . . The anticipated slowing . . . reflects the large projected rise in real product and associated productivity gains."
>
> June, 1972: "The rise in the [GNP implicit price] deflator is expected to . . . moderate. . . . The expected moderation reflects a moderation in the rise in unit labor costs."
>
> May, 1973: "The projected slowdown in the rise in the private GNP fixed weight price index reflects primarily the anticipation that food price increases will slow sharply."[12]

As this example illustrates, having leading indicators that began to signal rising inflation in June 1972 could have been valuable to forecasters. Another comparison can be seen by using the rightmost column of Table 1, in which each entry denotes the first date at which one could observe a 200 basis point change in the inflation rate after a turning point. The ILII signals turning points much sooner than that simple rule.

While the index appears to perform well, that judgment is based on the same data that were used to construct the index; its actual performance will be revealed by new data. The apparent performance of the index undoubtedly could have been improved by a systematic search over parameters such as the number of series, the weights on each series, the magnitude of the main signal, or the number of months required for either a main signal or an early signal. The future performance of an index so constructed undoubtedly would deteriorate, however. We therefore picked obvious values that seemed to work well, but a caveat remains. Any choice that we made would have been rejected if it conflicted with the data. The proof of how well the index works must await new data that were not used to construct it. An additional caveat is that we used the latest revisions of data, not data as originally released. That fact should be less important for this index than for the Commerce Department's CLI, however, since most of the individual series employed in this paper are not revised by substantial amounts.

Simulated Forecasts

Another check on whether the ILII contains useful information is to test whether it adds predictive power to lagged values of inflation. To test for additional

[12] Cullison's quotations are from the Greenbook, prepared by the staff of the Board of Governors of the Federal Reserve System prior to meetings of the Federal Open Market Committee. Karamouzis and Lombra (1989) have conducted a thorough examination of the quality of these forecasts and have concluded that the forecasts were "state of the art" in comparison with other macroeconomic forecasts.

predictive power, we constructed a bivariate VAR for monthly percentage changes in the core CPI and the level of the ILII. We first set lag lengths in the VAR by minimizing the Akaike Information Criterion over each lag length in each equation, resulting in the following equations:

$$P_t = \beta_{10} + \sum_{i=1}^{9} \beta_{11,i} P_{t-i} + \beta_{12} \mathrm{ILII}_t + e_{1,t}, \tag{1}$$

$$\mathrm{ILII}_t = \beta_{20} + \sum_{i=1}^{2} \beta_{21,i} P_{t-i} + \sum_{i=1}^{3} \beta_{22,i} \mathrm{ILII}_{t-i} + e_{2,t}, \tag{2}$$

where P is the percentage change in the core CPI from the previous month, ILII is the index of leading indicators of inflation, β is the model's coefficients, and e is the error term. For comparison we also estimated a univariate autoregression for the core CPI, using nine lagged values. The equations were estimated starting in 1958:1 and ending in 1969:12, and out-of-sample forecasts were made up to 12 months ahead. One month was then added to the period, the equations were reestimated, and new forecasts were made. We repeated this process to create series of 12-month inflation forecasts for the period 1970:12 to 1994:12.

Forecast errors were calculated as the difference between actual inflation and forecasted values, and summary statistics were calculated. The root mean squared error for the univariate forecasts was 2.19; it fell to 1.96 for the bivariate forecasts. Comparing the two series of squared errors, we found the difference to be significant at the 1 percent level according to a test proposed by Diebold and Mariano (1991). We conclude that the index does contain information with significant predictive value beyond that contained in the inflation series itself. The size of the forecast error, however, is a reminder of substantial remaining uncertainty in forecasts from this method. For perspective, consider that in the post-1983 period the average 12-month change in the core CPI was 4.2 percent. Taking the root mean squared error as an approximation of the anticipated standard error of current forecasts, even a 70 percent confidence interval, ± 2 percent, includes a wide range of outcomes.

We also estimated equation (1) over the entire sample period. The average error was again significantly lower when the ILII was included, indicating that it significantly improved one-month forecasts of inflation.

4. WHY THE INDEX APPEARS TO WORK, AND WHAT COULD CHANGE

On the basis of experience in the United States and other industrial countries before 1913, Wesley Mitchell (1941) presented an account that describes a

stylized business cycle. The behavior of prices played a key role in that account, as the following passages illustrate.

> A revival of activity, then, starts with this legacy from depression: a level of prices low in comparison with the prices of prosperity, drastic reductions in the cost of doing business [p. 150]. While the price level is often sagging slowly when a revival begins, the cumulative expansion in the physical volume of trade presently stops the fall and starts a rise [p. 151]. Like the increase in the physical volume of business, the rise in prices spreads rapidly; for every advance of quotations puts pressure upon someone to recoup himself by making a compensatory advance in the prices of what he has to sell. . . . Retail prices lag behind wholesale . . . and the prices of finished products [lag] behind the prices of their raw materials [p. 152]. [O]ptimism and rising prices both support each other and stimulate the growth of trade [p. 153]. Among the threatening stresses that gradually accumulate within the system of business during seasons of high prosperity is the slow but sure increase in the costs of doing business [p. 29]. The price of labor rises. . . . The prices of raw materials continue to rise faster on the average than the selling prices of products [p. 154]. [T]he advance of selling prices cannot be continued indefinitely . . . [because] the advance in the price level would ultimately be checked by the inadequacy of the quantity of money [p. 54]. [Once a downturn begins] with the contraction in trade goes a fall in prices [p. 160]. [T]he trend of fluctuations [in prices] continues downward for a considerable period. . . . [T]he lowest level of commodity prices is reached, not during the crisis, but toward the close of the subsequent depression, or even early in the final revival of business activity. The chief cause of this fall is the shrinkage in the demand for consumers' goods, raw materials, producers' supplies, and construction work [p. 134]. [E]very reduction in price facilitates, if it does not force, reductions in other prices [p. 160]. Once these various forces have set trade to expanding again, the increase proves cumulative, though for a time the pace of growth is kept slow by the continued sagging of prices [p. 162].

Zarnowitz (1992b) reviewed the literature and found that much of Mitchell's account has been consistent with cyclical data generated after he wrote it. There has been an important change, however. Under the gold standard there was little, if any, trend to the price level, and prices could fall in one phase of the business cycle and rise in another. In contrast, American monetary policy in the last 50 years has put in place an upward trend in prices. Thus, where Mitchell observed prices declining when cyclical contractions ended and expansions began, one now observes inflation being relatively low. Similarly, toward the end of expansions Mitchell saw price increases, but one now would see relatively high inflation.

Table 3 presents some evidence on this last point by looking at the behavior of inflation and other statistics over the business cycle. Expansions are divided into four segments of equal length, and contractions are divided into two equal segments. The inflation rate is calculated for each cycle, measured on a trough-to-trough basis. For each segment of the cycle, the average inflation rate for

the cycle is subtracted from the inflation rate for that segment; the result is a relative inflation rate for each cyclical segment. The relative rates can then be averaged over the last seven business cycles in order to depict the average cyclical behavior of inflation. The picture is clear: inflation is low early in a cyclical expansion, is relatively high in the last quarter of expansion, and peaks in the first half of recessions. Inflation is therefore procyclical in the sense that its rate increases during expansions and declines during contractions. It is also a lagging indicator in the sense that its highest rate usually occurs after the cyclical peak and its lowest rate usually occurs after the cyclical trough.[13] The leading indicator series anticipates that behavior by peaking in the third quarter of a typical expansion and hitting its low point in the last half of recessions.

It therefore appears that the leading indicator index is capturing a regular feature of the business cycle. High-frequency changes in inflation, which are clearly not sustained, are ignored by design. Changes in inflation rates between business cycles are also excluded from the picture. What is left are cyclical movements that have been reliable and predictable. An individual indicator can be a useful predictor if it has a definite place in the sequence of events of a typical business cycle.

Consequently, this index has a reason for working and does not simply reflect a spurious correlation. It is designed to continue to work under certain changing conditions. If any particular indicator were to change its cyclical behavior, its correlation with inflation would diminish and it would not be included in the index. Similarly, adding new indicators would be straightforward. The one event that could drastically change the role of the index would be a substantial change in the strategy of monetary policy. After all, the shift from the gold standard to a fiat money system that involved a particular central bank strategy changed the cyclical behavior of prices to the cyclical behavior of inflation. A different monetary strategy might cause another dramatic change that could change the role of this index.

For example, imagine a monetary strategy that eliminated the trend in prices by keeping inflation rates small in magnitude and centered on zero. Without sustained and substantial changes in inflation, would the index have any purpose? Certainly the strategy of choosing indicators by past correlations with inflation would need replacing. For a closely related example, imagine a monetary strategy that eliminated large fluctuations in inflation by keeping it relatively low but positive. In that case, the index would be much more

[13] Some authors, such as Cooley and Ohanian (1991), have asserted that prices are countercyclical. By their definition, prices are countercyclical if there is a negative correlation between the level of prices and the level of output when the same statistical transformation is applied to both series. For example, in Table 3 there is a negative comovement between real GDP growth and inflation: during the segment of the business cycle where one series peaks, the other series reaches its lowest value. Their finding does not contradict the statement that inflation is procyclical, using the usual NBER definition for procyclical.

Table 3 Cyclical Behavior of Inflation and Other Series

Statistic	Expansion First Quarter	Expansion Second Quarter	Expansion Third Quarter	Expansion Fourth Quarter	Recession First Half	Recession Second Half
Core CPI	−0.40	−0.45	−0.53	1.66	2.09	0.90
CPI	−1.28	−0.90	0.15	2.14	2.21	0.91
PPI, Finished Goods	−1.57	−1.44	0.95	1.75	3.22	0.34
ILII	−0.21	0.02	0.51	0.28	−0.11	−0.85
Real GDP	3.55	2.42	0.43	−0.61	−6.10	−3.52

Notes: Entries for all statistics except the ILII are relative rates of change. Each is calculated by subtracting the average rate of change over each business cycle from the rate of change during each segment of the business cycle, and then averaging over all post-Korean War cycles (the core CPI begins with the business cycle trough of 1958). The ILII is the relative value, calculated by subtracting the average value over each business cycle from the average value in each segment and then averaging over the post-1960 business cycles.

valuable if it could give a third signal, stable inflation, in addition to signals of inflationary increases and decreases.

This latter possibility can be illustrated with the ILII. The following rule for a stable price signal is added in order to identify periods in which the index is low and stable. If the level of the index, the 12-month change in the index, and the 12-month average value of the index are all less than 0.3, then a stable inflation period is signaled. This signal overrides the early signal of a turning point and, in turn, is overridden by a main signal.

That rule gave two signals that identify the two major periods of stable inflation in the sample period. The first signal was in May 1960; the inflation rate was within a two percentage point range from April 1957 until June 1965, with the low point in February 1960 and the first main signal of an upswing occurring in September 1964. The second was in March, 1984; the inflation rate was within a two percentage point range from September 1982 until July 1989, with the first main signal of an upswing occurring in May 1987. Based on those two observations, it appears that the index can be adapted to recognizing periods of stable inflation as well as signaling major changes in the inflation rate.

5. CONCLUSION

We have proposed a strategy for constructing an index of leading indicators for inflation. The goal is to recognize or predict sustained and substantial changes in the rate of inflation. A notable feature of our strategy is that it allows the composition of the index to change over time in response to changing economic conditions.

Our evaluation of the index emphasized its link to future inflation rates. In contrast, other evaluations of inflation indicators have often looked at less relevant lagging inflation rates. Our index appears to have value recognizing, and sometimes predicting, major swings in inflation. Important to its possible use is the fact that no false signals were generated and no turning points were missed. In each case, the index allowed the bulk of the change in inflation rates to be anticipated. And although the index was not designed to forecast the magnitude of inflation, it did help lower the forecast error for inflation rates in a simple model.

The performance of the index was related to typical movements of inflation over the business cycle. Whereas inflation is a procyclical but lagging indicator, the leading index typically peaks in the middle of expansions and has its lowest value in the first half of recessions. While this cyclical behavior should be robust in many environments, a major change in the strategy of monetary policy could substantially change the value of such an index. We illustrated the possibility of using the index to recognize periods of stable inflation.

It should be emphasized that the same data were used to construct the index and evaluate its performance. Since out-of-sample data will give the best test of the index's usefulness, the performance of the index outside the sample period will be studied in future research.

APPENDIX: SERIES USED IN THE INDEX OF LEADING INDICATORS FOR INFLATION

The appendix lists the series used to create the index of leading indicators for inflation. Table A1 contains series originally provided by the following sources: the Bureau of Labor Statistics (BLS), the Board of Governors of the Federal Reserve System (FRB), the Federal Reserve Bank of St. Louis (FSL), the National Association of Purchasing Management (NAPM), *The Wall Street Journal* (WSJ), the *Journal of Commerce* (JOC), the Commodity Research Bureau (CRB), and the *Treasury Bulletin* (TB). Data used in this article were obtained from secondary sources. The starting date, either January 1954 or the first month for which the transformed series is available, is affected by data availability and the particular method used for detrending data. Detrending methods are denoted by superscripts.

Table A2 provides further information on the series, as well as how often the individual series are included in the seven-series index. All three labor utilization measures are included more frequently than any other series. The NAPM price index is the only other series included more than half the time. Table A3 gives the composition of the leading index at times of inflation turning-point signals. Again, the three labor utilization measures and the NAPM price index are included more frequently than other series.

Table A1 Candidates for the Index of Leading Indicators for Inflation

Mnemonic	Definition	Source	Series Start
Labor Utilization			
U[a]	Civilian unemployment rate	BLS	1954:1
EP[a]	Employment to population ratio	BLS	1954:1
HR[a]	Index of aggregate weekly hours	BLS	1954:1
Money and Interest Rates			
M1[b]	M1 money supply	FRB	1954:1
M2[b]	M2 money supply	FRB	1954:1
MB[b]	Monetary base	FSL	1954:1
RFF[c]	Federal funds rate	FRB	1954:8
RT10[c]	Ten-year Treasury bond rate	FRB	1954:1
RSP	RT10−RFF		1954:8
Commodity Prices			
PN	Commodity price diffusion index	NAPM	1954:1
PAU[d]	Price of gold, London fix	WSJ	1967:12
PO[d]	Producer price index, crude oil	BLS	1960:1
PJC[d]	Price index of industrial commodities	JOC	1954:1
PCS[e]	Spot price index	CRB	1981:12
PCF[d]	Futures price index	CRB	1970:7
PPIF[b]	Producer price index—finished goods	BLS	1974:7
PPII[b]	Producer price index—intermediate goods	BLS	1974:7
PPIC[b]	Producer price index—crude goods	BLS	1974:7
Other Indicators			
SUP	Supplier deliveries diffusion index	NAPM	1960:1
LD	Lead time for orders and materials	NAPM	1977:1
XD	Trade-weighted value of the dollar	FRB	1960:1
W[b]	Average hourly earnings	BLS	1968:7
FD[b]	Federal government debt	TB	1958:7

[a] Each value is the ratio of the current month to the five-year average ending in the previous month.

[b] Each value is the six-month difference of logarithms of the variable.

[c] The series is used both in level form and in the difference over six months.

[d] The series is used in two forms; one is detrended by the method described in footnote a, and the other is detrended by the method described in footnote b.

[e] Each value is the ratio of the current month to the one-year average ending in the previous month.

Table A2 Candidate Series Selected for Leading Indicator Index

Candidate Series	Number of Months Available	Number of Months Included	Percent
U	444	275	62
EP	444	307	69
HR	444	247	56
M1	444	86	19
M2	444	118	27
MB	444	93	21
RFF, level	437	191	44
RFF, difference	431	64	15
RT10, level	444	64	14
RT10, difference	444	103	23
RSP	437	60	14
PN	444	226	51
PAU*	318	87	27
PAU, difference	372	55	15
PO*	372	55	15
PO, difference	372	91	24
PJC*	444	193	43
PJC, difference	444	132	30
PCS*	104	45	43
PCS, difference	109	34	31
PCF*	193	58	30
PCF, difference	246	51	21
PPIF	198	11	6
PPII	198	82	42
PPIC	198	39	20
SUP	372	120	32
LD	168	47	28
XD	372	67	18
W	337	50	15
FD	390	57	15

* Ratio of the value of the variable divided by a trailing five-year average (one-year average for PCS).

Notes: The first column lists each candidate series (see Table A1 for more complete descriptions). The second column lists the maximum number of months each series could enter the ILII. The third column lists the number of months the mechanical method outlined in the text selected each series to enter the index. The fourth column shows the ratio of column 3 to column 2.

Table A3 Composition of Index of Leading Indicators of Inflation at Dates of Turning-Point Signals

Series	Jan 58	June 59	Nov 69	June 72	Jan 75	May 77	June 80	Nov 83	Aug 89	Dec 93	Total
U	+	+		+	+	+	+	+	+	+	9
EP	+	+		+	+	+	+			+	7
HR	+	+		+	+	+	+			+	7
M1			+								1
M2			+							+	2
MB											0
RFF		+			+						2
ΔRFF				+							1
RT10	+										1
ΔRT10	+	+		+							3
RSP			+								1
PN	+	+			+	+	+	+	+		7
PAU			+					+	+	+	4
ΔPAU			+								1
PO											0
ΔPO						+					1
PJC	+	+				+		+		+	5
ΔPJC				+			+				2
PCS									+		1
ΔPCS											0
PCF								+		+	2
ΔPCF					+			+			2
PPIF											0
PPII								+	+		2
PPIC							+		+		2
SUP				+	+	+	+		+		5
LD											0
XD											0
W			+								1
FD			+								1

REFERENCES

Boughton, James M., and William H. Branson. "Commodity Prices as a Leading Indicator of Inflation," in Kajal Lahiri and Geoffrey H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records.* Cambridge: Cambridge University Press, 1991, pp. 305–38.

Burns, Arthur F., and Wesley C. Mitchell. *Measuring Business Cycles.* New York: National Bureau of Economic Research, 1946.

Cooley, Thomas F., and Lee E. Ohanian. "The Cyclical Behavior of Prices," *Journal of Monetary Economics,* vol. 28 (August 1991), pp. 25–60.

Cullison, William E. Appendix to Working Paper 88-6, "On Recognizing Inflation." Richmond: Federal Reserve Bank of Richmond, 1988.

Dasgupta, Susmita, and Kajal Lahiri. "A Leading Indicator of Inflation Based on Interest Rates," in Kajal Lahiri and Geoffrey H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records.* Cambridge: Cambridge University Press, 1991, pp. 339–53.

Diebold, Francis X., and Roberto S. Mariano. "Comparing Predictive Accuracy: An Asymptotic Test," Discussion Paper 52. Institute for Empirical Macroeconomics, Federal Reserve Bank of Minneapolis, 1991.

Diebold, Francis X., and Glenn D. Rudebusch. "Turning Point Prediction with the Composite Leading Index: An Ex Ante Analysis," in Kajal Lahiri and Geoffrey H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records.* Cambridge: Cambridge University Press, 1991, pp. 231–56.

Karamouzis, Nicholas, and Raymond Lombra. "Federal Reserve Policymaking: An Overview and Analysis of the Policy Process," *Carnegie-Rochester Conference Series on Public Policy,* vol. 30 (Spring 1989), pp. 7–62.

Klein, Philip A. "Leading Indicators of Inflation in Market Economies," *International Journal of Forecasting,* vol. 2 (1986), pp. 403–12.

Koenig, Evan F., and Kenneth M. Emery. "Misleading Indicators? Using the Composite Leading Indicators to Predict Cyclical Turning Points," Federal Reserve Bank of Dallas *Economic Review,* July 1991, pp. 1–14.

Lucas, Robert E., and Thomas J. Sargent. "After Keynesian Macroeconomics," Federal Reserve Bank of Minneapolis *Quarterly Review,* vol. 3 (Spring 1979), pp. 1–16.

McNees, Stephen K. "The Accuracy of Two Forecasting Techniques: Some Evidence and an Interpretation," *New England Economic Review,* March 1986, pp. 20–31.

Mitchell, Wesley Clair. *Business Cycles and Their Causes.* Berkeley: University of California Press, 1941.

Moore, Geoffrey H. "New Developments in Leading Indicators," in Kajal Lahiri and Geoffrey H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records.* Cambridge: Cambridge University Press, 1991, pp. 141–47.

_____. *Business Cycle Indicators, Volumes I and II: Contributions to the Analysis of Current Business Conditions.* Princeton: Princeton University Press, 1961.

Neftci, Salih N. "Optimal Prediction of Cyclical Downturns," *Journal of Economic Dynamics and Control*, vol. 4 (August 1982), pp. 225–41.

Niemira, Michael P., and Giela T. Fredman. "An Evaluation of the Composite Index of Leading Indicators for Signaling Turning Points in Business and Growth Cycles," *Business Economics*, vol. 26 (October 1991), pp. 49–55.

Niemira, Michael P., and Philip A. Klein. *Forecasting Financial and Economic Cycles.* New York: Wiley, 1994.

Roth, Howard L. "Leading Indicators of Inflation," in Kajal Lahiri and Geoffrey H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records.* Cambridge: Cambridge University Press, 1991, pp. 275–301.

Sims, Chris. "Comment" [on "New Indexes of Coincident and Leading Indicators"]. *NBER Macroeconomics Annual*, 1989, pp. 395–97.

Stock, James H., and Mark W. Watson. "New Indexes of Coincident and Leading Indicators," *NBER Macroeconomics Annual*, 1989, pp. 351–94.

_____. "A Procedure for Predicting Recessions with Leading Indicators," in Stock and Watson, eds., *Business Cycles, Indicators, and Forecasting.* Chicago: University of Chicago Press, 1993, pp. 95–156.

Webb, Roy H. "Forecasts of Inflation from VAR Models," *Journal of Forecasting*, vol. 14 (May 1995), pp. 267–86.

Zarnowitz, Victor. "Composite Indexes of Leading, Coincident, and Lagging Indicators," in *Business Cycles: Theory, History, Indicators, and Forecasting.* Chicago: University of Chicago Press, 1992a.

_____. "Cyclical Aspects of Cost and Price Movements," in *Business Cycles: Theory, History, Indicators, and Forecasting.* Chicago: University of Chicago Press, 1992b.