Monetary Policy in a Low Inflation Environment

J. Alfred Broaddus, Jr.

It's a pleasure to be here today and to have this opportunity to comment on conducting monetary policy in a low inflation environment. I've been at the Fed for more than thirty-two years and have had the privilege of either advising monetary policymakers or being a policymaker myself throughout my career. It has been quite a ride. For much of this period the Fed was struggling either to prevent inflation from rising further or to bring it down. As you know, over the last several years we have succeeded in reducing the inflation rate to about $1\frac{1}{2}$ percent as measured by the core personal consumption expenditures (PCE) price index—today's favored inflation index—and in stabilizing the rate at that level. In the parlance of the day, we have finally attained "price stability," meaning both low *actual* inflation and the credible expectation in the minds of financial market participants and the general public that it will persist, which together constitute the monetary policy equivalent of finding the Holy Grail.

In my brief remarks I want to do four things. First, I will compactly review several key aspects of the evolution of monetary policy over the last thirty years. To appreciate fully the nature of the challenge that lies ahead, it is essential to understand *how* price stability was lost in the 1970s and regained in the '80s and '90s. Second, I will try to convey the essence of the current strategy of Fed monetary policy. I'll then close with a brief discussion of the challenges monetary policymakers face in today's low inflation environment, as I see them, and a pitch for explicit inflation targeting as a means of preserving the substantial improvement in the effectiveness of monetary policy during the Volcker-Greenspan years. As usual, these views are my own

This article is the text of a speech given by J. Alfred Broaddus, Jr., president of the Federal Reserve Bank of Richmond, before the Public Policy Forum of the American Finance Association on January 3, 2003. The author wishes to thank his colleague Marvin Goodfriend for his assistance in preparing the article. The views expressed herein are those of the author and are not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

and don't necessarily reflect those of my Federal Open Market Committee (FOMC) colleagues. Also as usual, my views have been strongly influenced by discussions with, and the writings of, my longtime Richmond Fed colleague Marvin Goodfriend—in particular a preliminary version of a paper he will deliver at a National Bureau of Economic Research conference on inflation targeting later this month. He is not necessarily responsible, however, for anything I say here today.

1. HOW PRICE STABILITY WAS LOST AND REGAINED

It is probably fair to say that, after periods of moderate inflation in the 1950s, the economy returned to virtual price stability in the early 1960s. The core CPI inflation rate fluctuated in a narrow 1.0 to 1.6 percent range between 1960 and 1965. Subsequently, as you all know, it increased steadily to double-digit levels in the mid-1970s and again in the late '70s.

This extraordinary and debilitating increase in inflation has been attributed, in whole or in part, to many things: the two oil price shocks in the '70s, excess demand associated with the Vietnam War buildup and the Great Society social programs, the ineffectiveness of the Nixon Administration's price control program and the Ford Administration's "Whip Inflation Now" program, and even the failure of anchovy harvests off the coast of South America. Both theory and historical evidence, however, indicate that inflationary monetary policy was the central culprit.

The failure of monetary policy to contain inflation in this period can be approached from several directions. Economists of a monetarist persuasion argue that persistently above-target money supply growth, and the practice of adjusting the money supply target's base up each year to accommodate the preceding year's upside miss, was the principal operational deficiency. Currently, the more mainstream view focuses on the failure of the Fed's short-run interest rate policy to counter the rise of inflationary pressures promptly during business expansions. The 1970s and early 1980s are sometimes referred to as the period of "go-stop" monetary policy. Concerned about the potential impact of policy tightening on employment and production, the Fed would wait until a broad public consensus emerged that inflation was a serious problem before acting decisively to contain it. By then, however, it was generally too late to bring inflation down via tighter monetary policy without at the same time touching off a recession. The cycles surrounding the 1980 and 1981–82 recessions illustrate this pattern especially well.

More fundamentally, however, it is not an exaggeration to say that Fed monetary policy lost all or most of its *credibility* as an effective force against inflation in this period. As inflation began to rise, financial markets and the public—even in the early stages of expansions—quickly revised their inflation expectations upward. This reinforced the upward pressure on current

inflation, pushed up long-term interest rates, and in general helped foster the macroeconomic malaise described by the term *stagflation*.

Perhaps the most important lesson for monetary policy from this experience is how difficult and costly it is for the Fed to rebuild its credibility for low inflation once it has been lost—especially when, for all practical purposes, it has been totally lost, as in the late '70s and early '80s. Led by Chairman Volcker, the Fed had to raise nominal short-term interest rates to unprecedented levels and essentially induce one of the longest and deepest recessions in the entire post–World War II era just to *begin* the process of restoring its credibility for low inflation. It is highly doubtful that the process could have begun without this costly recession, in which real GDP declined 2.8 percent and the unemployment rate rose to 10.8 percent. And it has taken the Fed about twenty years—nearly a quarter century—to complete the process.

The essence of this process, in my view, has been the Fed's demonstration, particularly in two episodes, that it *can* preempt an increase in inflation without precipitating a recession, and its success in recent years in convincing the markets and the public that it *will* routinely do so in the future. One of these episodes came early in the process, in 1983 and 1984, as the economy recovered from the 1981–82 recession; the other was in 1994 when the recovery from the 1990–91 recession finally began to gather momentum. In both cases, incipient inflationary pressures were quickly picked up by financial markets, which produced what Goodfriend calls "inflation scares," characterized by sharply rising nominal bond rates. In both instances the Fed acted swiftly and decisively to preempt inflation. The 1994 episode was especially important since it occurred at a time when Fed policy had become much more transparent than earlier, as evidenced by its decision that year to announce publicly its federal funds rate target immediately following each FOMC meeting.

2. THE CURRENT STRATEGY OF MONETARY POLICY

The U.S. economy has now enjoyed virtual price stability since about 1996. There seems to be a growing consensus currently among monetary policymakers, close observers of the policy process in financial markets, Congress and the press, and individual Americans interested in policy that price stability and the Fed's credibility for low inflation should be sustained. This consensus is based partly on a broader public appreciation of the high costs of reestablishing lost credibility, as described above. More fundamentally, however, it appears to reflect a recognition that the Fed's revived credibility is beneficial to the economy—specifically, that the Fed can contribute meaningfully to an improved longer-term U.S. economic performance in the form of an increase in the sustainable growth of production and higher employment. Further, the public seems less concerned than earlier about possible short-run costs of low inflation, in terms of lower growth, perhaps because the transition to

low inflation has now been accomplished. And since low inflation is broadly expected to persist, the public would be surprised and disappointed if it were lost. Consequently, the consensus arguably sharpens the Fed's accountability for maintaining low inflation.

Against this background, I sense the emergence within the Fed of a more cohesive *strategy* of monetary policy than at any other time in the last three decades. To my mind it consists of two elements: (1) a strong commitment to maintaining high credibility for low inflation permanently, and (2) active management of real short-term interest rates to help stabilize the economy in the short run. Regarding the first, Goodfriend argues in his forthcoming paper that the Fed is now practicing "implicit" inflation targeting. As he points out, with the core inflation rate in the 1 to 2 percent range since the mid-1990s, it is hard to imagine the Fed now accepting a sustained inflation rate significantly above 2 percent. Nor would it be likely to accept a sustained rate significantly below 1 percent given the increased sensitivity to the risk of deflation and the proximity of the zero bound on nominal interest rates.

I personally believe that "implicit longer-term inflation targeting" is an accurate description of the first element of the Fed's current strategy. It is important to stress, however, that its ultimate objective is not price stability and high Fed credibility for its own sake, but the optimal financial foundation these conditions provide for strong real growth and high employment.

Moreover, these conditions enable the Fed to pursue the second element of the strategy: active countercyclical short-term interest rate policy. When the Fed's credibility was very low in the 1970s and early '80s, it was difficult—perhaps impossible—to conduct countercyclical interest rate policy effectively. With the Fed's long-run objective for inflation still unclear, the public could not confidently deduce the longer-term ramifications of particular short-term policy actions, and the Fed, in turn, could not confidently predict the public's reaction to its actions. With its credibility for low inflation now well established, the Fed *can* act both more promptly and more aggressively to counter the effects of unanticipated shocks and thereby stabilize the economy in the short run. Beginning exactly two years ago today, the Fed began to ease policy in response to the softening of the economy in the second half of 2000. It accelerated the easing process in the wake of 9/11, and over the course of the two-year period has reduced the federal funds rate $5\frac{1}{4}$ percentage points, from $6\frac{1}{2}$ percent to its present level of $1\frac{1}{4}$ percent. This is arguably the most aggressive series of policy easings taken to cushion a softening economy in the Fed's history and may well account for the apparent brevity of the recent recession despite the extraordinary decline in the stock market, 9/11, and other shocks.

The two elements of the Fed's current strategy, then, are complementary and mutually reinforcing. Implicit inflation targeting enhances the effectiveness of countercyclical interest rate policy. Conversely, active countercyclical

policy makes implicit inflation targeting acceptable, since the ability to act aggressively to stabilize the economy in the short run provides a clear and easily understood rationale for containing inflation.

3. CHALLENGES IN A LOW INFLATION ENVIRONMENT WITH IMPLICIT INFLATION TARGETING

After such a long struggle, one might expect that Fed monetary policymakers would be relatively comfortable now that price stability has been achieved and credibility for low inflation has been reestablished. And I think most, if not all, policymakers *are* more confident that the Fed can contribute constructively to the economy's longer-term performance rather than retarding it, as occurred when inflation was high and variable, and credibility was low.

But the Fed still faces significant policy challenges in the new low inflation environment. Historically, little practical attention has been given to the possibility of excessively sharp *dis*inflation and *de*flation. And with the press here I need to emphasize at the outset that I do not believe deflation is a serious present risk to the economy. But policymakers obviously need to think more about how they would deal with a deflationary threat, should one emerge unexpectedly, when inflation is in a 1 to 2 percent range than when it is at 6, 7, or 8 percent. This is especially so with the nominal funds rate, our principal short-term policy instrument, only 125 basis points above zero.

We have been thinking about it, and I am quite confident that we could deal with a deflationary threat successfully. In October 1999 the Fed sponsored a conference in Woodstock, Vermont, on conducting monetary policy in a low inflation environment attended by a large number of leading monetary economists. The participants gave substantial attention to deflation and how to deal with it should it arise in the future. More recently, Fed Governor Ben Bernanke nicely summarized current thinking on this issue in a speech to the National Economists Club. There is now broad agreement that the most effective way to deal with deflation is to prevent it from developing in the first place. In the present situation, the Fed's aggressive easing over the last two years appears to have preempted any significant drift toward either excessive disinflation or deflation. Moreover, even if disinflation unexpectedly intensified and the funds rate were reduced close to the zero bound, the Fed would still have a number of channels available to reestablish a comfortably positive inflation rate. For example, it could increase broad liquidity by purchasing long-term bonds.

The other potential policy challenge I see in today's low inflation environment is how to handle an incipient *increase* in inflation above its implicit target range. This possibility is not on many radar screens currently, but it is obviously a longer-term risk—arguably the most likely longer-term risk. I believe that the policy experience of the 1970s, '80s, and '90s summarized

above argues strongly for prompt action to preempt any sustained increase in inflation. If policymakers had precise, detailed foreknowledge of the relative costs, in terms of lost production, of alternative paths back to price stability, it might be feasible to tune the return more finely. There is little evidence, however, that we have such knowledge. Hence, it seems reasonable to resist any deviations from price stability promptly and strongly—and preferably preempt them altogether.

4. CONCLUSION—SUSTAINING THE PROGRESS

Hopefully these comments have convinced you that the conduct of monetary policy in the U.S. has improved significantly during the last two decades, and that this improvement of policy holds out the prospect of an improved longer-term economic performance going forward. The trick now is to sustain the progress. Much of the progress, in my view, is due to the exceptionally strong leadership since 1979 of, first, Paul Volcker and now Alan Greenspan. But, ultimately, high-quality monetary policy—i.e., sustained credibility for low inflation as a foundation for strong real growth—is too important to be dependent on exceptional leadership alone, which, after all, cannot be guaranteed over the long pull. The progress in recent years needs to be institutionalized—"locked in"—in some manner.

There are probably several ways this could be accomplished. Earlier I referred to one element of the Fed's current policy strategy as implicit inflation targeting. My personal preference for "hardening" our credibility is to make the implicit target both explicit and quantitative—specifically, 1 to 2 percent, based on the core PCE index. Explicit, quantitative inflation targeting is practiced by a number of other leading central banks around the world, and it would be consistent with the continuing evolution of Fed policy toward greater transparency and accountability. Most importantly, it would be a strong and visible step toward ensuring that the Fed's current high credibility for low inflation will be maintained indefinitely so that we can make our strongest possible contribution to maximum sustainable growth in the long run and economic stability in the short run.

Banking and Commerce: Tear Down This Wall?

John R. Walter

Many U.S. firms include both commercial and nonbank financial units. For example, General Motors Corporation encompasses not only units that manufacture automobiles but also those, such as General Motors Acceptance Corporation, that gather funding and make loans to individuals and businesses. Firms that handle both commercial and financial activities appear to reap significant benefits that create the appeal of such combinations. One byproduct of a commercial firm's activities may be information about its customers' financial situation. The financial affiliate might then use this information to inexpensively target products to particular customers, benefiting both the financial firm and its customers.

While finance/commerce combinations are widespread, combinations between *banks* and commercial firms are typically prohibited under various U.S. laws. Banks are distinguished from other financial firms by their ability to gather funding by issuing government-insured deposits such as checking and savings deposits. Despite prohibitions of banking/commerce combinations, firms have managed to find loopholes. Until recently the unitary thrift loophole was a popular means of circumventing the banking/commerce wall. The loophole allowed commercial companies to start or buy one, and only one, thrift (i.e., a savings bank or savings and loan association, both of which issue government-insured deposits), using the thrift as a conduit for providing financial services.

The unitary thrift loophole was closed in 1999, but another loophole remains open. Federal banking law allows commercial firms to own industrial loan corporations—essentially banks with somewhat restricted deposit-taking powers. (For additional discussion of the unitary thrift and industrial loan corporation loopholes, see the Appendix.)

The author benefited greatly from discussions with Kartik Athreya, Tom Humphrey, Ray Owens, and John Weinberg. The views expressed herein are not necessarily those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

Given the apparent benefits of combinations, why prohibit or restrict them? What hazards result from banking/commerce combinations? Traditionally, discussions of the threats focused on conflicts of interest and diminished competition. More recently, observers have been concerned that combinations might increase deposit insurance claims and expand the universe of economic activities protected by the government safety net. As will be discussed presently, the traditional concerns seem less relevant given the level of competition banks face in today's banking market. Over the last twenty-five years, competition has expanded as restrictions were eliminated on banks' ability to operate across state lines and to offer market rates on deposits. Also, new nonbank firms have arisen offering financial products competitive with most banking products. The concerns over increased deposit insurance claims and expansion of the safety net remain quite relevant. Nevertheless, over the last decade a number of legislators have argued for removal of the banking/commerce wall. I will analyze the threats and suggest some restrictions that would be necessary if the wall were removed.

1. THE STATUTES FORMING THE WALL

The building blocks of the wall between banking and commerce are various federal and state laws. The laws prevent banks from engaging in commercial activities. They also prevent banks from owning subsidiary commercial companies and from being owned by companies conducting commercial activities. Specifically, the building blocks are the National Bank Act of 1864, state banking laws, the Federal Deposit Insurance Corporation Improvement Act of 1991, and the Bank Holding Company Act of 1956. (For a detailed review of these statutes and their motivations see the Appendix.)

The National Bank Act limits the powers of national banks and their subsidiaries. National banks are those chartered and regulated by the U.S. Treasury Department's Office of the Comptroller of the Currency. The Act states that "a national banking association shall... have power to... exercise... all such incidental powers as shall be necessary to carry on the business of banking" (12 U.S.C. 24). While over the years the courts and the Comptroller have wrangled over the meaning of the phrase "business of banking," national banks have been allowed to engage in businesses similar to traditional banking services but not other commercial activities. This restriction of powers extends not only to activities of banks, but to activities conducted by subsidiaries owned by banks.

State banking statutes typically set limits on the nonbank activities of state banks and their subsidiaries, similar to the limits on national banks. Yet, over the years, a number of states have authorized activities well beyond those allowed national banks, some of which typically would be considered commercial powers. Concerns that these new powers might endanger state-chartered

banks, and thereby the taxpayer-backed deposit insurance fund, led Congress to restrict state legislatures' ability to grant powers to state-chartered banks. Specifically, under the Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA), insured state-chartered banks are prohibited from engaging in any activities impermissible for national banks unless the FDIC rules that such activities pose no threat to the deposit insurance fund.

Those who advocate removing the banking/commerce wall typically do not argue for allowing banks to conduct commercial activities or own commercial firms. Instead, they focus on allowing companies that own banks—that is, bank holding companies—to own commercial companies, too. In other words, they do not condone direct bank involvement in commerce but find it acceptable for bank holding companies to own commercial companies, allowing banks to affiliate with commercial companies.

The banking/commerce restrictions in the Bank Holding Company Act of 1956 were based on the view that "bank holding companies ought to confine their activities to the management and control of banks." The Act restricted bank holding companies such that they "would no longer be authorized to manage or control nonbanking assets unrelated to the banking business" (*U.S. Code: Congressional and Administrative News* [1956, 2482, 2484]).

To enforce this restriction, the Act defines a bank holding company as any company that owns a bank.[1] It prohibits bank holding companies from engaging in nonbanking activities. The Act allows an exception to the nonbanking prohibition in cases in which the Board of Governors of the Federal Reserve System determines the nonbanking activity "to be so closely related to banking as to be a proper incident thereto" (12 U.S.C. 1843c). Typically the Board defines activities closely related to banking as only those activities traditionally performed by banks. Certain additional activities also have been allowed, however, in cases in which they are tied to banking. For example, the Board has determined that a bank holding company may own a data-processing firm if it is primarily engaged in processing financial, banking, or economic data (Spong [2000, 156]).

The Gramm-Leach-Bliley Act, enacted in 1999, added securities underwriting and dealing, as well as insurance, to the list of activities in which banks—through bank-owned subsidiaries—and bank holding companies could engage. Those bank holding companies choosing to engage in securities or insurance activities are called *financial holding companies*. The Act also expanded the activities of bank-owning companies to include most financial activities, those determined by the Board and the Treasury Department to be "financial in nature," "incidental to a financial activity," or "complementary to

[1] The Bank Holding Company Act of 1956 applied only to companies owning two or more banks. Amendments enacted in 1970 extended the Act's provisions to single-bank holding companies.

a financial activity." Nonetheless, while authorizing a wide range of financial activities, the Act leaves in place the wall between banking and commerce.

Deciding whether an activity is commercial instead of banking or financial and whether it can be conducted by a banking company is often not simple. The decision was difficult under the old Bank Holding Company Act standard of being *closely related* and remains so under the new standard of being *financial in nature, incidental to a financial activity,* or *complementary to a financial activity*. The placement of an activity on one side or the other of the banking/commerce wall can be controversial and contentious. For example, in December 2000 the Board of Governors of the Federal Reserve System and the Secretary of the Treasury jointly released for comments a proposal to permit bank subsidiaries and financial holding companies to engage in real estate brokerage and management. Real estate industry trade groups quickly objected to the proposal, arguing that it would amount to an illegal mixture of banking and commerce. Banking trade groups argued in favor of the proposal. Beyond the real estate industry's objections, legislators introduced bills in both the U.S. House of Representatives and the Senate to prohibit these real estate activities in financial holding companies. In April 2002 the Secretary of the Treasury announced plans to put off a decision on the proposal until 2003.

2. RATIONALES FOR THE WALL

Three reasons are typically cited for maintaining a wall that prohibits banking/commerce combinations: conflicts of interest, monopoly power, and risk to the taxpayer-backed deposit insurance fund.[2] These three justifications will be examined below. As it turns out, because banking markets appear fairly competitive, the first two seem of relatively minor import. The third remains quite significant.

Conflicts of Interest

Observers have at times raised concerns over conflicts of interest that might arise if banks and commercial firms are owned by the same firm. They argue that such concerns justify keeping banking and commerce separate. Three conflicts have been described. First, a bank affiliated with a commercial firm would tend to deny loans to the affiliate's competitors. Second, a bank might use access to insured deposits to provide below-market-rate funding to its affiliates while charging higher interest rates to unaffiliated borrowers. Third, in the legislative history of the Bank Holding Company Act, legislators noted

[2] For discussions of these three reasons for maintaining the wall, see Krainer (2000, 21–23); Halpert (1988, 490–517); and U.S. GAO (1987, 11–12).

that a bank with a commercial affiliate might deny loans to individuals who do not purchase goods from the affiliate.

It seems natural that removing the banking/commerce wall would allow the first conflict to arise, since a bank with a commercial affiliate, say a restaurant, would not wish to provide funding to competing restaurants. Helping the competitor would tend to lower the profits of the affiliated restaurant. Yet, if competition is reasonably strong, denying loans to competitors only lowers overall profits of the consolidated banking/restaurant firm. If there are alternative lenders over which the affiliated bank has no price advantage, the competing restaurant would get a loan anyway and at the same interest rate the affiliated bank would offer. So, by failing to make the loan, the bank loses any profit it might have made on that loan, hurting the bank. Yet the affiliated restaurant suffers a loss in profits regardless.[3] Therefore, if competition is strong, this potential conflict of interest is unlikely to present a problem and cannot justify maintaining the banking/commerce wall.

But is banking competition strong? Since the 1970s, restrictions on bank-versus-bank competition have been greatly reduced. Restrictions on banks' ability to compete for deposits outside of their local markets, or at least outside of their home states, were severe before the late 1970s. While these restrictions did not apply to bank lending, banks generate a good bit of their lending in the same markets in which they gather loans. Consequently, these restrictions likely limited loan competition as well. Banks' ability to open branches statewide was greatly enhanced in the 1980s as many states removed branching restrictions. Restrictions on operating across state lines began to fall in the mid-1980s and were almost completely removed by the Riegle-Neal Interstate Banking Act of 1994. As a result, banks that had been protected from competition because of branching restrictions became subject to competitive pressure from nonlocal banks by the mid-1980s.

While the elimination of branching restrictions opened local banking markets to greater competition, other market and technological developments have expanded competition further. Consequently, if a bank denies a loan to a business firm because it competes with the bank's affiliate, that firm can find numerous alternative sources of funding in today's more competitive loan markets.

Competition among those who would lend to business borrowers has expanded along several dimensions. For large business borrowers, banks faced growing competition from the debt markets as commercial paper and bond issues increased significantly relative to bank lending over the last twenty-five years. While small businesses cannot issue commercial paper or bonds, today's small businesses have access to loans from a wide range of lenders.

[3] Owens (1994) makes this argument for bank lending in real estate.

The largest banks aggressively court small business borrowers throughout the country via their Web sites and toll-free phone lines. Further, small businesses enjoy a range of choices of nonbank lenders, including finance companies and leasing companies. Clearly, today's borrowers, both large and small, have a plethora of borrowing opportunities because of the competitive loan market.

If, in spite of these factors, some banking markets remain uncompetitive, policymakers can address the problem directly by removing any remaining barriers to entry. Alternatively, they might tackle monopoly power through antitrust enforcement. Maintaining a wall that separates banking and commerce at best addresses a symptom of an uncompetitive market rather than the lack of competition itself.

Still, using the restaurant example, one might argue that the bank with an affiliated restaurant may for some reason have a cost advantage over its bank competitors that lack such an affiliation. One reason for a cost advantage is that the bank acquires information about the restaurant business through its affiliation. While prohibiting affiliations might eliminate the advantage this bank (and its affiliates) has over competitors, the restriction would diminish economic efficiency because the least costly and most efficient means of producing banking services—through restaurant affiliation—would be denied.

Additionally, if bank/restaurant affiliations are allowed, banks that lack a restaurant affiliate can simply overcome the cost disadvantage by affiliating. This is just what the banking industry did when banks established branches in grocery and discount stores, though they did it through leasing agreements rather than affiliation. After perceiving the advantage gained by the innovative bank that first placed branches in such stores, other banks followed suit to achieve the same advantage. Soon the advantage was dissipated by competition.

Aside from the situation whereby a bank might deny loans to its affiliate's competitors, some observers note a second conflict of interest. They argue that banks may have access to inexpensive funding because of underpriced deposit insurance, and that this funding might be granted to banks' affiliates but not to other borrowers. Such funding would give affiliates an advantage over firms not so affiliated.[4] As long as the banking market is competitive, however, every firm that borrows from a bank gets equivalent access to low-cost funding whether affiliated with a bank or not. Access is equivalent because a bank only hurts itself (i.e., lowers its revenues) by not lending to its affiliate's competitors on equivalent terms to those offered its affiliate. If the bank with an affiliate does not lend to its affiliate's competitors, other banks would take those customers and profit from doing so. Further, Section 23A of the Federal Reserve Act, applicable to all banks, restricts the amount of such affiliate

[4] For discussions of the argument that access to bank funding could give bank affiliates an advantage, see Board of Governors (1987, 500) and Macey and Miller (1992, 377).

funding to at most 10 percent of the bank's capital. Section 23B of the Act requires such funding be on market terms.

In the legislative history of the Bank Holding Company Act, members of Congress describe a third possible conflict of interest. Specifically, they argue that a bank with a commercial affiliate might deny loans to individuals who do not purchase goods from the affiliate. The Senate Banking Committee's report that analyzes the features of the bill that later became the Bank Holding Company Act describes the concern as follows:

> The committee was informed of the danger to a bank within a bank holding company controlling nonbanking assets, should the company unduly favor its nonbanking operations by requiring the bank's customers to make use of such nonbanking enterprises as a condition to doing business with the bank. The bill's divestment provisions should prevent this fear from becoming a reality. (*U.S. Code: Congressional and Administrative News* [1956, 2486], as cited in Halpert [1988, 500])

Tying a loan (or other service) to the purchase of another product can only benefit a bank if the bank has monopoly power in its loan market. If it faces competition, denying loans to individuals who are not its affiliate's customers only hurts the bank, and so would not be undertaken. The bank is hurt because it forfeits revenues and helps its bank competitors who would make the loans (Owens [1994]). As noted earlier, when the Bank Holding Company Act was passed in 1956, banking markets were heavily regulated. Entry was restricted and prices were controlled. Monopoly power may have been significant, but most such restrictions have been removed. Additionally, even if banks maintain monopoly power in credit markets, the commercial affiliate must also have market power in order for tying to make consumers worse off. In the case in which the combined firm has market power in the banking and commercial markets, only under limited circumstances are consumers actually made worse off. In other cases consumers are unhurt by tying (Weinberg [1996]). Regardless, current statutes make tying by banking companies illegal.[5]

Proliferation of Monopoly

Some observers argue that, in addition to conflicts of interest, preventing the exercise of monopoly power is another reason for the banking/commerce separation. The legislative history of the Bank Holding Company Act makes it clear that Congress intended the Act to guard against the proliferation of

[5] See Weinberg (1996) for a review of bank anti-tying statutes and the economics of tying. See Krainer (2000, 22) for a discussion of banks denying credit to their commercial affiliates' competitors. In 1997 the U.S. General Accounting Office analyzed banks for evidence of tying bank loans to securities activities. It found little evidence of any such tying (U.S. GAO [1997]).

monopoly. For example, the Senate Banking Committee report on the conference bill notes that the Act was to provide "safeguards... against undue concentration of control of banking activities. The dangers accompanying monopoly in this field are particularly undesirable in view of the significant part played by banking in our present national economy" (*U.S. Code: Congressional and Administrative News* [1956, 2482–83]). Because such language is vague, it is difficult to determine whether the *undue concentration* discussed refers to horizontal or conglomerate concentration. As a result, it is uncertain whether proliferation of monopoly was behind the Bank Holding Company Act's banking/commerce restrictions. *Horizontal concentration* means combining a number of banks under one bank holding company such that this holding company controls a high percentage of banks. *Conglomerate concentration* means combining both banks and nonbanks under one holding company so that one conglomerate controls a significant percentage of business firms in banking and a nonbanking industry, or in several nonbanking industries.

Observers since have argued that Congress was indeed concerned with conglomerate concentration. A case in point was a 1974 Federal Reserve's denial, under the Bank Holding Company Act, of an application by BankAmerica Corporation to form an overseas joint venture with Allstate Insurance. Here the Fed said that "close working relationships abroad between large U.S. banking organizations and large U.S. insurance companies could in time weave a matrix of relationships... that could lead to an *undue concentration* of economic resources in the domestic and foreign commerce of the United States... not... consistent with the purposes of the Bank Holding Company Act."[6]

Frequently, when advocating the separation of banks from nonbanks on the basis of monopoly, proponents have argued that the combination would allow the monopoly power that banks hold in their product markets to be used by combined firms to raise prices in other areas.[7] But as discussed earlier, while in 1956 concerns about monopoly may have motivated Congress, during the 1970s and 1980s competition expanded greatly, among banks and between banks and nonbanks. Expanded competition significantly reduced any opportunity banks might have had to exercise monopoly power in banking services and to expand it to other businesses with which they might combine. Congress seems to have been cognizant of these changes. When Congress passed the Gramm-Leach-Bliley Act in 1999, it allowed combinations of large banks with large insurance and large securities firms. These were exactly the types

[6] Board of Governors (1974, 519) (italics added for emphasis). For a similar argument on another application, see Board of Governors (1981, 451), as cited in Halpert (1988).

[7] Halpert (1988, 500–505) discusses the argument that banking monopoly might proliferate into nonbank businesses.

of combinations—that is, large banks with large nonbanks—denied earlier by regulators based on undue concentration language in the Bank Holding Company Act.

Safety Net Concerns

For the reasons discussed above, conflicts of interest and fears of expanding monopoly power alone are probably insufficient reasons to maintain the current wall separating banking and commerce and deny firms the opportunity to benefit from combinations. Nevertheless, there is another set of hazards that could justify the continued presence of the wall separating banking and commerce or at least require that significant precautions be taken if the wall is removed. The hazards come in three forms, discussed in the following paragraphs, and each involves an increased chance of bank failures and a subsequent bailout financed by taxpayers. If bailouts occur, the government safety net, meant to protect bank depositors, could be extended to creditors of commercial companies. If extended, too many resources might flow to bank-affiliated commercial companies, and economic efficiency would be diminished. The threat to the safety net could arise because (1) losses might be shifted to banks to protect a combined firm's reputation with investors, (2) losses might be shifted to banks to take advantage of shareholder limited liability, and (3) the combined firm's riskiest assets might be shifted to the bank. Of the three hazards, the first two could justify continued banking/commerce separation. The third cannot justify separation but is discussed below because it is often mentioned as a hazard of bank/nonbank affiliations.

Loss Shifts that Protect Reputation

If banking/commerce combinations are allowed, a combined company can be expected under certain circumstances to withdraw resources from its bank to hide problems in its commercial subsidiary, damaging bank safety. The holding company is likely to choose this course when it can hide commercial subsidiary losses from investors and analysts by shifting commercial subsidiary losses to the bank. The holding company would benefit by hiding the loss, which if revealed would likely be perceived as negative information about the ability of the firm's management and the riskiness of its operations; that is, it would damage the firm's reputation. Such negative information would lead creditors to demand higher interest rates, lowering future profits. Yet, as discussed presently, while shifts to hide losses can be detrimental to banks, they can just as easily be beneficial: holding companies could choose to shift bank losses to commercial subsidiaries. Consequently, a concern that bank holding companies might engage in loss shifts is no reason to prohibit banking/commerce combinations. Instead, if one is to argue that the danger

of shifts can justify the banking/commerce wall, one must believe that loss shifts are more likely to flow toward bank than commercial subsidiaries.

Reputation-protecting shifts that can work to the detriment of bank health are likely to occur when two conditions are met. First, the commercial subsidiary suffers a loss large enough to create its insolvency. Second, the loss, if shifted to the bank subsidiary, would avert the bank's insolvency. The second condition would generally be met if the bank's net worth is considerably larger than that of the commercial subsidiary (before and after the shift). Under these conditions, a shift of a commercial subsidiary's loss to the bank would protect the holding company's reputation. An insolvency is certain to draw negative outsider attention, since it will likely involve either debt renegotiation or bankruptcy. In contrast, a mere loss or perhaps just an increase in the bank's reported expenses can be expected to draw far less attention. Even so, a loss shift necessarily weakens the bank.

Even if the commercial subsidiary experiences a loss that does not lead to its insolvency but is still significant, such a loss could still shift to a larger bank subsidiary. The bank holding company might choose to shift the loss because it might be less noticeable on the books of a large firm than on those of a smaller one. In addition, observers have traditionally argued that banks may have more opaque balance sheets than do commercial firms, so that losses can be better hidden in a bank subsidiary. Some recent research appears to support this view of bank opacity.[8] If banks are indeed more opaque, then losses are more likely to go unnoticed if shifted to the bank.

Nevertheless, the existence of this incentive to shift losses in order to hide them does not imply that commercial and banking firms should be kept separate. Under one set of circumstances already discussed—when the bank's net worth (meaning its capital) is larger than the commercial subsidiary's—a bank holding company can hide the loss by shifting it to the bank. However, under an equally likely set of circumstances—when the commercial subsidiary's capital exceeds the bank's capital—there is no benefit from shifting commercial subsidiary losses to the bank. Instead, if the bank produces losses, the bank holding company can benefit by shifting bank losses into the commercial firm. Therefore, prohibiting banking/commercial affiliations will not necessarily improve bank safety or protect taxpayers and the FDIC from losses.

Further, creditors of banks as well as commercial firms affiliated with banks are likely to be well aware of the incentive to shift losses in order to

[8] Morgan (2000) finds that banks and insurance companies are inherently more opaque than other firms. Morgan checks for opaqueness of banks and insurance firms versus nonfinancial firms by measuring the frequency of disagreements between the two major ratings agencies in their ratings of banks, insurance companies, and other firms. He finds that the two agencies disagree more frequently over banks (and over insurance companies) than over commercial firms. Morgan contends that the cause of the disagreement is the difficulty of evaluating opaque bank balance sheets.

hide them. By charging higher interest rates, both types of creditors will penalize affiliations that might shift losses to the detriment of their debtor (either commercial firm or bank). For example, a bank's creditors would be likely to view a combination with a risky commercial firm—that is, one that might produce shiftable losses—as dangerous. They would demand the bank pay an increased interest rate if such an affiliation were undertaken. Moreover, if the affiliated commercial firm's riskiness increased, the bank's creditors would impose an additional risk premium to account for the increased risk of a loss shift. If the risk became large, the increased premium might be sufficient to cause the holding company to divest either the commercial firm or the bank. The affiliated commercial firm's creditors would do the same.

Yet there is reason to think that shifts would tend more frequently to deplete bank resources rather than commercial firm resources. While creditors of commercial subsidiaries of bank holding companies would demand higher risk premia for affiliations likely to produce losses that can be shifted to the commercial affiliate, bank creditors have a reduced incentive to do so. Many of a bank's creditors—those holding insured deposits in the bank—would demand no additional compensation when the bank affiliates with a risky commercial firm. If losses are shifted to the bank, weakening it, its government-insured deposits are no less likely to be repaid. Therefore, while the creditors of commercial affiliates penalize risky combinations, those of banks do not. Consequently, combinations that could lead to loss shifts toward banks are likely to be more common than combinations that could produce loss shifts toward commercial firms.

Note that if bank deposit insurance premia were closely tied to individual bank riskiness and accounted for the risk of loss shifts, higher premia would discourage affiliations that could be risky to banks. As discussed in more detail below, observers typically argue that deposit insurance premia are not closely tied to bank risk.

Loss Shifts that Take Advantage of Limited Liability

While the previous section discusses a set of incentives that could lead a bank holding company to shift losses from a less capitalized subsidiary to a more capitalized one, another set of incentives can produce the opposite result. Under certain circumstances, by shifting losses from a more capitalized subsidiary to a less capitalized one, the bank holding company can reduce losses. The strategy, discussed below, is beneficial because of the protections offered shareholders by the principle of limited liability. In some cases it could work to the detriment of the FDIC, and ultimately, to the detriment of taxpayers. As in the case of reputation-protecting shifts, shifts that take advantage of limited liability seem at first to be just as likely to enhance bank safety as diminish it, suggesting that this argument cannot be used as a justification for maintaining the separation between banking and commerce.

On further analysis, however, it is clear that limited liability shifts would more likely work to the detriment of banks and therefore to the detriment of the FDIC and taxpayers. Therefore, maintaining the banking/commerce wall might be justified as a means of preventing these shifts.

The following example shows that with limited liability, a bank holding company can avoid losses if it shifts them. Suppose a holding company—Alpha Conglomerate Inc.—owns two subsidiaries, Bravo Dry Cleaners and Echo National Bank. Bravo has a net worth of $100 million, while Echo's net worth is $5 million. Alpha's only assets are its investments in the stock of Bravo and Echo, and it is the sole owner of both. Consequently, Alpha's net worth is $105 million, the sum of Bravo's and Echo's net worth. If Bravo suffers a $10 million loss (say it has bankrupt commercial customers to which it has made $10 million in loans), Bravo's net worth falls to $90 million. Also as a consequence of Bravo's loss, Alpha's stockholders suffer a $10 million loss since Alpha's net worth falls to $95 million (the sum of Bravo's $90 million net worth and Echo's $5 million).

Suppose instead that Alpha could arrange to have Echo take the loss. Echo could take the loss by purchasing the loans made to Bravo's bankrupt commercial customers for $10 million, even though the loans are worthless.[9] The shift of the $10 million loss to Echo, which had only $5 million in net worth before the shift, drives it into insolvency. Bravo has a net worth of $100 million after the shift, and Echo has a net worth of negative $5 million. Based on the principle of limited liability of shareholders, however, Alpha can suffer a loss of no more than its investment in Echo, or $5 million. The shift has saved Alpha's shareholders $5 million. In this case the FDIC, which insures Echo, suffers the remaining $5 million loss. In summary, a holding company can benefit by shifting a loss when that loss is smaller than the loss-producing subsidiary's capital, but greater than the other subsidiary's capital.

Alpha's incentive to protect its reputation, as discussed in the previous section, will tend to work to prevent it from employing a shift that will produce an insolvency. Echo's insolvency is certain to damage Alpha's reputation and raise its future borrowing costs. Further, such shifts could be illegal. Nevertheless, when the benefit from shifting losses is large, the shift might be undertaken regardless of reputation or legality.

While this incentive to shift losses could endanger bank health, and in fact such a shift led to a bank failure in 1953, holding companies owning bank and commercial affiliates initially seem no more likely to shift losses

[9] As will be discussed presently, federal law restricts a bank's ability to purchase loans made by its affiliates to a small percentage of the bank's net worth. Purchases of sufficient affiliate loans that lead to the bank's insolvency would be illegal under the restrictions. Nevertheless, in some cases, such purchases have occurred.

into banks than away from them.[10] In other words, combinations of banking and commerce are just as likely to enhance bank safety as reduce it. However, there is a greater chance that shifts will work against banks; for while banks' major creditors—insured depositors—are largely indifferent about the risks that affiliations with nonbanks might impose, commercial affiliates' creditors are very interested. Creditors of commercial affiliates will penalize, with demands of higher interest rates, affiliations that increase the likelihood of an affiliate failure.

Since commercial firm losses tend to be shifted toward banks, undermining bank health, banking/commerce affiliations increase the likelihood of FDIC payouts and ultimately of taxpayer bailouts of the FDIC. Consequently, a limited liability motive for loss shifts could provide a reason to favor prohibiting banking/commerce combinations.

Beyond the cost to the FDIC and perhaps to taxpayers, who provide the backstop for FDIC insurance, there is an additional cost of loss shifts (motivated by limited liability as well as reputation protection). If creditors of bank-affiliated commercial firms believe that these firms' losses can be shifted to banks and ultimately to the FDIC, then creditors will charge bank-affiliated commercial firms lower interest rates than they would absent the perceived ability to shift. As a result of this reduced cost of capital, affiliated firms would regard projects as viable that without this taxpayer-provided subsidy would be unprofitable. In sum, too much investment capital would flow to affiliated firms, and the economy's resources would be wasted.[11] This potential for resource waste may provide further reason to prohibit banking/commerce combinations, or at least to regulate combined firms to discourage shifts.

Risk Shifts

At first blush there appears to be one additional reason to maintain the separation between banking and commerce: combinations would allow *risky assets* to be shifted from the commercial firm to the bank. Doing so increases the bank's riskiness, putting taxpayer funds at risk. This possibility has caused some observers to raise concerns about affiliations between banks and nonbanks. As a justification for maintaining the banking/commerce separation, however, the argument is unconvincing.

To lower its total funding costs, a bank holding company with a commercial subsidiary can shift the commercial firm's riskiest assets to the bank. The commercial firm must borrow using uninsured debt, while the bank can gather

[10] See Walter (1996, 22) for a discussion of the case in 1953 in which a bank failure resulted from shifts from a bank holding company's nonbank subsidiary to its bank subsidiary, apparently motivated by an attempt to take advantage of limited liability.

[11] See Walter and Weinberg (2002, 373–75) for a more complete discussion of the economic costs of government subsidization of private firms' borrowing costs.

funds by issuing insured deposits. As a result, funding costs are lowered and holding company profits are increased when the commercial firm's riskiest assets are shifted to the bank. (Note that risk shifts differ from loss shifts, discussed earlier. Loss shifts occur when the bank purchases the assets from the commercial firm at a price that produces a loss for the bank. Risk shifts occur when the bank pays a price that produces no loss for the bank, since it is insensitive to risk.) For banks' costs to be less sensitive, deposit insurance premia and other supervisor-imposed costs must be imperfectly sensitive to bank riskiness. Observers argue that this could be the case for many banks.[12]

Risk shifts, however, do not justify the banking/commerce wall because affiliation creates no more incentive to shift risks than would exist without affiliation. If the penalty for holding risky assets is lower for banks than for commercial firms (or, for that matter, for any uninsured firm), then risky assets would flow into banks even if they have no affiliates. Banks would be willing to pay more for risky assets than would other firms and would bid them away from others.

For example, imagine that a commercial firm, Juliet Tool and Die, Inc., is currently paying its creditors 15 percent in annual interest payments to raise $100,000. It uses this $100,000 to make trade credit (i.e., a loan from a seller to its customer used by the customer to purchase the seller's goods) available to its customer, Kilo Millwork. Juliet's creditors charge this high rate because they view Kilo as risky, such that Juliet's loan to Kilo heightens the chance that Juliet will itself fail; in other words, the trade credit is a risky asset. Alternatively, Lima National Bank, which pays depositors only 10 percent, can raise the $100,000 from depositors with which it can provide funds to Kilo. Because of FDIC insurance, its depositors care little about the riskiness of Lima's assets. In such a case, Lima National can be expected to approach Juliet and offer to buy its asset, the trade credit to Kilo. Lima would be willing to pay more than the asset is worth to Juliet since Lima can fund the asset less expensively than can Juliet.[13]

A holding company with a bank subsidiary and a commercial subsidiary may well benefit from having its commercial firm sell its risky assets to the bank subsidiary because of underpriced deposit insurance. But a commercial firm with no affiliated bank would find that banks would want to buy their risky assets just as well. So, if deposit insurance is underpriced, whereby it is less expensive for banks than for commercial firms to hold risky assets, preventing affiliation would not prevent risky assets from being shifted to banks.

---

[12] See Walter (1998, 2–9) for a discussion of the means by which banks can receive risk-insensitive funding.

[13] Lima would only be willing to pay more for the loan than it is worth to Juliet if Lima's deposit insurance premia do not completely account for the risk the loan adds. Still, as already noted, for many banks, insurance premia may not accurately reflect their riskiness.

3. PROTECTIONS NEEDED IF THE WALL COMES DOWN

The previous section describes the hazard from corporate combinations between banks and commercial firms, and argues that the pertinent hazards arise from (1) loss shifts to protect bank holding company reputation, that is, shifts meant to hide the loss; and (2) loss shifts that take advantage of limited liability, allowing shareholders to avoid the loss by imposing them instead on creditors or on the FDIC. In either case, economic efficiency can be diminished and the loss can end up with taxpayers. One means of addressing the hazards is to prohibit banking/commerce combinations; in other words, maintain the legislative status quo. Alternatively, legislators may decide that the benefits of combinations are worth bearing some danger of loss shifts. If legislators took this latter view, what types of protections could they employ to reduce the frequency of loss shifts into banks and thereby possibly to the FDIC or taxpayers? Already in place are the firewalls established by Sections 23A and 23B of the Federal Reserve Act. These statutory provisions limit transactions between banks and their affiliates. The firewalls are enforced by regular (once every year or year and a half) supervisory examination and by the threat of penalty if violations are discovered.

Beyond these current protections, which apply to any affiliations, including any commercial affiliations that might be allowed in the future, supervisors might wish to mimic the types of limitations uninsured creditors would impose on risky affiliations. As noted earlier, one can expect uninsured creditors to penalize the firm they fund (their debtor) and thereby potentially prevent an affiliation that would tend to lend itself to loss shifts. If supervisors are to mimic creditors' actions, they will restrict the types of firms that banks can affiliate with to those least likely to produce loss shifts in the first place. In other words, supervisors would only allow banks to affiliate with healthy commercial firms possessing strong capital at the time of affiliation.

While at the time of acquisition a new commercial affiliate may be strong, its health could deteriorate or it could take on undue risks. If uninsured creditors find that their debtor's affiliates are suffering losses or assuming risky endeavors, thereby increasing the chance of losses that might be shifted to their debtor, they would demand higher interest payments to compensate for their added risk. So creditors can be expected to monitor carefully the health of their debtor's affiliates. The Federal Reserve mimics private creditors by performing such monitoring (called *umbrella supervision* by the Fed) of holding companies owning a bank and a securities or insurance company, under provisions specified in the Gramm-Leach-Bliley Act of 1999. Umbrella oversight might well be desirable for combinations of banks with commercial firms, but could be more difficult than umbrella oversight of financial firms, for reasons discussed presently. If monitoring reveals that the bank's commercial affiliate has increased its risk, supervisors could impose a monetary cost on the bank through the current mechanism by which insurance premia are set.

Because umbrella oversight of commercial firms may be more difficult than oversight of financial firms, any legislation that might remove the wall could add additional protections beyond those found in the Gramm-Leach-Bliley Act. For example, such legislation could also limit the size of commercial affiliates to those no larger than a fraction of the size of the bank. Such a limit could be beneficial because as noted in an earlier section, shifts large enough to sink the bank are most likely to derive from commercial affiliates that are large relative to the size of the bank affiliate.

Firewalls

The 23A and 23B firewalls are intended to stop exactly those loss shifts that present hazards for bank/commercial firm affiliations. Yet there have been several cases in which shifts have caused bank failures, regardless of firewalls. Additionally, the firewalls have not been tested in a period of widespread affiliations involving nonbanks large enough to produce dangerous losses. Consequently, while in principle firewalls should prevent loss shifts, supervisors will probably wish to take further protective steps if the banking/commerce wall is removed.

Since 1933, banks have been protected against shifts of losses from affiliates by firewalls. Firewalls are found in Sections 23A and 23B of the Federal Reserve Act and apply to all banks.[14] They limit and place controls on transactions between banks and their affiliates.[15] For example, the 23A firewalls limit transactions, such as loans and asset purchases, between a bank and any individual affiliate to 10 percent of the bank's capital, and with all of its affiliates in total to 20 percent of the bank's capital. The firewalls operate only in one direction—they prevent transactions that might shift affiliate losses to the bank, but do not prevent transactions that might shift bank losses to affiliates. For instance, the firewalls prohibit loans to affiliates beyond 10 percent of bank capital, but not the reverse—loans to the bank by affiliates. They also require that purchases of affiliate assets by the bank be on terms at least as favorable to the bank as market terms. In contrast, the nonbank affiliate can purchase assets from the bank on terms unfavorable to the affiliate. Penalties for firewall violations can be quite severe, extending to significant monetary penalties imposed on banks and their managers and directors.

Bank failures caused by the shifts that firewalls were designed to prevent have been infrequent, but there were at least two, one of which was quite large. The 1955 Senate Report on the Bank Holding Company Act noted that "no widespread abuse of this nature [loss shifts] has been brought to the attention

[14] More specifically, Sections 23A and 23B apply to all insured banks and savings institutions. They are found at 12 U.S.C. 371c and 12 U.S.C. 371c-1, respectively.

[15] See Walter (1996) for a discussion of firewalls.

of [Congress]" (*U.S. Code: Congressional and Administrative News* [1956, 2486]). The House Report on the Act did discuss one case, that of the 1953 failure of First State Bank of Elmwood Park, Illinois, which resulted from shifts of bad loans from a nonbank loan company to its affiliate bank—apparently to take advantage of limited liability protections (U.S. House [1955, 18–19]).[16] Similarly, a 1983 study of the causes of bank failures for the previous ten years found only one case out of 120 failures caused by transactions between a bank and its nonbank affiliates. Still, this case, the failure of Chattanooga-based, $461 million Hamilton National Bank in 1976, was the third largest in U.S. history up to that time (Walter [1996, 23]).

Historically, then, firewalls have proven less than perfectly impervious. Further, until recently, affiliations were quite limited, offering few opportunities to put the firewalls to the test. Bank holding companies were, for the most part, restricted to owning nonbank financial firms that conducted activities that were similar to banking. Until the Gramm-Leach-Bliley Act was enacted in 1999, banking companies were prohibited from broad securities and insurance powers. Because of the limitations on the types of nonbank firms that bank holding companies could own, nonbanks have typically been much smaller than their bank affiliates. Since they were smaller, they were unlikely to be capable of producing losses large enough to sink affiliated banks.

Due Diligence prior to Affiliation

Since the effectiveness of firewalls is uncertain, care must be taken to ensure that bank holding companies do not acquire especially risky nonbanks. Currently, supervisors evaluate the nonbank's financial health as part of their review of applications from bank holding companies to acquire nonbanks.[17] They conduct analyses similar to *due diligence* analyses performed by investment companies for unregulated acquirers. Supervisors look for many of the same signals of problems that a creditor would, such as excessive debt and weak earnings performance. Similar analyses would be necessary for bank holding company acquisitions of commercial firms, should the wall come down.

Still, it might seem that such analysis of commercial firms would be expensive for bank supervisors, requiring the development of a very different

[16] See FDIC (1953, 7–8) for details of the First State Bank case beyond those provided in the House Report.

[17] The Gramm-Leach-Bliley Act of 1999 allows financial holding companies to dispense with applying for supervisors' approval of many nonbank acquisitions. Financial holding companies simply notify supervisors of the acquisition, within 30 days of the acquisition (Spong [2000, 157]). Therefore, acquisitions of nonbanks by financial holding companies often do not involve a pre-acquisition review of the nonbank's financial health. For acquisitions by bank holding companies that have not chosen, under rules specified by Gramm-Leach-Bliley, to become financial holding companies, such reviews still occur.

skill set. Bank supervisors who specialize in application review are experienced in examining the health of financial firms, not of commercial firms. Yet other supervisory employees—those who review banks' loans for their repayment prospects—are practiced in analyzing commercial firms, since most large bank loans go to such firms. Today's bank examiners also engage in industrywide analysis as part of their review of syndicated lending to large commercial firms. Therefore, these skills might be brought to bear fairly cheaply.

Under current procedures, if the supervisory review of the firm to be acquired turns up potential risks, the supervisor can deny the application or require that the risk be ameliorated. An application review of acquisitions of commercial firms would likely include the same options.

Beyond these procedures, supervisors might add another requirement, because analyses of commercial firms could be more difficult than analyses of financial firms. Shifts of losses large enough to sink the bank and take advantage of limited liability are most likely to occur when the commercial firm is large relative to its bank affiliate. Consequently, supervisors might also limit the size of commercial firms acquired to a fraction of the size of affiliated banks. Doing so would reduce the chance of bank-sinking or bank-endangering loss shifts. Such a requirement is not unprecedented, as relative size limits were imposed on bank holding company merchant-banking acquisitions under provisions of Gramm-Leach-Bliley.

Umbrella Supervision

While careful analysis of potential affiliates might prevent bank holding companies from purchasing troubled or initially risky commercial firms, problems at a commercial firm could arise well after its acquisition. Because of concern for this possibility, supervisors may wish to maintain ongoing oversight of the health of banks' commercial affiliates. The aim of such oversight is to determine whether the commercial affiliate has suffered losses or is expanding its riskiness. If supervisors find losses or heightened riskiness of commercial affiliates, they could indirectly impose a monetary penalty on the bank by lowering its supervisory rating. All banks are graded by supervisors on their financial health, riskiness, and management expertise. When a bank's grade (supervisory rating) declines, its insurance premiums can rise. Beyond this monetary penalty, when commercial affiliates suffer losses or increases in riskiness, supervisors might also watch more carefully for loss shifts (i.e., firewall violations). Further, they might even prohibit all transactions between the bank and its troubled commercial affiliate. In doing so, the supervisor mimics the monitoring that bank creditors would be expected to perform in the absence of deposit insurance.

Oversight of this sort currently occurs under provisions of the Gramm-Leach-Bliley Act, which make the Fed the *umbrella supervisor* of all financial holding companies. In this role, the Fed is to ensure that problems in a securities or insurance affiliate do not endanger the bank. For information on the health of securities and insurance affiliates, the Fed relies on financial reports from the Securities and Exchange Commission and state insurance commissioners. In some cases the Fed will participate in examinations of insurance companies performed by insurance commissioners. One main point of its umbrella oversight is to ensure that bank resources are not being shifted to nonbank affiliates. In the extreme, the Gramm-Leach-Bliley Act gives the Fed the authority to require that nonbank affiliates are divested. Divestiture provides the ultimate prohibition on loss-shifting transactions.

Umbrella oversight of commercial firms may be more difficult than oversight of securities and insurance firms. Securities and insurance firms already face strict regulation by agencies with long-standing experience as supervisors. Moreover, insurance companies receive regular examinations for financial health. Most commercial firms are less regulated and are not subject to examination by governmental supervisors. Developing such processes for commercial firms affiliated with banks could be quite expensive for an umbrella supervisor of combined bank/commercial firms and could impose large regulatory costs on the combined firms themselves.

Nevertheless, public firms—those whose securities trade in public markets—must release a great deal of financial information. Such information could provide much of the data necessary to judge financial health. While publicly available information may be less accurate and complete than that typically available to bank regulators, who have the power to require the release of any additional information they may deem useful, it is the set of information private investors rely upon when deciding whether to invest.[18] Therefore, for unregulated commercial firms, umbrella supervisors could rely upon publicly available information to a significant extent. Additionally, in passing the Sarbanes-Oxley Act, enacted in July 2002, legislators attempted to enhance the reliability of disclosures made by publicly traded companies.

4. CONCLUSION

Clearly, firms benefit from combinations of commercial and financial units, since for years they have chosen to mix them. Yet many experts have maintained that banking and commerce should remain separate. Two predominant reasons for maintaining the separation are concerns with conflicts of interest and the proliferation of monopoly. The most credible reason—indeed, one

[18] Of course, if private investors believe losses suffered by their bank-affiliated commercial firms can be shifted, then they have less reason to demand accurate accounting information.

that poses a significant hazard from combining banking with commerce—is that such affiliations could provide, at least under certain circumstances, incentives for loss shifts. While it turns out those circumstances are somewhat limited, they are not inconsequential.

Loss shifts can impose costs on taxpayers and waste resources. If losses are shifted from commercial firms to affiliated banks, taxpayer-funded bailouts may result. If creditors become convinced that firms affiliated with banks can shift losses to insured banks, then these firms will enjoy below-market borrowing costs. Below-market funding means that too many resources will flow to bank-affiliated firms. If so, productivity and financial market efficiency are diminished; in other words, scarce resources are wasted.

Nonetheless, since the Gramm-Leach-Bliley Act of 1999 allowed securities and insurance firms to affiliate with banks, potentially producing the same loss shifts that commercial affiliation might engender, why not also allow commercial affiliations? One reason that legislators might prefer not to open that door is that commercial firms are largely unregulated so the demands on supervisory resources are likely greater when protecting against shifts from largely unregulated commercial firms.

On the other hand, if legislators decide that the benefits of banking/commerce combinations could outweigh the hazards, what means of protection might they employ to minimize them? Several come to mind, including (1) careful analysis of the financial condition of commercial firms that bank holding companies wish to acquire, prior to acquisition; (2) the maintenance of firewalls to prevent loss shifts; and (3) umbrella supervision to provide the means of reducing the hazard. In addition to these means, the requirement that commercial firms be significantly smaller than any banks they affiliate with offers further protection. Size limits are likely to be valuable since a commercial firm is unlikely to produce a loss large enough to threaten a much larger bank affiliate.

APPENDIX: BACKGROUND ON THE STATUTES FORMING THE WALL

National Bank Act of 1864

The National Bank Act restricts the opportunities for national banks to undertake commercial activities. National banks are those chartered and regulated by the U.S. Treasury Department's Office of the Comptroller of the Currency. The Act states that "a national banking association shall. . . have power to. . . exercise. . . all such incidental powers as shall be necessary to carry

on the business of banking; by discounting and negotiating promissory notes, drafts, bills of exchange, and other evidences of debt; by receiving deposits; by buying and selling exchange, coin, and bullion; by loaning money on personal security" (12 U.S.C. 24). While courts and the Comptroller have, over the years, wrangled over the meaning of the *business of banking* clause, courts have generally taken a fairly conservative view of activities that might qualify. As decided in an influential court ruling, for example, banks are generally limited to conducting businesses that are functionally interchangeable with traditional banking services (*M&M Leasing Corp. v. Seattle First National Bank*, 563 F.2d 1377, 1383 [9th Cir. 1977] as cited by Halpert [1988, 487]). In sum, under the Act courts have allowed national banks to engage in businesses similar to banking but not other commercial activities. This restriction of powers extends not only to activities of banks, but to activities conducted by subsidiaries owned by banks (Halpert [1988, 486]).[19]

State Laws and FDICIA

For state-chartered banks, the banking/commerce wall is constructed of a mix of elements from state laws, the Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA), and the National Bank Act. State banking statutes typically set limits on the nonbank activities of state banks and their subsidiaries similar to the limits on national banks (Spong [2000, 37–41]). Over the years a number of states have authorized activities beyond those allowed national banks. Yet state banks' opportunity to expand further than the activities allowed under the National Bank Act was largely ruled out by the FDICIA. Specifically, Section 24 of the Federal Deposit Insurance Act as amended by the FDICIA prohibits insured state-chartered banks from engaging in any activities impermissible for national banks unless the FDIC rules that such activities pose no threat to the deposit insurance fund (sec. 303 of Public Law 102-242).

Bank Holding Company Act

Enacted in May 1956, the Bank Holding Company Act was based on the view that "bank holding companies ought to confine their activities to the management and control of banks." Legislators appear to have been motivated by two concerns. First, that conflicts of interest might arise if one company

[19] Halpert (1988, 497) argues that it was of minor significance to Congress whether banks engaged in nonbank activities when writing this language of the National Banking Act. He maintains that Congress "never affirmatively required banks to stay out of nonbanking business," but "[r]ather, subsequent interpretations of the statute by comptrollers of the currency and various courts provided its restrictive cast."

owned both a bank and a commercial firm. For example, such a conflict arises when a bank receives a request for a loan from one of its commercial affiliate's competitors. Second, though the legislative history is less clear on this point, legislators appear to have also been worried that combinations might lead to the growth of monopoly power.

To address these concerns, the Act restricted bank holding companies such that they "would no longer be authorized to manage and control nonbanking assets unrelated to the banking business" (*U.S. Code: Congressional and Administrative News* [1956, 2484, 2492]). At the time the Act was passed, banking companies were growing rapidly through mergers. In a few cases these companies included nonbanking businesses. The widest-ranging example was found in Transamerica Corporation. It combined in one firm, banking, insurance, and a relatively small amount (as a percentage of Transamerica's total assets) of metals manufacturing and fish processing (Halpert [1988, 498]).

The Act required that companies wishing to purchase a bank first seek approval from the Board of Governors of the Federal Reserve System. Further, the Act prohibited the Board from approving purchases by companies engaged in activities that were not *closely related to banking*, thereby prohibiting commercial companies such as manufacturers from purchasing banks. Commercial firms like Transamerica that owned banks were given several years in which to divest either the bank or alternatively their commercial activities. Through the next forty years the Board developed a list of activities that would be considered closely related, excluding activities most observers would consider commercial.

In 1999, the Gramm-Leach-Bliley Act was enacted. It added securities underwriting and dealing as well as insurance to the list of activities in which banks—through bank-owned subsidiaries—and bank holding companies could engage. Until that time, banks and their subsidiaries and bank holding companies had been prohibited from the securities business by the 1933 Glass-Steagall Act.[20] Insurance activities were likewise highly restricted before Gramm-Leach-Bliley by the Bank Holding Company Act and other laws.

The Glass-Steagall Act's separation of securities activities from banking was driven by legislators' concerns over conflicts of interest, excessive stock market speculation by bank-owned securities firms, and threats to the health of banks from securities activities. Likewise the Bank Holding Company Act's separation of banking and insurance was part of that law's general separation of banking from nonbank activities, driven by concerns over conflicts of interest and monopoly power. By the time the Gramm-Leach-Bliley Act was passed, legislators and other observers had various reasons for removing the walls that separated banks from securities and insurance activities. These reasons

[20] Bank holding companies began to engage in limited securities activities starting in 1987 through a loophole in Glass-Steagall.

fall into three categories. First, there is little evidence of conflicts of interest or other problems when banks were combined with nonbank firms. Second, market developments, such as growing competition in banking markets, had rendered these problems less important by the 1990s. Third, the concerns could be dealt with effectively by regulating the combined firms.

Ultimately, the Gramm-Leach-Bliley Act specified that only healthy, fairly low-risk bank holding companies were to be allowed to undertake this broader array of financial activities. Those that do so are called *financial holding companies*. Further, the Act allows these new financial holding companies to engage in merchant banking, whereby under certain conditions financial holding companies may purchase the equity of (in other words, become owners of) any type of corporation, commercial or otherwise. Financial holding companies' merchant banking subsidiaries are restricted to holding the equity of firms for a limited period of time and are prohibited from active management of the firms. Beyond securities and insurance, Gramm-Leach-Bliley allows the Board of Governors, in conjunction with the Treasury Department, to also authorize financial holding companies to undertake additional activities that are "financial in nature" or "incidental to financial activities." It also authorizes the Board to approve activities that are "complementary to a financial activity." So the Gramm-Leach-Bliley Act expands the activities of bank-owning companies beyond those previously allowed by the Bank Holding Company Act to include most financial activities, but leaves in place the wall between banking and commerce.

Loopholes in the Bank Holding Company Act Section of the Wall

Loopholes have been employed to allow banking/commerce combinations, at least to a limited extent. The *unitary thrift loophole*, closed by the Gramm-Leach-Bliley Act in 1999, was one such opening. Through it, companies owning only one thrift (thus the phrase *unitary thrift*) could also own commercial firms. The loophole existed because thrift institutions (meaning primarily savings and loans, and savings banks) are not covered by the Bank Holding Company Act, which prevents banking/commerce ties. Instead, thrifts are regulated under the Savings and Loan Holding Company Amendments of 1967, which allow commercial activities in unitary thrift holding companies (Seidman [1998, 7]). Gramm-Leach-Bliley closed the loophole though it grandfathered existing unitary thrift holding companies, allowing them to continue to engage in commercial activities.

An additional loophole was partially closed in 1987, but remains open to a limited degree. Before 1987, the Bank Holding Company Act defined a bank as a firm that both offered demand deposits (a type of checking account) and made commercial loans. This definition prevented commercial firms from owning a typical bank, which offers both demand deposits and commercial

loans. Nevertheless, commercial firms could form a bank that did not offer one or the other. By doing so, commercial firms could own banks that did not fall within the Bank Holding Company Act definition of a bank and could circumvent the Act's prohibition of mixing banking and commerce. These banks, known as *nonbank banks*, did not fit the Act's definition of a bank but did offer most banking services. A number of firms established nonbank banks, both as a means of combining banking and commerce and as a means of banking across state lines, which was difficult until the 1990s. In 1987, Congress closed the loophole by tightening the definition, but allowed states with existing laws authorizing the chartering of *industrial loan corporations* (a type of nonbank bank that funds itself with insured deposits but does not offer demand deposits) to continue to charter these ILCs. Several states had such laws as of 1987. This option remains in force as a means of combining banking and commerce in these states.

REFERENCES

Bank Holding Company Act. 1956. *U.S. Code*. Title 12, sec. 1843c.

Board of Governors of the Federal Reserve System. 1974. "Order Denying Investment in Allstate International S.A., Zurich, Switzerland." *Federal Reserve Bulletin* 60 (July): 517–19.

_____. 1981. "Order Approving Proposed Bookkeeping and Data Processing Activities and Denying Proposed Finance, Loan Servicing, Leasing and Insurance Activities." *Federal Reserve Bulletin* 67 (May): 449–51.

_____.1987. "Order Approving Applications to Engage in Limited Underwriting and Dealing in Certain Securities." *Federal Reserve Bulletin* 73 (June): 473–508.

Federal Deposit Insurance Corporation (FDIC). 1953. *Annual Report*. Washington, D.C.: FDIC.

Federal Deposit Insurance Corporation Improvement Act (FDICIA). 1991. U.S. Public Law 102-242, sec. 303. 102nd Cong., 1st sess.

Federal Reserve Act. 1913. *U.S. Code*. Title 12, secs. 371c, 371c-1.

Halpert, Stephen K. 1988. "The Separation of Banking and Commerce Reconsidered." *Journal of Corporation Law* 13 (Winter): 481–533.

Krainer, John. 2000. "The Separation of Banking and Commerce." Federal Reserve Bank of San Francisco *Economic Review*: 15–25.

Macey, Jonathan R., and Geoffrey P. Miller. 1992. *Banking Law and Regulation*. Boston: Little, Brown and Company.

Morgan, Donald P. 2000. "Rating Banks: Risk and Uncertainty in an Opaque Industry." Federal Reserve Bank of New York Staff Report 105 (April).

National Bank Act. 1864. *U.S. Code*. Title 12, sec. 24.

Owens, Raymond E. 1994. "Commercial Real Estate Overbuilding in the 1980s: Beyond the Hog Cycle." Federal Reserve Bank of Richmond Working Paper 94-6 (May).

Seidman, Ellen. 1998. Statement on Financial Modernization before the Committee on Banking, Housing, and Urban Affairs. U.S. Senate. 105th Cong., 2nd sess. 25 June.

Spong, Kenneth. 2000. *Banking Regulation: Its Purposes, Implementation, and Effects*, 5th ed. Kansas City, Mo.: Federal Reserve Bank.

U.S. Code: Congressional and Administrative News. 1956. St. Paul, Minn.: West Publishing Company.

U.S. General Accounting Office. 1987. "Bank Powers: Insulating Banks from the Potential Risks of Expanded Activities." Washington, D.C.: U.S. GAO. April.

_____. 1997. "Bank Oversight: Few Cases of Tying Have Been Detected." Washington, D.C.: U.S. GAO. May.

U.S. House. 1955. *Report to Accompany H.R. 6227*. 84th Cong., 1st sess. H.R. 609.

Walter, John R. "Firewalls." 1996. Federal Reserve Bank of Richmond *Economic Quarterly* 82 (Fall): 15–39.

_____. 1998. "Can a Safety Net Subsidy Be Contained?" Federal Reserve Bank of Richmond *Economic Quarterly* 84 (Winter): 1–20.

Walter, John R., and John A. Weinberg. 2002. "How Large Is the Federal Financial Safety Net?" *The Cato Journal* 21 (Winter): 369–93.

Weinberg, John A. 1996. "Tie-in Sales and Banks." Federal Reserve Bank of Richmond *Economic Quarterly* 82 (Spring): 1–19.

Unemployment Insurance and Personal Bankruptcy

Kartik Athreya

Personal bankruptcy allows households to stop or delay the repayment of debts. In so doing, bankruptcy provides a form of insurance to households. In particular, bankruptcy allows households some flexibility in timing repayments in a way that allows for sudden unforeseen contingencies. As an implicit form of insurance, bankruptcy may augment, substitute for, or even limit other forms of insurance. Conversely, the presence of other forms of insurance against life's vicissitudes may enhance or limit the usefulness of bankruptcy. In this article, I investigate the interaction between one of the largest social insurance schemes, the U.S. unemployment insurance system (UI), and the personal bankruptcy system.

An overwhelmingly large proportion of those filing for bankruptcy (over two-thirds) have recently experienced a job disruption (Sullivan et al. [2000] and Domowitz and Sartain [1999]). Further, Cochrane (1991) finds that prolonged spells of unemployment are poorly insured and therefore result in large drops in consumption levels. How does the level of unemployment insurance available to workers affect the benefits of bankruptcy protection? Conversely, how do the benefits of bankruptcy alter the benefits generated by UI? Lastly, how does the presence of bankruptcy alter the consequences of scaling back unemployment insurance?

My findings are as follows: First, in the benchmark economy, introducing bankruptcy under even low UI replacement ratios lowers welfare. Second, reducing the UI replacement ratio increases bankruptcy rates.[1] Additionally,

The author would like to thank Tom Humphrey, Ned Prescott, John Weinberg, Roy Webb, and seminar participants at Colgate University and Hamilton College. The views expressed herein are those of the author and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] One reason why this should not be seen as obvious is that penalties for bankruptcy involve ejection from credit markets. Therefore, lowering the replacement ratio hurts bankruptcy filers more than before, which may imply fewer, not greater, annual filings.

reducing the UI replacement ratio worsens consumption smoothing less when bankruptcy is allowed than when it is not. However, while welfare falls slightly with the replacement ratio, the fall is nearly independent of bankruptcy law. Third, bankruptcy lowers asset trade, which in turn implies a more equal long-run distribution of wealth (as fewer households hold either very low or very high asset levels to deal with income shocks). However, asset trading behavior is not affected greatly by changes in the UI replacement ratio. Fourth, UI is more important than bankruptcy: if society must choose either UI or bankruptcy, it should choose UI. Last, bankruptcy's role in providing insurance is clearly dependent on the existing social safety net. In summary, unemployment insurance appears to materially affect the desirability of bankruptcy protection, but allowing bankruptcy does not, in the benchmark economy, alter the consequences of scaling back UI.

The environment here is an extension of the environment studied in Athreya (2002b), augmented to include unemployment. Athreya (2002b) examines the welfare implications of recent "means-testing" proposals. The present work is perhaps closest to the work of Livshits et al. (2002) and Fisher (2002). The work of Fisher (2002) is the first empirical study of the effects of public insurance on the personal bankruptcy decision. With respect to bankruptcy, this work is also related to recent research of Chatterjee et al. (2001) and Li and Sarte (2002). With respect to positive analyses of unemployment insurance and its consequences, the work is related to, but simpler than, the models of Hansen and Imrohoroglu (1992) and Alvarez and Veracierto (2001). The two preceding articles study unemployment insurance in general and the effect of severance payments on job security, respectively.

The sudden fall in earnings associated with a layoff or firing or an inability to continue working due to illness has long been cited by bankruptcy scholars as an important correlate of bankruptcy.[2] Thus, it stands to reason that the treatment received by those who become separated from their employers will influence their decision whether or not to file for bankruptcy. I turn now to a simple dynamic general equilibrium model of consumption and savings in the presence of some uninsurable income risks, including the risk of losing one's job. To simplify matters, I abstract from production decisions as well as the impact of moral hazard in increasing the costs of administering an unemployment insurance system. In ongoing research (Athreya [2002a]), I pursue a more complete analysis to incorporate moral hazard and production.

1. THE MODEL

Bankruptcy allows a borrower to essentially design a state-contingent repayment plan, whereby repayment is made only when outcomes for the borrower

[2] Of course, this is no more causal than is having too little income or too much debt, given one's income stream.

are relatively good. In this sense, the amount of a household's income dedicated to loan repayment can be varied, allowing it to apply limited income in a difficult period towards consumption rather than debt service. Unemployment insurance and antipoverty programs, conversely, act directly on the income of the household and help it remain above a threshold. Both of these programs can help households insure themselves within a period against uncertain job or health prospects. However, both programs must be paid for.

Allowing bankruptcy implies paying more for loans, as households are also purchasing the right to suspend or completely avoid repayment, subject to penalties. The high rate on loans also means that as households attempt to avoid borrowing, each saves so much that the return to savings may fall relative to an economy without bankruptcy. In turn, this fall mutes the effectiveness of savings to carry consumption across periods. Unemployment insurance, for its part, must be paid for via (possibly distortionary) taxes. Furthermore, as is well known, UI may introduce inefficiency, as both the effort expended by currently employed households and the job search efforts of currently unemployed households may fall. Moreover, a major penalty for filing for bankruptcy is exclusion from credit markets. In contrast, while UI may directly lower the need for borrowing and subsequent bankruptcy, generous UI makes exclusion from credit markets less painful. Thus, while bankruptcy and UI act in different ways, the presence of each is likely to affect the other.

Preferences and Endowments

Individuals maximize the present value of expected lifetime utility, given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\alpha} - 1}{1 - \alpha}, \tag{1}$$

where E_0 is the expectations operator, conditional on time 0 information, $\beta \in (0, 1)$ is the discount factor, c is consumption, and α is the measure of both risk aversion and the desire for intertemporal consumption smoothing. A full description of the household's optimization problem will be given after more notation is introduced.

Consumers in this market, intended to represent U.S. households, are assumed to be risk-averse price takers. They face uncertain labor incomes and other uninsurable idiosyncratic risks. The economy is composed of many long-lived households. At the beginning of each period, all households receive a random level of labor income that depends on their employment status. Households in the economy retain employment in each period with probability ρ and are subject to the risk of losing employment in a given period with probability $(1 - \rho)$. Once employment is lost, regaining employment occurs with probability ξ. An unemployed worker receives unemployment insurance

in only the first period of unemployment, that is, when the worker is newly unemployed. In subsequent periods, there is a subsistence level of income given to households. The endowment structure for unemployed households is meant to reflect the current practice of the use of a flat "replacement ratio" and the limited length of UI benefits in the United States. Newly unemployed households receive $\theta\overline{Y}$, where $\overline{Y}$ is mean labor income and $\theta \in [\underline{\theta}, 1]$ the replacement ratio. After the first period of unemployment, households, if unemployed, will receive the subsistence transfer of $Y_{\min} > 0$.[3]

Given the exogenously imposed flow of households out of unemployment, the replacement ratio for UI benefits θ, and the long-run average employment rate μ_e, it is easily shown that the per-period lump-sum tax η_u necessary to finance the UI system is given by $(1-\rho)\mu_e\theta\overline{Y}$.[4]

The endowments of employed households are random and cross-sectionally independent but are serially dependent. Agents are identical ex ante in terms of expected income, assets, and consumption. When employed, the after-tax endowment of a household in period t can take two values, $\widetilde{Y} = y_l$ and $\widetilde{Y} = y_h$, where the subscripts h and l denote high and low labor income, respectively, such that $y_l < y_h$.[5] Defining unemployment as a separate state for the endowment process is what allows for an analysis of how UI benefits interact with bankruptcy law.

There is a transition function over the income of employed households whereby $P(y' = y_l | y = y_l) = p_{ll}$ and $P(y' = y_h | y = y_h) = p_{hh}$. That is, p_{ll} is the probability that the labor income shock remains low in the next period, given that it is low in the current period. Similarly, p_{hh} is the probability that the labor income shock remains high in the next period, given that it is high in the current period. The assumption of serially dependent income introduces anticipation effects in asset holdings and default behavior and determines the effectiveness of using assets to smooth consumption. The parameters of the income process will be chosen to be broadly consistent with post-transfer income variability in U.S. data.

---

[3] To focus on the interaction of bankruptcy and explicitly financed unemployment insurance, I avoid tracking the collection of taxes with which subsistence income payments are made. When analyzing changes in bankruptcy law, $Y_{\min}$ will remain fixed, so there is no harm in treating it as an endowment.

To keep matters simple, I did not specify UI replacement to depend on previous income. Doing so would entail tracking households flowing into employment separately, which increases the cost of computing solutions. Moreover, as job loss is exogenous with respect to income, the average household flowing into unemployment will have $\overline{Y}$ as the previous period's labor income.

[4] This is a simple example of a "bathtub" model of unemployment, whereby the exogenous flows into and out of employment are set such that there is a constant level of employed (and unemployed) households in the economy. See Ljungqvist and Sargent (2000). The flow into unemployment is given by $(1-\rho)\mu_e$, and the cost of insurance payments to each household is $\theta\overline{Y}$. Therefore, per capita taxes, η_u, must satisfy $\eta_u = (1-\rho)\mu_e\theta\overline{Y}$.

[5] I define endowments as "after-tax" income for simplicity of notation and exposition.

Assets

Agents may save using risk-free private bonds or risk-free government debt and may borrow on an unsecured credit market. Government debt is incorporated both for descriptive accuracy as well as to avoid artificially constraining households to the use of private borrowing and lending alone.[6] Household borrowing is subject to a liquidity constraint, and households may default on previously acquired debt. The stock of private risk-free debt is issued by diversified competitive financial intermediaries in order to finance loans to households. The market for privately issued unsecured credit in the United States is characterized by a large, competitive marketplace where price-taking lenders issue credit through the purchase of securities backed by repayments from borrowers. These transactions are intermediated principally by credit card issuers. As the typical credit card contract is described by a fixed interest rate and credit line, the interest rates charged by credit card issuers may be viewed as being set to cover the aggregate default rate rather than being individually tailored for each account. Further, interest rates do not appear to vary systematically with individual debt levels, even though the marginal likelihood of default may change.[7]

There will be two prices quoted for assets: a loan rate, r^l, for those who borrow and a deposit rate, r^d, for those who save. These two rates are different, because with the bankruptcy option, a certain fraction of households will default in equilibrium, and in order to break even, financial intermediaries will have to charge higher interest rates on loans than they pay for deposits. The stock of government debt is denoted D and is financed by a lump-sum tax $\eta_D = r^d D$ on all households, where r^d is the interest rate on risk-free savings deposits.

The assumption that all debt is unsecured is less restrictive than it may seem. The model best represents the section of U.S. households with little or no collateral and higher than average labor income risk that rely on unsecured debt to smooth consumption. Therefore, the welfare implications developed here apply directly to a population most affected by bankruptcy reform. Additionally, Gropp, Scholz, and White (1997) argue that in many cases those considering filing use unsecured credit to pay off secured debts and then discharge this debt in bankruptcy, thereby making the distinction between these types of debt less clear in practice.

---

[6] The stock of government debt per capita will, however, be held fixed throughout all the policy experiments I conduct.

[7] On the existence of competitive equilibrium in a model where interest rates on loans cover average repayment rates, see Dubey et al. (2000).

Bankruptcy

Bankruptcy in the model will most closely resemble Chapter 7 "total liquidation" bankruptcy. If a household files for bankruptcy, its income and assets become known to the credit market, and if it qualifies, its unsecured debt is discharged but is then constrained for an uncertain period of time from borrowing. Households may, however, save during this time. The principal motivation for a random period of restricted credit access is that it reduces significantly the computational burden of solving the household's optimization problem. Specifically, the assumption allows one not to distinguish between households on the basis of the length of their credit market exile.[8] In each period following a bankruptcy, a borrowing-constrained household remains constrained with probability $(1 - \psi)$. Therefore, the average time that a household is constrained from borrowing and prohibited from filing again is given by $1/(1 - \psi)$.

The Cost of Bankruptcy and Deadweight Loss

Bankruptcy involves three types of costs. First, as was just discussed, it results in at least some exclusion from credit markets. Second, there are explicit time costs arising from court dates and other legal proceedings. Finally, societal disapproval or "stigma" may play a role (see Dubey et al. [2000]; Fay et al. [1996]; and Gross and Souleles [2000]).

An important drawback of using bankruptcy to provide insurance is that the penalties listed above typically do not involve a transfer of wealth from debtors to anyone, let alone creditors. I denote by λ all costs of bankruptcy beyond credit market exclusion. That is, λ represents the "deadweight" costs of bankruptcy. I will set λ to match observed bankruptcy filing rates among homeowners, given the current average length of credit market exclusion.

The Household's Problem

At any point in time, households belong to one of two mutually exclusive classes of credit market status and three mutually exclusive classes with respect to employment status. For credit status, households are either solvent or constrained from borrowing. Solvent households are those that have full access to credit markets and have the option of filing for bankruptcy. Borrowing-constrained households are those that have filed for bankruptcy in the past but have not yet been readmitted to credit markets.[9] With respect to employment

[8] For more on stochastic punishment spells in bankruptcy, see Athreya (2002b) and Chatterjee et al. (2001).

[9] Therefore, while the move from solvent to borrowing-constrained status is a choice for households, the release from borrowing-constrained status is exogenous.

status, households are either employed, newly unemployed, or unemployed for more than one period.

In each period, given their current income and beginning-of-period assets, households must choose consumption, c, and asset holdings to carry forward into the next period, denoted a'. From the individual's point of view, all saving is risk-free and earns the same rate of return. Therefore, the household makes no distinction between government debt and private bonds when choosing how to allocate its savings. Depending on whether it chooses to be a net borrower or lender, it faces either the net rate of interest on loans, r^l, or deposits, r^d, where $r^l > r^d$.

I restrict borrowing according to a household's credit status as follows. For solvent households, assets a' must be greater than $\underline{a}^S$, a negative number indicating that solvent households may borrow. Households that have filed for bankruptcy face a more severe restriction than solvent households in their ability to borrow. Their borrowing limit, denoted $\underline{a}^B$, therefore is given by $a' \geq \underline{a}^B$, where $\underline{a}^B > \underline{a}^S$. Similarly, households that are constrained from borrowing are also restricted in their borrowing, with a limit denoted $\underline{a}^{BC}$, whereby $a' \geq \underline{a}^{BC}$, where $\underline{a}^{BC} > \underline{a}^S$.

When a household is solvent, it must first choose whether or not to file. It then chooses assets subject to the constraints for solvent or borrowing-constrained households, depending on its employment status and default decision. The current period state vector, conditional on credit status, is denoted (e, a, y), indicating employment status, asset holdings, and current income, respectively.

Current labor income is denoted $y(e)$, where e denotes beginning-of-period labor market status. A worker's employment status belongs to one of three categories, that is, $e \in \{e_0, e_1, e_2\}$, where e_0 denotes an employed worker, e_1 a newly unemployed worker, and e_2 a worker who has been unemployed for more than one period. The law of motion for labor income is simple. In any period, an employed worker may lose his job with probability $(1 - \rho)$. He is then classified as "newly unemployed" and is eligible for UI benefits. In the following period, he finds employment with probability ξ, in which case he receives a (random) endowment of $y(e_0)$.[10] If he fails to find employment in this period, he is classified as "unemployed" and is therefore no longer eligible for UI benefits and receives labor income $Y_{\min} > 0$.

I denote the value of being solvent by V^S, the value of not filing for bankruptcy as W^S, the value of filing for bankruptcy as W^B, and the value of being borrowing-constrained as V^{BC}. The value of solvency is given as follows:

[10] This income is drawn from the conditional probability distribution of income, as if the household had received income shocks while unemployed. This simplifies the analysis by avoiding the use of a separate income process once released from unemployment.

$$V^S(e, a, y) = \max[W^S(e, a, y), W^B(e, a, y)], \tag{2}$$

where

$$W^S(e, a, y) = \max\{u(c) + \beta E V^S(e', a', y')\} \tag{3}$$

s.t.

$$c + \frac{a'}{1 + r^{d,l}} \leq y(e) + a \tag{4}$$

s.t.

$$a' \geq \underline{a}^S. \tag{5}$$

When the household chooses to file for bankruptcy, it has its debt removed, pays the nonpecuniary cost, λ, and then is automatically sent to the borrowing-constrained state, where it obtains value V^{BC}. Therefore, the value of filing for bankruptcy, W^B, satisfies

$$W^B(e, a, y) = \max\{u(c) - \lambda + \beta E V^{BC}(e', a', y')\} \tag{6}$$

s.t.

$$c + \frac{a'}{1 + r^d} \leq y(e) \tag{7}$$

s.t.

$$a' \geq \underline{a}^B. \tag{8}$$

To define V^{BC} above, note that households in the borrowing-constrained state face a lottery, whereby with probability ψ, they are returned to solvency (i.e., they are free to borrow and default in the following period), and with probability $(1 - \psi)$, they are still restricted from borrowing or defaulting. Thus, we have

$$V^{BC}(e, a, y) = \max\{u(c) + \psi\beta E V^S(e', a', y') + (1 - \psi)\beta E V^{BC}(e', a', y')\} \tag{9}$$

s.t.

$$c + \frac{a'}{1 + r^d} \leq y(e) + a \tag{10}$$

s.t.

$$a' \geq \underline{a}^{BC}. \tag{11}$$

I turn now to the definition of equilibrium in the model.

Equilibrium

The consumer choice problem above captures the decisions of a very large number of households. However, given the absence of perfect income insurance, households that have received many bad income shocks are likely to find themselves in debt, while those that have been lucky may have large levels of savings. Their choices are governed by a *decision rule*, which, for a household of type i, specifies asset holdings as a function of interest rates, employment status, income, current assets, and borrowing constraints.

An equilibrium consists of a decision rule for each type of agent and interest rates r^l and r^d such that four requirements are met. First, given these interest rates, decision rules solve the optimization problem described above for each type of household. Second, total economy-wide borrowing by households equals total economy-wide saving. Third, the spread between loan and deposit rates is such that financial intermediaries exactly cover their costs, given the observed bankruptcy rate. Fourth, the payments to newly unemployed households each period must be covered by tax revenues (i.e., the government runs a balanced budget while maintaining the stock of debt D). In addition, I restrict attention to steady state equilibria where the bankruptcy rate and the proportion of agents in the population with a given level of assets are stationary, that is, the same at every date.

Welfare Measurement

The welfare criterion used here measures the percentage change in consumption, in all states and at all dates, that would make a household indifferent between living in an economy in which a given policy experiment prevailed and one in which the benchmark setting prevailed. Let this increment/decrement to consumption be denoted by ϕ. A negative value for ϕ implies that households are worse off, and a positive value implies the reverse. Multiplying ϕ by mean household income then converts ϕ into a dollar measure of annual welfare gains or losses per household.[11]

[11] With the utility function used here, the welfare measure is given as follows. The desirability of outcomes will be evaluated according to the following expression:

$$\Lambda = \int_X V(x) d\mu, \tag{12}$$

where $V(x)$ is the maximal attainable utility from being in a given state x and μ is the long-run stationary distribution (C.D.F.) of households across states. Therefore, Λ is the expected value function of households over assets, income, *and credit status*. This is a utilitarian social welfare function that weights all households equally. It measures ex ante welfare. I use this measure to estimate the increment/decrement to consumption under a given bankruptcy policy, at all dates and states, that makes households indifferent between the economy defined by the proposed bankruptcy policy and the benchmark economy. I denote this increment/decrement ϕ. Let Λ^{bench} denote benchmark welfare, and Λ^{policy} denote welfare under a proposed policy. Given the preferences used here, ϕ will satisfy the following:

Beyond this measure of welfare, I will also examine the behavior of some other statistics in assessing the interaction between bankruptcy and UI. Given that changes in the replacement ratio alter the mean level of after-tax income for households in the model, it is useful to have a measure of consumption volatility that does not depend on average income, such as the *coefficient of variation* (denoted c.v.). The c.v. will also be useful when exploring the role of bankruptcy in altering the level of asset accumulation and decumulation.

To measure inequality, I use a traditional tool, the Gini Coefficient. Roughly speaking, this coefficient measures the departure of a given distribution of wealth, consumption, or income from a perfectly equal distribution. A Gini of one indicates, for example, that the very richest household holds the entirety of wealth, while a coefficient of zero indicates that all households hold exactly equal levels of wealth. A more disaggregated measure of inequality is the distribution of income, consumption, and wealth by various percentiles, which I also report below.

Parameterization

The model parameters are set to match observed bankruptcy rates under plausible levels of income shock persistence and volatility and are summarized in Table 1. For brevity, rather than including a full discussion here, I refer the interested reader to the details in Athreya (2002a, b).

With respect to unemployment insurance, I follow Hopenhayn and Nicolini (1997), who use the estimates of Meyer (1990). In particular, Meyer (1990) finds that the average length of insured unemployment is thirteen weeks, with a replacement rate of 66 percent and a 10 percent chance of reemployment at the end of the spell. I therefore set the model period at thirteen weeks, set $\theta = 0.66$, and set $\xi = 0.10$. The credit limit is set by noting that median unsecured debt among bankrupt households in recent years has fluctuated between one-fourth and one-half of annual median income (see Sullivan et al. [2000, 65–66, 122]). Credit card debt, to which the debt in the model corresponds most closely, was approximately $9,500 in 1997, equal to U.S. median quarterly income (Sullivan et al. [2000]). Given the period of thirteen weeks, or one quarter, I therefore set $\underline{a}^s = \overline{Y}$. For simplicity, I set $\underline{a}^B$ and $\underline{a}^{BC}$ to zero.

An important parameter in the model with respect to bankruptcy is the one governing credit market exclusion, ψ. While ψ is not easily observable, lenders in the unsecured credit market still allow agents access to loan markets

$$\phi = \left(\frac{\Lambda^{policy} + \frac{1}{(1-\alpha)(1-\beta)}}{\Lambda^{bench} + \frac{1}{(1-\alpha)(1-\beta)}} \right)^{\frac{1}{1-\alpha}} - 1. \tag{13}$$

Under this criterion, $\phi > 0$ implies that households are better off under a proposed policy than in the benchmark case, and $\phi < 0$ implies the reverse.

Table 1 Parameters

Parameter	Value	Source
β (annual)	0.914	Calibrated
α	2.00	Aiyagari (1994)
p^i_{hh}	0.74	Heaton and Lucas (1997)
p^i_{ll}	0.74	Heaton and Lucas (1997)
ρ	0.006	Calibrated
μ_e	0.943	Alvarez and Veracierto (2001)
ξ	0.10	Meyer (1990)
θ (benchmark)	0.66	Meyer (1990)
y_h	1.25	Heaton and Lucas (1997)
y_l	0.75	Heaton and Lucas (1997)
$Y_{\min}$	0.40, 0.10	
λ	2.80	Calibrated
$\underline{a}^s$	$-\overline{Y}$	Huggett (1993); Sullivan et al. (2000)
$\underline{a}^B, \underline{a}^{BC}$	0	

following default or bankruptcy within a year or two. I set $\psi = 0.25$ such that the average period of exile from credit markets is four model periods, or one year.[12] The level of income received by unemployed households after unemployment benefits are exhausted, $Y_{\min}$, is set in the benchmark case to 0.40, to provide 40 percent of median household income, as a proxy for the various income support and transfer programs available to U.S. households. This level amounts to $1,332 per household per month.[13] Subsequently, $Y_{\min}$ will be set to a much lower 0.10, or $333 per household per month, to examine the role played by social insurance beyond unemployment compensation.

The parameter λ, which is the cost of bankruptcy in excess of credit market restrictions, will be inferred by the level that it must take in order to match observed bankruptcy filing rates. In terms of bankruptcy rates, total nonbusiness bankruptcy filings have been stable at roughly 1.3 million annually. Of these, roughly 70 percent are Chapter 7, "total liquidation" bankruptcies, implying an annual incidence of 0.9 percent.

2. RESULTS

To begin, I define the benchmark case, against which policy experiments will be compared.

[12] In this model, exclusion from borrowing hurts the households without helping anyone else. It is therefore a deadweight penalty and could have been left unmodeled by combining it with the general nonpecuniary penalty, λ.

[13] The transfers received by households beyond UI come from the major public assistance programs in the United States: Supplemental Security Income (SSI), General Assistance, Medicaid, and Temporary Assistance to Needy Families (TANF).

Table 2 Welfare Effects of Introducing Bankruptcy

θ	r^l	r^d	Bankruptcy Rate	Welfare Change ($)	Utility
0.66	4.39%	4.39%	—	—	−37.43
0.66	13.00%	2.57%	0.90%	−$66.88	−37.57
0.50	4.21%	4.21%	—	—	−37.49
0.50	13.00%	2.35%	0.99%	−$70.46	−37.63
0.40	4.06%	4.06%	—	—	−37.52
0.40	12.55%	2.31%	1.03%	−$62.70	−37.65

Definition 1 *Throughout the analysis, the "benchmark" economy is defined specifically to be the case where bankruptcy is allowed, the replacement ratio, θ, is set at* 0.66*, and $Y_{\min}$ = 0.40.*

I first study the consequences, when $Y_{\min} = 0.40$, of introducing bankruptcy into a setting where unemployment is already partially insured. The clear conclusion in this case is that bankruptcy protection is harmful, as seen in Table 2. Introducing bankruptcy is damaging even when the unemployment insurance system is very strict. The quarterly cost to the household of introducing bankruptcy ranges from $66.88, when $\theta = 0.66$, to $70.46, when $\theta = 0.50$, to $62.70, when $\theta = 0.40$. With respect to prices, I find that when bankruptcy is introduced, the interest rate on savings falls, while the rate on borrowing rises. For example, when $\theta = 0.66$, r^d falls from 4.39 percent to 2.57 percent, while r^l rises sharply from 4.39 percent to 13.00 percent. Such changes in interest rates are associated with worsened consumption smoothing, as the return to savings is low, while borrowing becomes very expensive. On the other hand, the option of bankruptcy allows households new consumption smoothing possibilities. On net, however, welfare appears to suffer. The welfare measure reported in Table 2 captures the change in welfare generated by the introduction of bankruptcy, holding the replacement ratio fixed. This result is summarized in Result 1.

Result 1 *In the benchmark economy, introducing bankruptcy under even low UI replacement ratios lowers welfare, increases interest rates on loans, and reduces interest rates on savings.*

Perhaps unsurprisingly, lower replacement ratios produce systematically lower utility levels. For example, when bankruptcy is not allowed, the expected utility of households falls from −37.43 to −37.49 to −37.52 as θ drops from 0.66 to 0.50 to 0.40 (see Table 2). The intuition here is simple. As the replacement ratio falls, the income risk faced by households rises, leaving more room for bankruptcy to be a useful form of implicit insurance.

Table 3 Effects of Lower UI Replacement Ratios

Panel A: Welfare Effects of Lower UI Replacement, without Bankruptcy

θ	r^l	r^d	Bankruptcy Rate	Welfare Change ($)	Welfare Change ($) Rel. to Benchmark
0.66	4.39%	4.39%	—	—	+$66.88
0.50	4.21%	4.21%	—	−$23.13	+$43.71
0.40	4.06%	4.06%	—	−$40.68	+$26.14

Panel B: Welfare Effects of Lower UI Replacement, with Bankruptcy

θ	r^l	r^d	Bankruptcy Rate	Welfare Change ($)
0.66	13.00%	2.57%	0.90%	—
0.50	13.00%	2.35%	0.99%	−$26.82
0.40	12.55%	2.31%	1.03%	−$36.60

Because bankruptcy causes less harm when the replacement ratio is low than when it is high, it appears that bankruptcy does play an insurance role. To see this, consider Panel A of Table 3 for the results when bankruptcy is not allowed. Welfare (relative to the case where $\theta = 0.66$, and bankruptcy is *not* allowed) falls slightly with the replacement ratio, by the equivalent of $23.13 when θ falls from 0.66 to 0.50, and by $40.68 when θ falls from 0.66 to 0.40.

As seen in Panel B of Table 3, when bankruptcy is allowed, the bankruptcy rate rises systematically when θ falls from 0.66 to 0.40, from 0.90 percent in the benchmark to 1.03 percent, an increase of 100,000 filings annually. This effect is supported in recent empirical work of Fisher (2002), who finds that higher UI benefits are associated with lower bankruptcy rates.

When welfare is measured relative to the benchmark economy, as shown in Table 4, the welfare effect of eliminating bankruptcy, while always positive, becomes smaller as θ rises. The gain from eliminating bankruptcy, relative to the benchmark, is $66.88 when $\theta = 0.66$ but drops to $26.14 when θ falls to 0.40. As noted earlier, all else equal, the effect of an increased interest rate on borrowing and a lowered rate on savings deposits would be to worsen consumption smoothing. Yet such interest rate changes are actually associated with small improvements in consumption smoothing, as seen in the column "c.v.-Cons." Panels A and B of Table 4 show that when $\theta = 0.66$, the c.v. of consumption falls slightly, from 0.1347 without bankruptcy to 0.1336 when bankruptcy is allowed. This suggests that bankruptcy must be providing some offsetting consumption benefits. Nonetheless, the costs of implementing a bankruptcy system, from both the socially wasteful penalty of credit market

Table 4 Distributional Effects of Lower UI Replacement Ratios

Panel A: Distributional Effects of UI, without Bankruptcy

θ	Gini-Cons.	c.v.-Cons.	c.v.-Assets	Gini-Assets	Avg. Borr. (% of $\overline{Y}$)
0.66	0.0663	0.1347	1.6878	0.9497	−17.21%
0.50	0.0667	0.1363	1.6809	0.9465	−17.17%
0.40	0.0671	0.1374	1.6829	0.9477	−17.12%

Panel B: Distributional Effects of UI, with Bankruptcy

θ	Gini-Cons.	c.v.-Cons.	c.v.-Assets	Gini-Assets	Avg. Borr. (% of $\overline{Y}$)
0.66	0.0694	0.1336	1.4211	0.8027	−11.64%
0.50	0.0699	0.1349	1.4211	0.8029	−11.63%
0.40	0.0698	0.1352	1.4189	0.8014	−11.67%

exclusion against filers, as well as the nonpecuniary costs, cause overall welfare to fall.

Interest rate spreads are relatively stable, but the deposit rate does fall from 2.57 percent in the benchmark to 2.35 percent and 2.31 percent as θ falls from 0.66 to 0.50 to 0.40, respectively (see Panel B of Table 3). The fall in deposit rates is the consequence of households needing to save more in the face of greater income loss from unemployment than before. As all households attempt to save more, the interest rate on savings falls. Conversely, as the cost of funds for banks falls, the increased bankruptcy rate does not result in an increase in the level of the interest rate on loans, relative to the benchmark. In terms of consumption smoothing, however, the presence of bankruptcy helps in the face of reduced replacement ratios. In Panel A of Table 4, the c.v. of consumption rises from 0.1347 to 0.1363 to 0.1374, with reductions in θ, when bankruptcy is not allowed. When bankruptcy is allowed (see Panel B of Table 4), the c.v. of consumption rises by less, from 0.1336 to 0.1349 to 0.1352. We therefore have the following:

Result 2 *Reducing the UI replacement ratio lowers welfare slightly and increases bankruptcy rates. However, the* ***fall*** *in welfare is nearly independent of whether or not bankruptcy is allowed. Additionally, reducing the UI replacement ratio worsens consumption smoothing less when bankruptcy is allowed.*

By making repayment optional, bankruptcy has the potential to reduce the need to actively accumulate and decumulate savings in the face of income shocks. Furthermore, in an economy where bankruptcy is allowed, the interest rate on loans might be prohibitively high, while that on savings very low, thereby retarding the ability of households to smooth consumption by borrowing and saving frequently. Indeed, for both reasons, bankruptcy appears to

significantly lower asset trade. In particular, whenever bankruptcy is allowed, the volume and volatility of asset trade fall sharply, as seen in Panels A and B of Table 4. For example, compare the case when bankruptcy is allowed under the benchmark replacement ratio (Panel B) to the case where bankruptcy is eliminated under benchmark replacement ratios (Panel A). The coefficient of variation of assets jumps from 1.42 to 1.68 and the Gini Coefficient for assets similarly rises from 0.80 to 0.95. The average volume of borrowing, denoted "Avg. Borr.," also jumps from a roughly 11.6 percent debt-income ratio (approximately \$4,000 per household annually), which is close to the 8.5 percent level found in the data (CBO [2000]), to roughly 17 percent of median annual income (or \$7,000 per household).[14] Note, however, that in all cases, the response of asset trading to reductions in the replacement ratio is very modest. Therefore, we have the following:

Result 3 *Bankruptcy lowers asset trade and makes the distribution of wealth more equal. However, changes in the UI replacement ratio do not greatly alter asset trade.*

As seen above, when bankruptcy is prohibited, the premium on borrowing falls. This fall is in turn associated with a great deal more borrowing. For equilibrium to obtain in the credit market, however, it must also be the case that households save more in good times. In turn, one might expect the interest rate (recall that there is only one interest rate in the absence of bankruptcy) to rise. Indeed, when bankruptcy is eliminated, the rate of interest on bank deposits rises sharply from 2.57 percent under benchmark UI replacement ratios to 4.39 percent when bankruptcy is eliminated.

Given that both unemployment insurance and bankruptcy protection provide some insurance, it is useful to ask the following: If households had to choose either one, but not both, which would households prefer? Table 5 shows the results for four polar cases. Not surprisingly, it is unemployment insurance that is quantitatively much more important than bankruptcy. Welfare is lowest when bankruptcy is allowed and UI is driven down to $Y_{\min}$ by setting $\theta = 0.40$. The latter generates a utility level of -37.65 units. When bankruptcy is allowed but $\theta = 0.66$ (the benchmark case), utility rises to -37.57 units. When bankruptcy is not allowed and $\theta = 0.40$, welfare climbs further to -37.52 units. Last, allowing UI alone, with $\theta = 0.66$, produces the highest welfare, -37.44 units. In dollar terms, the quarterly welfare consequences range from $-\$36.60$ when bankruptcy is allowed and $\theta = 0.40$, to \$66.88 when bankruptcy is not allowed and $\theta = 0.66$, to $+\$26.14$ when bankruptcy is not allowed and $\theta = 0.40$. For exposition, let $Welf(Bk = \{Yes, No\}, \theta)$ denote the welfare under a regime where bankruptcy is either allowed (whereby

[14] Specifically, this measures, conditional on borrowing, the mean level of unsecured debt held by households.

$Bk = Yes$), or not ($Bk = No$), and a UI replacement ratio, θ. We can express the following rank ordering for welfare:

Result 4 $Welf(No, \theta = 0.66) > Welf(No, \theta = 0.40) > Welf(Yes, \theta = 0.66)[Benchmark] > Welf(Yes, \theta = 0.40)$. *Therefore, if society must choose either UI or bankruptcy, it should choose UI. Furthermore, even if it could choose both UI and bankruptcy, a society should choose UI alone.*

Also, as mentioned above, not allowing bankruptcy even when UI is very strict ($\theta = 0.40$) improves welfare relative to allowing bankruptcy when UI is generous ($\theta = 0.66$). This is the sense in which bankruptcy is quite damaging. The intuition for this is that better UI coverage mutes the consequences of exclusion from the credit market and makes bankruptcy more attractive. This raises a more general issue.

Remark 1 *Any program that smooths a household's income lowers the need for access to credit markets. Therefore, bankruptcy becomes most attractive precisely when it is least necessary.*

Thus far, I have held the subsistence level of income, $Y_{\min}$, fixed while altering the replacement ratio and bankruptcy law. The subsistence level of income is meant to represent the combined effects of all social insurance programs beyond unemployment insurance. One abstraction is that the period is thirteen weeks long, when eligibility for unemployment benefits is typically at least twenty-six weeks. In the benchmark setting, $Y_{\min}$ could be thought of as representing these extra benefits in the remaining thirteen weeks (if one qualifies), after which other income support programs might take over. I now briefly note the effects of cutting UI off after one period, followed by only minimal public assistance. To this end, I set public assistance to cover just 10 percent of median household income, whereby $Y_{\min} = 0.10$. In this case, the household that is no longer qualified for UI receives the equivalent of only $333 monthly in public assistance.[15] I will not discuss these results in detail, but will note the following findings: First, both savings and borrowing interest rates fall, as precautionary savings rise. Second, welfare rises as bankruptcy is allowed. Third, welfare rises by increasing amounts as the replacement ratio falls, consistent with an increased insurance role. Last, the reductions in welfare emerging from reductions in the UI replacement ratio are smaller when bankruptcy is allowed than when it is not. Therefore, because the results for $Y_{\min} = 0.10$ reverse those where $Y_{\min} = 0.40$, we are led to the following conclusion:

Result 5 *Bankruptcy's role in providing insurance is clearly dependent on the existing social safety net. For example, when $Y_{\min}$ is lowered to 0.10, allowing bankruptcy improves welfare relative under all UI replacement ratios.*

[15] Assuming a $40,000 median annual income.

Table 5 Bankruptcy and UI: Four Polar Cases

Bkrptcy., θ	r^l	r^d	Bkrptcy. Rate	c.v.-Cons.	Gini-Cons.	Welfare Change ($)	Utility
Bench							
[Yes, $\theta = 0.66$]	13.00%	2.57%	0.90%	0.1336	0.0694	—	−37.57
Yes, $\theta = 0.40$	12.55%	2.31%	0.99%	0.1352	0.0698	−$36.60	−37.65
No, $\theta = 0.66$	4.39%	4.39%	—	0.1347	0.0663	+$66.88	−37.44
No, $\theta = 0.40$	4.06%	4.06%	—	0.1374	0.0671	+$26.14	−37.52

Corollary 1 *The debate over bankruptcy protection (in the presence of existing insurance programs) should be centered on the quantitative aspects of income uncertainty.*

This result is also consistent with the recent work of Livshits et al. (2002), who find, in a life-cycle setting, that the presence of large uninsured medical shocks allows bankruptcy to play a role in improving welfare.

3. FINAL REMARKS

I have developed a stylized model of employment, unemployment, and bankruptcy in order to better understand how the consumption "insurance" provided by bankruptcy interacts with that provided by explicit unemployment insurance programs.

Five results are worth noting. First, in the benchmark economy, introducing bankruptcy under even low UI replacement ratios lowers welfare. Second, reducing the UI replacement ratio lowers welfare slightly and increases bankruptcy rates. Although the fall in welfare is nearly independent of whether bankruptcy is allowed or not, reducing the UI replacement ratio worsens consumption smoothing less when bankruptcy is allowed. Third, bankruptcy lowers asset trade and makes the distribution of wealth more equal. However, asset trading behavior is not affected greatly by changes in the UI replacement ratio. Fourth, UI is more important than bankruptcy: If society must choose either UI or bankruptcy, it should choose UI.

Last, bankruptcy's role in providing insurance is clearly dependent on the existing social safety net. In summary, unemployment insurance appears to materially affect the desirability of bankruptcy protection. Were other social assistance to be scaled sharply back, the results suggest that bankruptcy could serve a useful insurance role in the United States. However, as currently practiced, income risk, broadly defined, does not appear high enough to justify bankruptcy in the presence of unemployment insurance. Indeed, when

unemployment insurance is set to current levels, bankruptcy actually appears to harm the efficacy of the UI system.

A potentially important abstraction in the model is the absence of moral hazard that could limit the extent of socially desirable insurance protection. Specifically, unemployed households in the model do not alter their job search efforts in the face of insurance payments but rather face an exogenous probability (ξ) of return to employment. Hansen and Imrohoroglu (1992) find that when households are allowed to reject job offers while unemployed, but are subject to random (or imperfect) auditing by the government, the welfare maximizing level of insurance is much lower than otherwise. Furthermore, effort expended by workers *while employed* may fall with the promise of generous unemployment insurance. Also, the availability of bankruptcy will help reduce the incentive effects of strict unemployment insurance and may further increase moral hazard. The experience rating of employers lowers the willingness of firms to fire lazy workers, leading again to the possibility of reduced effort. Moral hazard in an economy where output is explicitly produced using labor leads in turn to lower output, quite unlike the pure endowment setting employed here. The model also places fixed limits on credit availability that do vary with bankruptcy law. It is possible that a strict bankruptcy code would improve access to credit.

Nevertheless, there are reasons to suspect that the simple environment developed here does provide a useful first pass at the interactions between the UI system and the personal bankruptcy system. In particular, the unemployment insurance in the model is strictly capped at one period, and the re-entry probability of 0.10 by no means provides comfortable income prospects for those who fail to find work. With respect to the robustness of using fixed credit limits, note that the elimination of bankruptcy is treated here synonymously with the prohibition of default. To the extent that informal default would become more prevalent were bankruptcy outlawed, the expansion of credit availability might be limited. With that said, in ongoing work (Athreya [2002a]), I augment the model developed here to include moral hazard in job search effort, as well as capital accumulation and the production of output where labor effort matters. This article is therefore a first step in the analysis of how the interactions between bankruptcy and an existing social insurance program determine the desirability of changes to each one in isolation.

APPENDIX: DEFINING EQUILIBRIUM

A stochastic stationary equilibrium is defined as follows: Let $X = A \times \widetilde{Y} \times CS$ denote the state space for households, where $CS = \{S, BC\}$. Let χ_B be the Borel σ-algebra on X. The household's asset decision rule is denoted $a(x)$.

The decision rule and the uncertainty of income together imply a stochastic process for consumption and asset holdings, with an associated transition function $Q(x, Z), \forall Z \in \chi_B$ on the measurable space (X, χ_B). This transition function implies a stationary probability measure $\mu(Z)$ for all $Z \in \chi_B$. This is a measure on subsets of X that describes the joint distribution of households on asset holdings, current income, and credit market status. For a measure to be stationary, it must satisfy the following fixed-point condition:

$$\mu(Z) = \int_X Q(x, Z) d\mu.$$

This implies fixed interest rates on loans and deposits and a constant fraction of bankrupt households. Not every stationary probability measure, however, qualifies as part of an equilibrium. Since the private bond market must clear, aggregate holdings of private bonds must be zero. Additionally, all public debt must be held in equilibrium. Therefore, market clearing requires that the aggregate supply of bonds equals the stock of public debt, D.

Next, as the banking sector is competitive, profits also must be zero. The zero-profit constraint is motivated as follows: First, let $X_{neg} = \{x \in X | a < 0\}$ denote the subset of the state space X such that households hold negative asset balances. In the stationary state, there is a time-invariant mass of households, whose total borrowing is given by $\int_{X_{neg}} a(x) d\mu$. The total revenue for the intermediary will therefore be $(1 + r^l)(|\int_{X_{neg}} a(x) d\mu|)$. Analogously, the total cost of funds for the intermediary is determined by total borrowing times the gross deposit interest rate, $|(1 + r^d) \int_{X_{neg}} a(x) d\mu|$. The losses from default are on both interest and principal from those who borrow. Define $\pi(x)$ to be the probability that a household in state x will default. Total principal losses are therefore $|\int_X a(x)\pi(x) d\mu|$. The zero profit condition on intermediaries is then: $(1 + r^l)((|\int_{X_{neg}} a(x) d\mu|) - (|\int_{X_{neg}} a(x)\pi(x) d\mu|)) - |(1 + r^d) \int_{X_{neg}} a(x) d\mu| = 0$. (Note that the aggregate default rate is then given by $\Pi \equiv \int_X \pi(x) d\mu$.) Lastly, the unemployment insurance system must collect revenues equal to outlays, i.e., $\eta_u = (1 - \rho)\mu_e \theta \overline{Y}$. The following five equations will therefore define equilibrium.

Definition 2 *A stationary equilibrium of the model is a four-tuple,* $(a(x), \pi(x), \mu(Z), (r^l, r^d))$, *that satisfies four conditions.*

1. The decision rule, $a(x)$, *is optimal, given* r^d *and* r^l.

2. $\mu(Z)$ *is stationary:* $\mu(Z) = \int_X Q(x, Z) d\mu$ *for all* $Z \in \chi_B$.

3. Asset market clearing: $\int_X a(x) d\mu = D$.

4. Zero profits: $(1 + r^l)((|\int_{X_{neg}} a(x) d\mu|) - (|\int_{X_{neg}} a(x)\pi(x) d\mu|)) - |(1 + r^d) \int_{X_{neg}} a(x) d\mu| = 0$.

5. Unemployment insurance fund breaks even: $\eta_u = (1 - \rho)\mu_e \theta \overline{Y}$.

I use simple discrete state approximations to the value functions, conditional on income and credit market status, and then use Monte Carlo integration with antithetic variates to compute all integrals. I then bisect on both r^l and r^d until I simultaneously clear markets and satisfy the zero-profit condition.

REFERENCES

Aiyagari, S. R. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *Quarterly Journal of Economics* 109: 659–84.

Alvarez, F., and M. Veracierto. 2001. "Severance Payments in an Economy with Frictions." *Journal of Monetary Economics* 47: 477–98.

Athreya, K. B. 2002a. "Personal Bankruptcy and Unemployment Insurance." Federal Reserve Bank of Richmond. Mimeo.

_____. 2002b. "Welfare Implications of the Bankruptcy Reform Act of 1999." *Journal of Monetary Economics* 49: 567–95.

Chatterjee, S., D. Corbae, M. Nakajima, and J. V. Rios-Rull. 2001. "A Quantitative Theory of Unsecured Consumer Credit with Risk of Default." University of Pennsylvania. Mimeo. Available at http://www.ssc.upenn.edu/~vr0j.

Cochrane, John. 1991. "A Simple Test of Consumption Insurance." *Journal of Political Economy* 99: 957–76.

Congressional Budget Office. 2000. "Personal Bankruptcy: A Literature Review." CBO paper. Washington, D.C.

Domowitz, I., and R. Sartain. 1999. "Determinants of the Consumer Bankruptcy Decision." *Journal of Finance* 54(1): 403–20.

Dubey, P., J. Geanakoplos, and M. Shubik. 2000. "Default in a General Equilibrium Model with Incomplete Markets." Cowles Foundation Discussion Paper 1247. Available at http://cowles.econ.yale.edu/P/cd/.

Fisher, J. D. 2002. "The Effect of Transfer Programs on Personal Bankruptcy." Bureau of Labor Statistics Working Paper 346. Available at http://www.bls.gov/ore/orecatlg.htm.

Hansen, G., and A. Imrohoroglu. 1992. "Optimal Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard." *Journal of Political Economy* 100: 118–42.

Heaton, J., and D. Lucas. 1997. "Market Frictions, Savings Behavior, and Portfolio Choice." *Macroeconomic Dynamics* 1: 76–101.

Hopenhayn, H., and J. P. Nicolini. 1997. "Optimal Unemployment Insurance." *Journal of Political Economy* 105: 412–38.

Huggett, M. 1993. "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies." *Journal of Economic Dynamics and Control* 17: 953–69.

Li, W., and P.-D. G. Sarte. 2002. "The Macroeconomics of U.S. Consumer Bankruptcy Choice: Chapter 7 or Chapter 13?" Federal Reserve Bank of Richmond Working Paper 02-1 (September).

Livshits, Igor, James MacGee, and Michele Tertilt. 2001. "Consumer Bankruptcy: A Fresh Start." Federal Reserve Bank of Minneapolis Working Paper 617 (December). Available at woodrow.mpls.frb.fed.us/research/wp.

Ljungqvist, L., and T. Sargent. 2000. *Recursive Macroeconomic Theory*. Cambridge: MIT Press.

Meyer, B. 1990. "Unemployment Insurance and Unemployment Spells." *Econometrica* 58: 752–82.

Sullivan, T., E. Warren, and J. Westbrook. 2000. *The Fragile Middle Class: Americans in Debt.* New Haven: Yale University Press.

Economic Fundamentals and Bank Runs

Huberto M. Ennis

Recently there has been a renewed discussion in the literature about the determinants of bank runs. Two alternative theoretical explanations are usually provided. According to the first theory, bank runs are exclusively driven by changes in economic fundamentals, such as a deterioration in the return on investment. The second theory views bank runs as a consequence of the existence of multiple equilibria. In the latter case, which equilibrium obtains depends on the realization of an extrinsic random variable, often called "sunspots." *Extrinsic uncertainty* is uncertainty in economic outcomes that does not originate directly in changes of economic fundamentals (see Shell and Smith [1992]). The word "sunspots" is intended to convey the idea that these random variables do not directly influence the economic fundamentals of the economy.[1] However, sunspots can still influence economic outcomes to the extent that people believe they do. In this sense, sunspots can be viewed as coordination devices for agents' expectations in decentralized market economies. This is the view adopted in the bank-run literature and in this paper.

Some scholars have recently argued that the multiple-equilibria-plus-sunspots explanation of bank runs is inconsistent with available evidence

Research Department, Federal Reserve Bank of Richmond, huberto.ennis@rich.frb.org. Some of the ideas discussed in this article are the product of my joint work with Todd Keister from Instituto Tecnológico Autónomo de México. I would like to thank Emilio Espino, Tom Humphrey, Ned Prescott, John Weinberg, Alex Wolman, and especially Todd Keister for comments on an earlier draft. All errors are of course my own. The views expressed here do not necessarily reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

[1] Shell and Smith (1992, 602) write: "The 'sunspot' terminology is a bit of a spoof on the work of Jevons (1884) and his followers, who related the business cycle to the cycle of actual sunspots. To the extent that actual sunspots do affect economic fundamentals this is intrinsic uncertainty, but the overall effects of actual sunspots on economic fundamentals are probably not major. Then, if actual sunspot activity does have substantial impacts on the economy, it must be that it serves a role beyond its effects on fundamentals. Cass-Shell (1983) sunspots are highly stylized; by definition, they represent purely extrinsic uncertainty."

showing that bank runs have historically been strongly correlated with deteriorating economic fundamentals (see Gorton [1988]; Allen and Gale [1998]; and Schumacher [2000]). In this paper I will argue that such a conclusion is not well justified. More specifically, I will show that the multiple-equilibria model of bank runs, combined with a reasonable (and well-accepted) equilibrium selection concept, can provide theoretical justification for the correlation observed in the data. In other words, the presence of an empirical correlation between bank runs and poor economic fundamentals cannot be used to discriminate between the two competing theories. Furthermore, the equilibrium selection story presented here strongly accords with the long-standing belief that some bank runs can be characterized as events resulting from exogenous waves of pessimism and that those mood shifts are more likely when economic conditions are bad or deteriorating.

The empirical evidence that links bank runs to economic conditions has been well documented. Gorton (1988) discusses what he calls the "recession hypothesis," according to which bank panics are closely associated with the business cycle. In a related paper, Miron (1986) presents evidence in favor of the "seasonal hypothesis," which is that bank runs tend to be correlated with seasonal fluctuations in the liquidity needs of depositors. Saunders and Wilson (1996) and Schumacher (2000) discuss evidence on the selectivity of depositors: not all banks are equally likely to experience a run during a panic, and in particular a questionable solvency position prior to the run tends to increase the probability of depositors running on a particular bank.[2]

Gorton (1988) studies bank panics during the National Banking Era (1865–1914). Using data for national banks, Gorton investigates whether the model and variables that explain the behavior of depositors during no-panic situations also explain their behavior during panics. In this sense, panics would not be purely random events; rather, they would be directly correlated with the arrival of new information that determines depositors' desire to withdraw funds from the bank. Gorton finds no evidence for something special happening during panics that cannot be explained by the model that describes the behavior of depositors in no-panic situations. Instead, the evidence seems to suggest that panic events are just the consequence of extreme realizations of the circumstances that explain behavior during normal times. It is important to note, however, that Gorton finds examples in which shocks of equal magnitude to those usually associated with runs did not cause a panic (for example, the November 1887 spike in the liability of failed businesses did not induce a

[2] Calomiris and Mason (1997) find evidence of depositors' confusion during the June 1932 bank panic, but they also find that solvent banks were able to support each other to avoid failure.

Table 1 Financial Panics, 1890–1908 (Miron, 1986)

Major Panics	September 1890
	May 1893
	December 1899
	May 1901
	March 1903
	October 1907
Minor Panics	February 1893
	September 1895
	June 1896
	December 1896
	March 1898
	September 1899
	July 1901
	September 1901
	September 1902
	December 1904
	April 1905
	April 1906
	December 1906
	March 1907
	September 1908

panic, while the smaller increase in June 1884 did). Finally, in Tables 1 and 2 we can see that there is some disagreement as to what constitutes a panic. For example, Gorton does not consider the episodes of May 1901 and March 1903 as panics. Furthermore, and more germane to this paper, Tables 1 and 2 suggest that there were several bank panics in periods with no economic recession. Of course, seasonality may be part of the answer in those cases (as discussed by Miron [1986]).[3]

These are interesting findings, but they are not enough to rule out the possibility that, in some cases, banking panics are associated with the existence of multiple equilibrium outcomes (that is, situations where both the panic and the no-panic outcomes are possible). These stylized facts refute only the simplest way of modeling multiple equilibria and even then only under fairly specific conditions. Showing that reasonable theories of multiple-equilibria bank runs are not refuted by the available evidence is important since policy prescriptions depend on the assessment of the economic conditions that generate those bank runs. It would be helpful for policymakers to be able to conclude that multiple-equilibria bank runs are not the norm. However, as I will show here, the evidence discussed above does not allow us to reach that conclusion.

[3] Gorton (1988) finds no evidence of seasonal effects as causes for panics using his definition.

Table 2 Business Cycle and Bank Panics (Gorton, 1988)

NBER Cycle (Peak-Trough)	Panic Date
October 1873 - March 1879	September 1873
March 1882 - May 1885	June 1884
March 1887 - April 1888	No panic
July 1890 - May 1891	November 1890
January 1893 - June 1894	May 1893
December 1895 - June 1897	October 1896
June 1899 - December 1900	No panic
September 1902 - August 1904	No panic
May 1907 - June 1908	October 1907
January 1910 - January 1912	No panic
January 1913 - December 1914	August 1914

The paper is organized as follows. In the next section I discuss a simple model of bank runs that is now standard in the economic literature. I then study the conditions under which multiple equilibria arise, and I review different theories of how an equilibrium is selected in those cases. I show that some of the more appealing equilibrium selection mechanisms are indeed compatible with the available evidence. Finally, in the conclusion I discuss some policy implications.

1. MODELING BANK RUNS

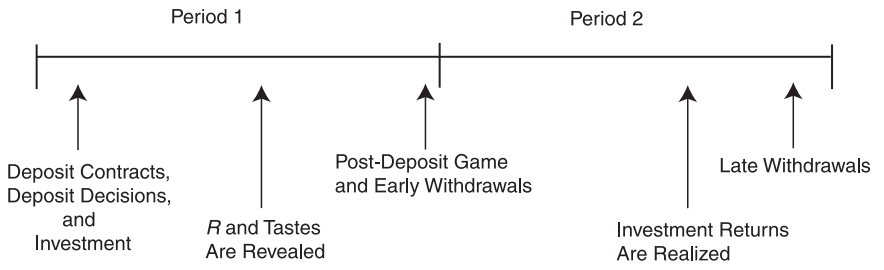
The Environment

The environment is similar to that in Diamond and Dybvig (1983), except that the return on investment is stochastic. There are two time periods, $t = 1, 2$, and a large number of ex ante identical agents (a continuum of agents with unit mass). Each agent is endowed with a consumption good at the beginning of date 1 and none after that. Agents are uncertain about their preferences: some will be impatient and will need to consume at the end of period 1; the rest will be patient and can wait to consume in period 2. At the beginning of period 1 agents do not know whether they will be patient or impatient, but they know that the probability of being impatient at the end of the period is u. Preferences are represented by the following utility function:

$$v(c_1, c_2) = \begin{cases} \frac{1}{\gamma}(c_1)^{\gamma} & \text{with probability} \quad u \\ \frac{1}{\gamma}(c_1 + c_2)^{\gamma} & \text{with probability} \quad 1 - u \end{cases},$$

where c_1 is consumption at the end of period 1, c_2 is consumption at period 2, and $\gamma < 1$. The realization of preference types is independent across agents, implying that u will also be the fraction of the population that becomes impatient. Agents' types are not observable and hence patient agents can

Figure 1 Timing



always pretend to be impatient if they wish to do so (impatient agents could pretend to be patient, but this is never the case for the contracts studied below).

There are two saving technologies available: storage and investment. One unit of consumption placed in storage yields one unit of consumption at any future time. For the investment technology, one unit of consumption placed in investment at the beginning of period 1 yields R units in period 2. The return on investment R is a random variable taking values greater than unity and with a probability density function given by $f(R)$. Note that the expected value of R is necessarily greater than one and hence investment is a better technology than storage to save consumption for the second period (that is, for funds that are needed with certainty in the second period). If investment is liquidated early (at the end of period 1), then it yields $x < 1$ units of consumption per unit invested. Hence, investment is an illiquid asset that yields a higher return than storage if held to maturity, but a lower return if liquidated early.

Timing

Since agents do not know their preferences until after the opportunity to invest has passed, they pool their endowments in banking coalitions. These banks then allocate some resources into the illiquid investment and provide insurance to their members in case they happen to become impatient at the end of period 1.

Competition in the banking industry drives the banks to offer the best possible available contract to consumers. I restrict the type of contracts that banks can offer to simple deposit contracts that are subject to a sequential service constraint (Wallace [1998]). Under this type of deposit contract, an agent gets the right to either a fixed payment at the end of period 1 (as long as the bank has funds) or a contingent payment in period 2. The sequential service constraint prevents the bank from adjusting the payment to early withdrawers according to the number of agents that decide to withdraw early. The bank must pay a fixed amount until it runs out of funds. This kind of contract is in

the tradition of Diamond and Dybvig (1983) and Cooper and Ross (1998). I use it here mainly because of its simplicity and potential descriptive content.[4]

The timing of events is as follows. At the beginning of period 1, the bank, without knowing the value of R, chooses a deposit contract and a portfolio of assets (investment is possible only at this point). This choice can be summarized by the pair (a, η), where a is the payment that the bank will give to depositors if they decide to withdraw early and η is the proportion of total deposits that the bank decides to keep in storage (with $(1-\eta)$ being the proportion that the bank puts in the illiquid investment technology). Also at this time, agents decide whether or not to deposit their funds in the bank. At the end of period 1, the uncertainty about preferences and technology is resolved: agents find out whether or not they are impatient and the value of R is revealed.[5] At this time, then, agents decide whether or not to go to the bank to withdraw their deposits. Impatient agents have no choice but to withdraw early. Patient agents, however, could choose to wait until period 2, which they will do if they are not better off imitating the impatient agents. Whether a patient agent would be better off withdrawing his or her deposits early depends, in general, on what all the other patient agents are doing. Hence, patient agents play a strategic game at the end of period 1. Following Peck and Shell (2003), I shall call it the "post-deposit game." In period 2, the return on the illiquid technology is realized and those agents that did not withdraw their deposits early (at the end of period 1) go to the bank and share the total remaining resources equally.

2. THE POST-DEPOSIT GAME

The source of multiplicity of equilibria in the model lies in the post-deposit game played by patient agents. The expected outcome of this game will determine the bank's investment decisions and the willingness of agents to make deposits in the bank. The details of those problems are presented in Section 4. What is important here is to understand that solving those problems requires knowing what could happen in the post-deposit game. For this reason, I turn next to the study of this game.

---

[4] See also Ennis and Keister (2003b). In this environment, there are potential gains from making the early payments contingent on the realization of the return on investment R. The contracts studied here do not allow for this possibility. Gale and Vives (2002) and Allen and Gale (1998) do not assume sequential service, but the optimal contract has a structure similar to the deposit contract in the sense that for high values of R the payoff to early withdrawers is not contingent. This is because investment cannot be liquidated (it has zero liquidation value), and for high enough values of R (so that late consumers get more than early consumers), early consumers just divide the available liquid funds among them, resulting in a fixed quantity for each, independent of the value of R. The costly state verification literature provides another justification for the debt contracts (see, for example, Williamson [1986]).

[5] This value of R is common to all investment in the economy. No diversification is possible.

Table 3 Notation

u	Probability of being impatient
γ	Coefficient of relative risk aversion
R	Return on the risky investment
x	Return from early liquidation
$f(R)$	Probability distribution of R
a	Bank payment for early withdrawal
η	Proportion of total deposits held in storage
$\overline{\overline{u}}$	Probability of getting paid in case of run
R^*	Multiple-equilibria threshold for R
$\widehat{R}$	Risk-dominance threshold for R
p_r	Risk factor of the bank-run equilibrium
π	Probability of a bank run

I concentrate only on symmetric pure strategy equilibria.[6] At the end of period 1, the patient agents are faced with the decision of whether to withdraw their deposits early or leave them in the bank until period 2. Let r denote the decision to go to the bank to withdraw (i.e., to run) and n the decision to wait until the next period (i.e., not to run). Let us define as $P_{ij}(R; a, \eta)$ the payoff to a patient agent following action i $(i = r, n)$ given that all other patient agents are following action j $(j = r, n)$. We need only to consider those payoffs because we are looking at symmetric equilibria, where all patient agents act in the same manner. The normal form of the post-deposit game played by patient agents is given by the following matrix:

		Other Patient Agents: Run	Other Patient Agents: No Run
Patient Agent	Run	$P_{rr}(R; a, \eta)$	$P_{rn}(R; a, \eta)$
	No Run	$P_{nr}(R; a, \eta)$	$P_{nn}(R; a, \eta)$

Note that the payoff $P_{ij}(R; a, \eta)$ depends on the return on investment R and on the deposit contract chosen by the bank (a, η). (Note also that deviations by a single player do not change the payoff to the rest of the players because we are assuming that there is a large number of players.)

It is easy to state conditions under which this game has multiple equilibria. In particular, if $P_{rr}(R; a, \eta) > P_{nr}(R; a, \eta)$ and $P_{nn}(R; a, \eta) > P_{rn}(R; a, \eta)$, then running to the bank at the end of period 1 and waiting until period 2 to withdraw are both equilibria of the game. To see this, note that when $P_{rr}(R; a, \eta) > P_{nr}(R; a, \eta)$ holds, if the patient agent thinks that all other patient agents will run to the bank, then it is in her best interest to run as

[6] Symmetry implies that in equilibrium all impatient agents play the same strategy and all patient agents play the same strategy (but perhaps different from the one played by the impatient agents). Pure strategies are those strategies that do not involve randomization over different possible actions (each agent plays a single action with probability one).

well. Therefore, if all patient agents believe that a run will occur, the run does occur and running is a Nash equilibrium of the game. Likewise, when $P_{nn}(R; a, \eta) > P_{rn}(R; a, \eta)$ holds, if the patient agent thinks that no other patient agent will run to the bank, then it is in her best interest not to run. Therefore, if all patient agents believe that there will be no run, there is indeed no run and not running is a Nash equilibrium of the game. In equilibrium, then, all players play the same strategy, and I will denote each equilibrium by the strategy being played in it. Thus, I call the run equilibrium (if it exists) "equilibrium r," and the no-run equilibrium "equilibrium n."

Another important characteristic of this post-deposit game is that the multiple equilibria are usually Pareto-ranked.[7] One equilibrium is better than another equilibrium in the Pareto sense if all players in the former receive a payoff at least as high as in the latter and one or more players receive a strictly higher payoff. In the game studied here, if $P_{nn}(R; a, \eta) > P_{rr}(R; a, \eta)$, then the no-run equilibrium n is Pareto-preferred to the run equilibrium r.

Given the possibility of multiple equilibria, the natural next step is to ask, how does one of the equilibria get selected? I will discuss the answer to this question in the next section.

Before going into the equilibrium selection issue, it is worth noting that we can further characterize the payoff matrix of the post-deposit game. Studying these payoffs will give us a better idea of the conditions that determine the existence of multiple equilibria in the game.

Since the bank chooses the contract before observing the return R, the values of η and a depend only on the probability distribution of R and not on the particular realizations of R. The bank will never choose a contract such that $ua > \eta$ holds. In such a case, the bank will be certain to need to early-liquidate some of the investment in order to pay depositors (even if no patient agent runs). Since early liquidation is costly, this contract is never optimal. I will study the problem of the bank later, but for now let us assume that the distribution of R is such that the bank chooses a contract (a, η) that satisfies $\eta + x(1 - \eta) < a$. This inequality implies that if every agent goes to the bank early, then the bank would run out of resources before being able to pay the promised amount a to each withdrawer. Furthermore, if the inequality does not hold, then there would be no runs in equilibrium. These two inequalities allow us to determine the value of waiting when there is a run, $P_{nr}(R; a, \eta)$, and the value of running when there is no run, $P_{rn}(R; a, \eta)$. First, we have that $P_{nr}(R; a, \eta) = 0$ because if (almost) every agent goes to the bank to withdraw early, then the bank will run out of funds and no payments will be made in the second period. Second, we have that $P_{rn}(R; a, \eta) = P_{rn}(a) = a^{\gamma}/\gamma$ because when only impatient agents withdraw early, total withdrawals are equal to ua

[7] Games with multiple Pareto-ranked equilibria are called "coordination games" in the literature (for a general review, see Cooper [1999]).

and the bank has access to enough liquid funds, $\eta + x(1-\eta)$, to cover that amount.

Let us now define $\overline{\overline{u}} \equiv [\eta + x(1-\eta)]/a < 1$ as the probability of being paid when every agent goes to the bank early. This formula is a direct consequence of assuming that agents take random positions in the line formed at the bank's window and that there is a sequential service constraint. Thus, we have that $P_{rr}(R; a, \eta) = P_{rr}(a, \eta) = \overline{\overline{u}} a^{\gamma}/\gamma$. It is important to note that P_{nr}, P_{rn}, and P_{rr} are not functions of the particular realization of R. The only payoff that is a direct function of the realization of R is that for late withdrawals when there is no run, that is

$$P_{nn}(R; a, \eta) = \frac{1}{\gamma}\left(\frac{R(1-\eta) + (\eta - ua)}{1-u}\right)^{\gamma}.$$

Note that $P_{nn}(R; a, \eta)$ is a continuous, increasing, and unbounded function of R. Hence, there exists a threshold value R^* such that if $R > R^*$, we have that $P_{nn}(R; a, \eta) > P_{rn}(a) = a^{\gamma}/\gamma$ and the post-deposit game is a multiple-equilibria coordination game. If $R < R^*$, the post-deposit game has a unique equilibrium in which all agents withdraw their deposits at the end of period 1. In summary, the payoff matrix for the post-deposit game is:

		Other Patient Agents: Run	Other Patient Agents: No Run
Patient Agent	Run	$\frac{1}{\gamma}\overline{\overline{u}} a^{\gamma}$	$\frac{1}{\gamma} a^{\gamma}$
	No Run	0	$\frac{1}{\gamma}\left(\frac{R(1-\eta)+(\eta-ua)}{1-u}\right)^{\gamma}$

3. EQUILIBRIUM SELECTION IN THE POST-DEPOSIT GAME

There is an extensive literature on equilibrium selection in games. This literature has concentrated some attention on 2×2 games with multiple equilibria. The post-deposit game of the previous section can be thought of as just an example of a 2×2 symmetric game with the potential for multiple equilibria (i.e., a 2×2 symmetric coordination game).[8] In this section, I will review some of the basic ideas from this literature and discuss how they apply to the bank-run problem at hand.

It is useful at this point to introduce the concept of *equilibrium selection mechanism* (ESM). An ESM is a probability distribution that assigns, to each equilibrium of the game, a probability indicating how likely it is to be the result of play. For the post-deposit game under consideration, an ESM is a function that for each possible triplet (R, a, η) assigns a probability π to

[8] Usually we refer to a 2×2 game as a game that is played by two individuals who each have two possible pure strategies that they can choose to play. In the post-deposit game, agents play a "game" against the population that is often called a "macroeconomic game." See Cooper (1999) for an extensive discussion on the subject.

the run equilibrium (r) and a probability ($1 - \pi$) to the no-run equilibrium (n). These probabilities must be feasible in the sense that, for given values (R, a, η), if the run equilibrium does not exist, then $\pi = 0$, and if a run is the only equilibrium, then $\pi = 1$. It is important to note that there is a degree of coordination being assumed from the outset: agents know that the only possible outcomes are those where all the rest of the agents play in the same manner (and this coordination is common knowledge). The ESM provides some structure to the coordination problem but does not explain why and how coordination arises. In this sense, the concept of an ESM can be thought of as a generalization of the traditional sunspot approach: there is still in place an exogenous coordination device on which all agents base their actions. The innovation is that the general ESM allows for the probability of each equilibrium to depend on exogenous and endogenous variables in the model.

The next natural question is, where does the function $\pi(R, a, \eta)$ come from? In the traditional sunspot approach, the function π is a constant between zero and one when feasible (i.e., when both equilibria exist). Another commonly used criterion for equilibrium selection is to assume that the best equilibrium (in the Pareto sense) will be selected. In this case, the ESM is such that the probability π is equal to zero as long as the no-run equilibrium exists and switches to unity when only the run equilibrium exists. Yet there are other possible forms that the function π may take and that can be reasonably justified. I review some of these forms next.

Let us start by defining the *risk factor* of equilibrium j, for $j \in \{r, n\}$ as the smallest probability p such that if a player believes that with probability strictly greater than p all the other players are going to play action j, then action j is the unique optimal action to take (see, for example, Young [1998]). Hence, the risk factor of the run equilibrium (r) is given by the solution to the following equation:[9]

$$p_r P_{rr} + (1 - p_r) P_{rn} = p_r P_{nr} + (1 - p_r) P_{nn}.$$

Therefore,

$$p_r = \frac{P_{nn} - P_{rn}}{(P_{rr} - P_{nr}) + (P_{nn} - P_{rn})}$$

is the risk factor of the run equilibrium. When both equilibria exist (run and no-run), the only payoff that depends on R is P_{nn}, and this payoff is increasing in R. Hence, p_r is an increasing function of R. This result is rather intuitive. It says that the higher the return on investment R, the higher the belief probability of a run p must be in order to induce a patient agent to run on the bank.

[9] The payoffs are still a function of the triple (R, a, η), but I choose not to explicitly write this dependence in order to simplify notation.

An equilibrium j is *p-dominant* if the equilibrium action j is the unique best response to any belief of the player that puts probability at least $p \in [0, 1]$ on the other players playing action j (see Morris, Rob, and Shin [1995]). Hence, the run equilibrium is p_r-dominant.

If the risk factor of the run equilibrium p_r is less than or equal to one-half, then the run equilibrium is *risk dominant* (Harsanyi and Selten [1988]). *Risk dominance* has been used as a criterion for equilibrium selection: the risk-dominant equilibrium will be the one selected and played. This criterion has an appealing interpretation. If each player is uncertain about the action of the other players, it is plausible that he or she would assign equal probability to each of the possible outcomes (a flat or diffuse prior). If the risk factor of equilibrium j is less than one-half, that is, if equilibrium j is risk dominant, and if players have flat priors about the actions of the other players, then equilibrium j will be the one played. In the post-deposit game, when each player assigns equal odds to all of his or her opponents playing either action r or n, then the players will choose to play the action of the risk-dominant equilibrium. In terms of the definition of ESM, the risk dominance criterion assigns probability one to the risk-dominant equilibrium.

Another way of motivating an equilibrium selection rule in games with multiple equilibria is to study learning dynamics under repeated iterations of the static (stage) game. See, for example, Kandori, Mailath, and Rob (1993); Young (1998); and Matsui and Matsuyama (1995). These papers concentrate on games with two players and assume that there are frictions limiting the ability of agents to adjust their strategies. Kandori, Mailath, and Rob also assume bounded rationality on the part of the agents playing the dynamic game (in the form of myopic behavior and some propensity to make mistakes). It is interesting to note that the learning dynamics under these assumptions tend to select (as the frictions or the probability of mistakes vanish) the risk-dominant equilibrium as the one most likely to be played. Temzelides (1997) extends this work and applies it to the bank-run model.

Ennis and Keister (2003a) study a learning model that induces a probability distribution over the possible equilibria of a 2×2 macroeconomic coordination game. We show that the probability of equilibrium j induced by this learning process is strictly decreasing in the risk factor of equilibrium j and can take values strictly lower than one even when equilibrium j is risk dominant. In terms of the previous ESM terminology, we have that the function π is a decreasing function of p_r and may take values strictly between zero and one. Since p_r is an increasing function of R (the fundamentals), we have that the probability of a run π is a decreasing function of R. That is, the better the fundamentals (R), the less likely is a bank-run event. In Ennis and Keister (2003b) we apply these ideas to study the effect of bank runs on economic growth.

Let us now go back to the case of equilibrium selection based on the traditional sunspot approach. Assume that the return on investment R takes values only in the interval (R^*, ∞), where R^* is the threshold such that for values of R greater than R^* there are multiple equilibria of the post-deposit game. In other words, assume that the contract is such that the no-run equilibrium exists for every possible value of R. Assume also that a binomial sunspot random variable determines which equilibrium is selected. Because both equilibria exist for every value of R, the probability of a bank run is always given by the constant probability associated with the sunspot realization that coordinates agents to "run" to the bank. This is the sense in which the previous literature on bank runs has dismissed the sunspot explanation for not conforming with the observed correlation of bank runs with economic fundamentals.

However, note that if R *can* be below R^* with positive probability, then for those realizations, regardless of the sunspot variable, the probability of a run will be equal to unity. In such a case, even though sunspots still play an important role in coordinating the agents when there are multiple equilibria, the probability of bank runs will be higher for lower values of R, and indeed the probability of observing a bank run will be the highest (equal to one) when the fundamentals deteriorate sufficiently (that is, when $R < R^*$). In this sense, even the traditional sunspot approach can account for some of the correlation of bank runs with economic fundamentals. Economic fundamentals determine whether multiple equilibria exist, and then probabilities have to adjust to reflect this fact.[10]

Furthermore, the traditional sunspot approach seems too simplistic for this environment, and the risk-dominance-based selection mechanism appears to be a reasonable extension. We can think that the risk dominance ESM is the case where the particular sunspot variable that coordinates patient agents to run to the bank is correlated, in a specific way, with the stochastic variable R determining fundamentals. Risk dominance provides discipline and intuition to this correlation.

In particular, the risk dominance criterion divides the support of the distribution of R into two sets: the set where $R < \widehat{R}$, in which the run equilibrium is risk dominant, and the set where $R > \widehat{R}$, in which the no-run equilibrium is risk dominant. We can think that there is an associated sunspot random

[10] Ironically, the model in the second part of the paper by Allen and Gale (1998) can be used to provide a good example of this situation. For some parameter values their model has multiple equilibria. Their equilibrium analysis delineates three relevant regions for the possible realization of the return on the risky asset R. When R is very low, the equilibrium has a bank run; when R is very high, there are no bank runs in equilibrium; and for intermediate values of R, there are multiple equilibria: both having a bank run and not having a bank run are possible equilibrium outcomes. Therefore, just using a simple sunspot variable to determine which of the two equilibria will be observed in the intermediate region of R would deliver the historical correlation: as fundamentals deteriorate (as R goes from high to low levels), the probability of bank runs first goes from zero to positive (the value associated with the sunspot) and then to unity when fundamentals are so poor that a bank run is unavoidable.

variable s, perfectly correlated with R, such that whenever R takes values in the interval $[1, \widehat{R}]$, the variable s takes the value r, and whenever R takes values in the interval $(\widehat{R}, \infty)$, the sunspot variable s equals n. If agents associate values of $s = r$ with a run situation and values of $s = n$ with a no-run situation, the equilibrium selection process is still driven by sunspots (the variable s), but it generates a correlation of bank runs with the behavior of fundamentals. It is worth noting that for most values of R, both equilibria still exist, even though one of them is risk dominant. What determines which equilibrium will be played is a matter of how agents get coordinated. Coordination is driven by the sunspot variable s. Risk dominance can be thought of as the justification for why the particular sunspot random variable s has been selected as a coordination device over all possible variables that may be available. Note that there is a higher level of coordination among agents in the choice of the relevant sunspot variable. This interpretation of sunspots is in fact associated with another argument that has been used to explain the appearance of such coordination devices: sunspots can be viewed as the limiting case of situations in which agents are overreacting to some small movement in economic fundamentals. Manuelli and Peck (1992) formalize this argument.

Finally, it should be clear at this point that the more general ESM approach (Ennis and Keister [2003a]), in which the probability of a bank run π is a decreasing function of R, is also consistent with both the multiplicity of equilibria and the correlation of bank runs with economic fundamentals. In fact, with this approach the probability π can be strictly between zero and unity and at the same time be dependent on R. This feature seems very appealing, since the historical correlation was never perfect: sometimes bank runs did not occur even though economic fundamentals were as bad as or worse than in periods where a bank run did occur (see Gorton [1988]).

4. THE BANK'S PROBLEM

In Section 2 we assumed that agents would be willing to deposit their funds in the bank and that the bank would choose a contract with some specific properties. This section provides the justification for those assumptions.

Given that the banking system may be subject to runs, agents might choose not to participate in the banking system.[11] In that case, their payoff would be given by the following "autarky" problem

$$V_A \equiv \max_{\eta} \int \left(u \frac{(\eta + x(1-\eta))^{\gamma}}{\gamma} + (1-u) \frac{(\eta + R(1-\eta))^{\gamma}}{\gamma} \right) f(R)\, dR,$$

[11] For the sake of simplicity, I am restricting agents to deposit either *all* their resources in the bank or nothing at all. Ennis and Keister (2003b) consider the case where agents can deposit just part of their initial resources in the bank. This is an important extension in environments where bank runs can happen with positive probability, as is the case in this paper.

subject to $0 \leq \eta \leq 1$. At the beginning of period 1, the agent decides how to split the endowment between storage (η) and investment ($1 - \eta$). At the end of period 1, the agent finds out whether she is patient or impatient. If she is impatient, then she liquidates the investment and consumes (funds are useless for her in the second period). If she is patient, then she stores the liquid funds and consumes in the second period both the liquid funds and the return on investment (recall that we are assuming that $R > 1 > x$).

A bank could always choose a contract that eliminates the possibility of experiencing a run. I will call the best contract with such property the "run-proof contract." A contract is run-proof if there is enough liquidity in the bank to pay all agents the amount a at the end of period 1. But because the contract is run-proof, patient agents actually wait until the second period to withdraw. The problem of a bank choosing the run-proof contract is the following:

$$V_{RP} \equiv \max_{a,\eta} \int \left(u \frac{a^{\gamma}}{\gamma} + (1-u) \frac{1}{\gamma} \left(\frac{R(1-\eta) + (\eta - ua)}{1-u} \right)^{\gamma} \right) f(R)\, dR,$$

subject to

$$a \leq \eta + x(1-\eta),\ a \geq 0, \text{ and } 0 \leq \eta \leq 1.$$

The first constraint is the run-proof constraint. It says that even if all agents go to the bank in the first period (i.e., early), the bank will not run out of funds.

Finally, after having studied equilibrium selection in the post-deposit game, we are now in a position to write down the problem faced by the bank at the beginning of period 1. It is important to note that the probability of a run may depend on the contract chosen by the bank and hence the bank will take this effect into account when determining the best possible contract. Formally, the bank's problem is given by

$$\begin{aligned} V \equiv \max_{a,\eta} \int & [\pi(R,a,n) P_{rr}(a,\eta) + \\ & (1 - \pi(R,a,\eta)) \left(u \frac{a^{\gamma}}{\gamma} + (1-u) P_{nn}(R,a,\eta) \right)] f(R)\, dR, \end{aligned}$$

subject to $a \geq 0$ and $0 \leq \eta \leq 1$. Note that P_{rn} does not enter the problem directly. It may, however, enter the problem indirectly through the determination of $\pi(R,a,n)$, as in the case of the ESM based in risk dominance or adaptive learning.

When we have $V_A < \max\{V, V_{RP}\}$, the agents will choose to deposit their funds at the bank. When we have $V > V_{RP}$, the bank will choose the contract that allows for the possibility of bank runs according to the ESM that is operating in the economy (that is, according to the given function $\pi(R,a,\eta)$). It is important to note that if there exist values of R such that $R < R^*$ and $f(R) > 0$, then for those values of R we must have that $\pi(R,a,\eta) = 1$ because the post-deposit game has a unique (run) equilibrium for those values of R.

Diamond and Dybvig (1983) show that when the return on investment R is not stochastic (and greater than unity) and the probability π is arbitrarily set at zero, the bank chooses a contract (a, η) for which a bank run is a possible equilibrium of the post-deposit game played by the patient agents. Hence, using arguments of continuity, it can be shown that there exist functions $f(R)$ and $\pi(R, a, \eta) > 0$ such that a bank solving the problem V described above will also choose a contract that admits runs (that is, a contract such that $\eta + x(1 - \eta) < a$ holds).

5. CONCLUSION

I have shown that even when bank runs are driven by self-fulfilling expectations in environments with multiple equilibria, the historical correlation of bank runs with poor economic fundamentals can still be accounted for. More evidence would be necessary to reject the case of bank runs originating in situations with multiple equilibria. For now, when we observe a bank run, we cannot in principle confidently discard the possibility that another equilibrium with no bank run was also possible. This conclusion is important from a policy standpoint. In some cases, multiple-equilibria bank runs can be avoided by the design of off-equilibrium policies that are hence never observed. For example, the suspension of convertibility could make the run situation I have presented no longer an equilibrium of the post-deposit game (as proposed by Diamond and Dybvig in their original paper). But because suspension would occur only when there is a run and runs are not equilibrium outcomes anymore, the suspension of payments will not be observed. An important qualification is that, like many other off-equilibrium threats, this policy entails a certain ability of the bank to commit to actually implementing the policy if it becomes necessary.

There is another important policy implication of the ideas presented here. In the multiple-equilibria case, bank runs are usually not optimal and in general the policymaker would like to avoid them (or at least lower their probability). Contrary to this position, Allen and Gale (1998) present the case of bank runs that are not the consequence of a coordination failure and that are in fact part of the optimal arrangement for risk sharing in the economy. The policymaker would not want to avoid the Allen-Gale type of bank runs. Determining which of the two cases is driving a particular episode is an important issue that the policymaker would need to carefully evaluate.

REFERENCES

Allen, Franklin, and Douglas Gale. 1998. "Optimal Financial Crises." *Journal of Finance* 53 (August): 1245–84.

Calomiris, Charles, and Joseph Mason. 1997. "Contagion and Bank Failures During the Great Depression: The June 1932 Chicago Banking Panic." *American Economic Review* 87 (December): 863–83.

Cass, David, and Karl Shell. 1983. "Do Sunspots Matter?" *Journal of Political Economy* 91 (April): 193–227.

Cooper, Russell. 1999. *Coordination Games*. Cambridge: Cambridge University Press.

_____, and T. W. Ross. 1998. "Bank Runs: Liquidity Costs and Investment Decisions." *Journal of Monetary Economics* 41 (February): 27–38.

Diamond, Douglas, and Philip Dybvig. 1983. "Bank Runs, Deposit Insurance, and Liquidity." *Journal of Political Economy* 91 (June): 401–19.

Ennis, Huberto, and Todd Keister. 2003a. "Government Policy and the Probability of Coordination Failures." Centro de Investigación Económica Discussion Paper 03-01, ITAM (February).

_____. 2003b. "Economic Growth, Liquidity, and Bank Runs." Forthcoming in the *Journal of Economic Theory*.

Gale, Douglas, and Xavier Vives. 2002. "Dollarization, Bailouts, and the Stability of the Banking System." *Quarterly Journal of Economics* (May): 467–502.

Gorton, Gary. 1988. "Banking Panics and Business Cycles." *Oxford Economic Papers* 40 (December): 751–81.

Harsanyi, John, and Reinhard Selten. 1988. *A General Theory of Equilibrium Selection in Games*. Cambridge: MIT Press.

Jevons, W. Stanley. 1884. *Investigation in Currency and Finance*. London: Macmillan.

Kandori, Michihiro, George Mailath, and Rafael Rob. 1993. "Learning, Mutation, and Long Run Equilibria in Games." *Econometrica* 61 (January): 29–56.

Manuelli, Rodolfo, and James Peck. 1992. "Sunspot-Like Effects of Random Endowments." *Journal of Economic Dynamics and Control* 16 (April): 193–206.

Matsui, Akihiko, and Kiminori Matsuyama. 1995. "An Approach to Equilibrium Selection." *Journal of Economic Theory* 65 (April): 415–34.

Miron, Jeffrey. 1986. "Financial Panics, the Seasonality of the Nominal Interest Rate, and the Founding of the Fed." *American Economic Review* 76 (March): 125–40.

Morris, Stephen, Rafael Rob, and Hyun Song Shin. 1995. "p-Dominance and Belief Potential." *Econometrica* 63 (January): 145–57.

Peck, James, and Karl Shell. 2003. "Equilibrium Bank Runs." *Journal of Political Economy* 111 (February): 103–23.

Saunders, Anthony, and Berry Wilson. 1996. "Contagious Bank Runs: Evidence from the 1929–1933 Period." *Journal of Financial Intermediation* 5 (October): 409–23.

Schumacher, Liliana. 2000. "Bank Runs and Currency Run in a System Without a Safety Net: Argentina and the 'Tequila' Shock." *Journal of Monetary Economics* 46 (August): 257–77.

Shell, Karl, and Bruce Smith. 1992. "Sunspot Equilibrium." In *The New Palgrave Dictionary of Money and Finance*. Vol. 3. New York: Macmillan, 601–7.

Temzelides, Ted. 1997. "Evolution, Coordination, and Banking Panics." *Journal of Monetary Economics* 40 (September): 163–83.

Wallace, Neil. 1988. "Another Attempt to Explain an Illiquid Banking System: The Diamond and Dybvig Model with Sequential Service Taken Seriously." Federal Reserve Bank of Minneapolis *Quarterly Review* 12 (Fall): 3–16.

Williamson, Stephen. 1986. "Costly Monitoring, Financial Intermediation, and Equilibrium Credit Rationing." *Journal of Monetary Economics* 18 (September): 159–79.

Young, Peyton. 1998. *Individual Strategy and Social Structure*. Princeton, N.J.: Princeton University Press.