Federal Reserve Bank of Chicago

Call for Papers

2003 Conference on Bank Structure and Competition

Fourth Quarter 2002

Economic perspectives

Economic perspectives

Economic Perspectives is published by the Research Department of the Federal Reserve Bank of Chicago. The views expressed are the authors' and do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.

Single-copy subscriptions are available free of charge. Please send requests for single- and multiple-copy subscriptions, back issues, and address changes to the Public Information Center, Federal Reserve Bank of Chicago, P.O. Box 834, Chicago, Illinois 60690-0834, telephone 312-322-5111 or fax 312-322-5515.

Economic Perspectives and other Bank publications are available on the World Wide Web at http://www.frbchi.org.

chicagofed.org

ISSN 0164-0682

Contents

Fourth Quarter 2002, Volume XXVI, Issue 4

The challenges facing community banks: In their own words

Robert DeYoung and Denise Duffy

Introduction and summary

When economists analyze an industry, they typically do so at arms length, using a combination of theoretical models and large amounts of statistical data. The theoretical models describe the interplay between the structure of the industry and the competitive behavior of the firms that populate the industry. The statistical data—which may include financial ratios, industry trends, and peer group comparisons—serve to personalize the sterile, one-size-fits-all nature of the theoretical models. But most industry studies never get especially close to the people most responsible for the industry data: the managers and owners who make long-run strategic plans that shape the data, who make short-run competitive decisions in response to the data, and whose careers and companies are ultimately defined by the data.

In this article, we analyze the U.S. community banking sector—a sector populated by small firms that hold a shrinking share of an increasingly competitive and technology-based financial services industry—but we rely on an atypical approach to perform the analysis. We use numerous first-hand observations made by individual community bankers, collected during a Federal Reserve survey in August 2001 (Federal Reserve System, 2002), to complement the usual data-intensive industry analysis. Although the survey itself was an effort to learn about the evolving payments services needs of community banks, the surveyed bankers also made wide-ranging observations on a variety of other topics, including the fundamental mission of community banks; the threats and opportunities posed by large banks; perceptions that the playing field is not always level; and the growing tension between traditional high-touch relationship banking and potentially more efficient high-tech banking.

Augmenting systematic industry data with bankers' anecdotal observations humanizes our analysis. The bankers tended to be more optimistic about the future viability of the community banking business model than many industry observers and, not surprisingly, they tended to be less sanguine about the regulatory and technological changes that have increased the competitive pressures on community banks. But aside from these and a few other differences, the assessments of the two groups were quite consistent—despite being stated from different perspectives and arrived at using different (and, in the case of the bankers, implicit) analytic frameworks. The consensus view is that industry consolidation and technological change are providing opportunities as well as posing threats for community banks; that community banks can profitably coexist with large multi-state banks in the future; but, to do so, community banks must be efficiently operated, well-managed, and must continue to innovate.

Forces of change

The past decade has witnessed tremendous changes in how banks are regulated, how they use technology to produce financial services, and how they compete with each other. These transformations have important consequences for the typical community bank, for the community banking sector as a whole, and by extension for the households and small businesses that purchase financial services from community banks.

Robert DeYoung is a senior economist and economic advisor in the Economic Research Department and Denise Duffy is an economic capital specialist in the Global Supervision and Regulation unit at the Federal Reserve Bank of Chicago.
The authors wish to thank Carol Clark, Zoriana Kurzeja, and David Marshall for helpful comments and suggestions.

Geographic deregulation

The McFadden Act of 1927 restricted U.S. commercial banks from branching across state borders. In addition, most state governments have historically restricted bank branching within state borders. These restrictions reduced the efficiency of the U.S. banking system by artificially limiting the size of commercial banks. But state governments began to gradually relax their geographic branching restrictions beginning in the mid-1970s, and by 1994 the federal government had passed the Riegle–Neal Act which eliminated virtually all prohibitions against interstate banking in the U.S. Both large and small banking companies have taken advantage of geographic deregulation by acquiring banks in other counties, states, or regions. Growth via acquisition is a fast way to expand into a new geographic market, because the expanding bank can begin its operations in the new market with an established physical presence and an established customer base.

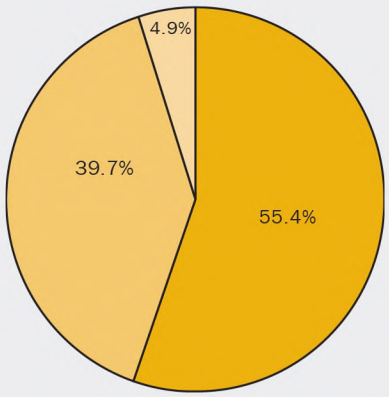
The most visible evidence of these geographic-expansion mergers is the substantial reduction in the number of community banks in the U.S. As shown in figure 1, over half of all U.S. bank mergers since 1985 have combined two community banks (defined here as having less than $1 billion in assets), and in most of the remaining mergers a larger bank has acquired a community bank.[1] Figure 2 illustrates the dramatic change in the size distribution of U.S. commercial banks caused by these mergers. The number of small community banks (less than $500 million in assets) has nearly halved since 1985, while the numbers of large community banks ($500 million to $1 billion), mid-sized banks ($1 billion to $10 billion), and large banks have remained relatively constant.

Perhaps the primary motivation for community banks to merge is to capture scale economies, reductions in per unit costs or increases in per unit revenues that occur as small banks grow larger.[2] By growing larger via merger, a community bank can make loans to bigger firms; offer a broader array of products and services; attract and retain higher quality managers; diversify away some of its riskiness by lending into new geographic markets; generate network benefits from integrating systems of branches and ATMs (automated teller machines) in different geographic areas; gain access to new sources of capital; or operate its branch offices and computer systems closer to full capacity. Another motivation for community banks to merge is to become large relative to the local market: A combination of two community banks that operate in the same small towns may increase their pricing power in those towns. But increased size can also have a downside: A community bank that grows too large, too geographically spread out, or otherwise too complex may become unable to deliver the same level of personalized service that attracted many of its business and retail customers in the first place.

Market-extension mergers have approximately doubled the geographic reach of the typical U.S. bank holding company over the past two decades. The average bank holding company affiliate with more than $100 million in assets was located about 160 miles from its holding company headquarters in 1985; by 1998 this distance had increased to about 300 miles (Berger and DeYoung, 2001). But as banking companies have used mergers to arc across geographic boundaries, the structure of local banking markets has changed very little. Since 1980, the nationwide share of deposits held by the ten largest U.S. banks has doubled from about 20 percent to about 40 percent, but there has been little upward trend in concentration in local banking markets (DeYoung, 1999). As a result, the bank merger wave is unlikely to have resulted in a systematic increase in local market power. On the contrary, recent studies suggest that the merger wave has intensified

FIGURE 1

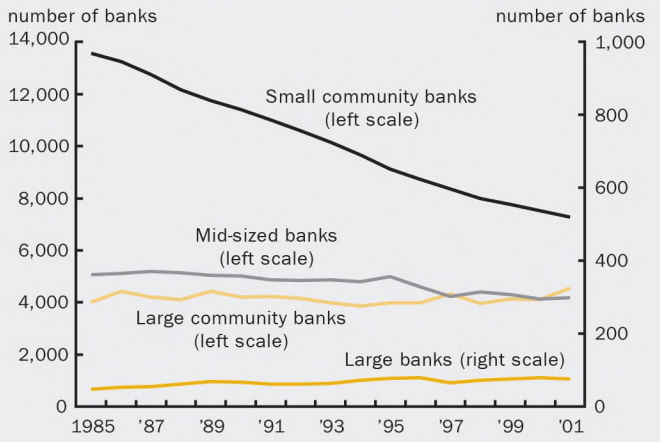
Breakdown of commercial bank mergers and acquisitions, 1985–99



Note: Large banks have over $10 billion in assets. Mid-sized banks have between $1 billion and $10 billion in assets. Community banks have less than $1 billion in assets. All figures are in 1999 dollars.
Source: Authors' calculations using Federal Reserve data.

FIGURE 2

Size distribution of U.S. commercial banks, 1985–2001



Notes: Large banks have over $10 billion in assets. Mid-sized banks have between $1 billion and $10 billion in assets. Large community banks have between $500 million and $1 billion in assets. Small community banks have less than $500 million in assets. Assets are in 1999 dollars.

Source: Authors' calculations using call reports.

competition among banks in local markets: Banks tend to operate at higher levels of efficiency after one of their local competitors is acquired by an out-of-market bank.[3]

Product market deregulation

Deregulation has also broadened the scope of financial services that banks are permitted to offer their customers. The Gramm–Leach–Bliley Act of 2000 ended or greatly relaxed restrictions that for decades had limited the financial activities of commercial banks; the most famous of these restrictions was the Glass–Steagal Act of 1933, which prohibited commercial banks from engaging in investment banking. Commercial banking companies are now permitted to produce, market, and distribute a full range of financial services, enveloping the previously separate areas of commercial banking, merchant banking, securities brokerage and underwriting, and insurance sales and underwriting.[4]

Product market deregulation has had a subtler impact on community banks than geographic deregulation. Community banks have traditionally offered a limited array of banking products, generating interest income from loans and investments and generating a limited amount of noninterest income (service charges) from deposit accounts. Larger commercial banks offer these traditional interest-based banking services as well, but they also sell a variety of additional financial services that generate fees and noninterest income. Large banks are more likely to securitize their loans; they collect little interest income because these loans are not held for long on their books, but collect potentially large amounts of noninterest income from originating and servicing these loans. Large banks often write back-up lines of credit for their large business customers; they receive fees for this service but receive interest income only in the rare case that the client draws on the credit line. Large banks can generate large amounts of noninterest income by charging third-party access fees at their widespread ATM networks. And, compared with community banks, large banks tend to charge high fees to their own depositors.[5]

Figure 3 shows that noninterest income accounts for a relatively small percentage of community bank revenue and has increased slowly over time relative to its growth at larger banks. This suggests a growing differentiation between the business strategies of small community banks and larger commercial banks. Whether community banks can continue to be profitable by offering a relatively narrow range of services, while their largest rivals are becoming "financial supermarkets," is an important question for determining the future size and viability of the community banking sector.

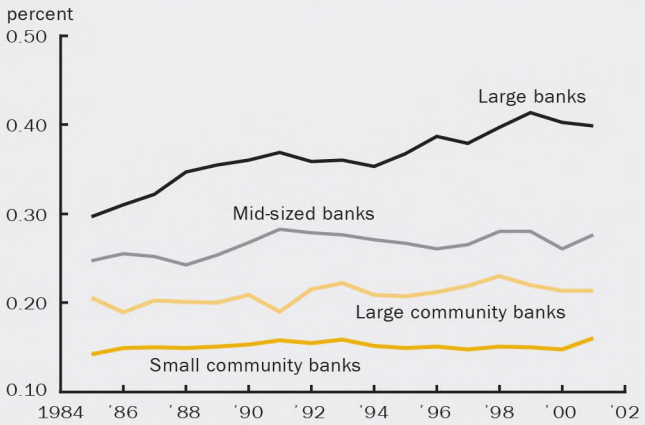
New technologies

Like deregulation, advances in information, communications, and financial technologies over the past two decades have increased the competitive pressures on commercial banks. For example, mutual funds, online brokerage accounts, and money market funds have provided attractive investment options for depositors; as a result, core deposits have become less available for all size classes of banks.[6] Because community banks have fewer non-deposit funding options than large banks (for example, small banks typically do not have access to bond financing), it costs them more to attract and retain core deposits.[7] New financial instruments, combined with improved information about borrower creditworthiness, have intensified competition on the asset side of banks' balance sheets. Commercial paper has become an attractive alternative to short-term bank loans for large, highly rated business borrowers, and junk bond financing has become an alternative to long-term bank loans for riskier business borrowers.

In some cases, banks have been able to fight back by deploying new financial technologies of their own. Virtually all banks are using ATMs—and an increasing number are using transactional Internet websites—to

FIGURE 3

Noninterest income as a percentage of net revenue, U.S. commercial banks, 1985–2001



Notes: Large banks have over $10 billion in assets. Mid-sized banks have between $1 billion and $10 billion in assets. Large community banks have between $500 million and $1 billion in assets. Small community banks have less than $500 million in assets. Assets are in 1999 dollars.

Source: Authors' calculations using call reports.

offer increased convenience to their depositors. Many banks offer sweep accounts and proprietary mutual funds to limit the number of small business and retail customer defections to nonbank competitors. And as discussed above, some banks have reoriented their business mix toward off-balance-sheet activities like back-up lines of credit, so they can continue to earn revenues from business customers that switched from loan financing to commercial paper financing.

Technology has also allowed banks to fundamentally change the way they produce financial services. Securitized lending is a prime example. By bundling and selling off their loans rather than holding them on their balance sheets, banks can economize on increasingly scarce deposit funding while simultaneously generating increased fee income. Securitized lending operations exhibit deep economies of scale, so banks that originate and securitize large amounts of loans can operate at low unit costs. As a result, the cost savings and increased revenues generated by securitized lending are generally not available to small banks. However, a securitized lending strategy can limit the strategic options of a large bank. Securitization only works for standardized loans like credit cards, auto loans, or mortgage loans—"transactions" loans that can be underwritten based on a limited amount of "hard" financial information about the borrower that can be fed into an automated credit-scoring program.[8]

Securitized bundles of transactions loans share many of the same characteristics as commodities: They are standardized products, easily replicable by other large banks, and they are bought and sold in competitive markets. As a result, securitized lending is a high-volume, low-cost line of business in which monopoly profits are unlikely. In contrast, "relationship" lending requires banks to collect a large amount of specialized "soft" information about the borrower in order to ascertain her creditworthiness. The classic example of relationship lending is the small business loan made by community banks. The uniqueness of these lending relationships gives banks some bargaining power over borrowers, which supports a relatively high profit margin.

Internet website technology is relatively inexpensive, so both large banks and community banks can theoretically use the Web to do business in local markets anywhere in the nation. But in reality, community banks face a disadvantage at using this new technology. First, small banks often do not have a large enough customer base to efficiently utilize this delivery channel.[9] Moreover, profitable entry into a new market is not just a technological feat, but also a marketing feat. Getting noticed in a new market generally requires expensive advertising; getting noticed on the World Wide Web is even more difficult, and requires substantial advertising expenditures beyond the resources of the typical community bank. One way that banks have attracted customers' attention on the Web is by offering above-market rates on certificates of deposit, so that the bank's name gets posted on financial websites that list high-rate payers. But this strategy is itself a costly substitute for advertising, and usually attracts one-time sources of funds that do not develop into long-lasting relationship clients.[10]

Implications of these changes for community banks

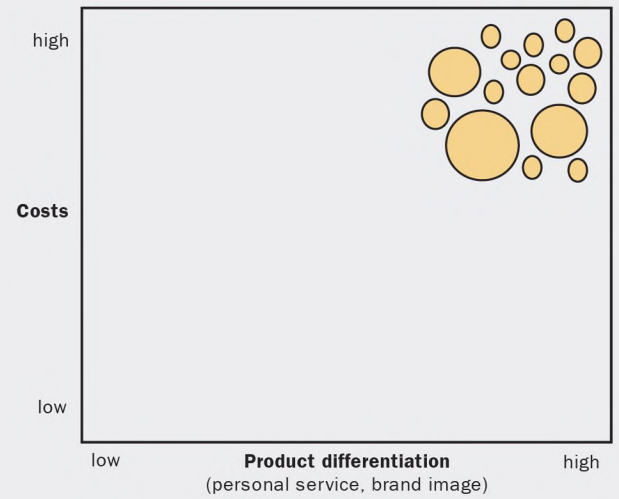
Many of these developments appear to favor large banks at the expense of small local banks. However, some have argued that well-managed community banks may be able to turn these competitive threats into opportunities. One case in point concerns the market for small business loans, a prime product line for small community banks.[11] The idiosyncratic nature of small business relationship lending is in many ways inconsistent with automated lending technology. Thus, when a large bank shifts toward an automated lending culture, traditional community banks may stand to pick

up profitable small business accounts. Similarly, the movement of large banks toward charging explicit (and often higher) fees for separate depositor services may provide an opportunity for community banks to attract relationship-based deposit customers who prefer bundled pricing.

DeYoung and Hunter (2003) argue that the banking industry will continue to feature both large global banks and small local banks. They illustrate this argument using the strategic maps in figures 4 and 5. The maps are highly stylized depictions of three fundamental structural, economic, and strategic variables in the banking industry: bank size, unit costs, and product differentiation. The vertical dimension in these maps measures the unit costs of producing retail and small business banking services. The horizontal dimension measures the degree to which banks differentiate their products and services from those of their closest competitors. This could be either actual product differentiation (for example, customized products or person-to-person service) or perceived differentiation (for example, brand image). For credit-based products, this distinction may correspond to automated lending based on "hard" information (standardization) versus relationship lending based on "soft" information (customization). In this framework, banks select their business strategies by combining a high or low level of unit costs with a high or low degree of product differentiation. The positions of the circles indicate the business strategies selected by banks, and the relative sizes of the circles indicate the relative sizes of the banks.

FIGURE 4

Strategic map of U.S. banking industry, pre-deregulation period



Source: DeYoung and Hunter (2003).

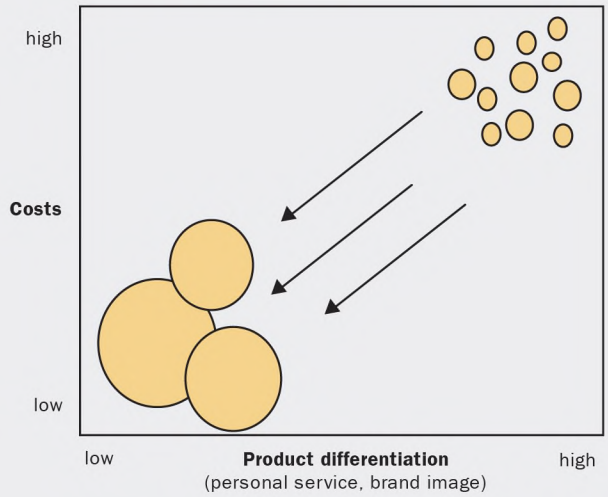
Figure 4 shows the banking industry prior to deregulation and technological change. Banks were clustered near the northeast corner of the strategy space. The production, distribution, and quality of retail and small business banking products were fairly similar across banks of all sizes. Small banks tended to offer a higher degree of person-to-person interaction, but this wasn't so much a strategic consideration as it was a reflection that delivering high-touch personal service becomes more difficult as an organization grows larger. Large banks tended to service the larger commercial accounts, but bank size often wasn't a strategic choice; the economic size of the local market and state branching rules often placed limits on bank size.

Deregulation, increased competition, and new financial technologies created incentives for large banks and small banks to become less alike. Large banks began to get larger, at first due to modest within-market mergers, and then more rapidly due to market-extension megamergers. Increases in bank size yielded economies of scale, and unit costs fell.[12] Increased scale also gave these growing banks access to the new production and distribution technologies discussed above, like automated underwriting, securitization of loans, and widespread ATM networks. These technologies reduced unit costs even further at large banks, but in many cases gradually altered the nature of their retail business toward a high-volume, low-cost, and less personal "financial commodity" strategy.

The combined effects of these changes effectively drove a strategic wedge between the rapidly growing large banks on one hand and the smaller community banks on the other hand. The result is shown in figure 5. Large banks have moved toward the southwest corner of the strategy space, sacrificing personalized service for large scale, a more standardized product mix, and lower unit costs. This allows large banks to charge low prices and still earn a satisfactory rate of return. Although many community banks have also grown larger via mergers, they remain relatively small and have continued to occupy the same strategic ground, providing differentiated products and personalized service. This allows small banks to charge a high enough price to earn a satisfactory rate of return, despite low volumes and unexploited scale economies.[13] In the following section, we consider these trends from the

FIGURE 5

Strategic map of U.S. banking industry, post-deregulation transition



Source: DeYoung and Hunter (2003).

community bankers' point of view, based on the results of the August 2001 Federal Reserve survey.

The survey

In August 2001, the Federal Reserve System's Customer Relations and Support Office (CRSO), located at the Federal Reserve Bank of Chicago, conducted a series of interviews with officers and employees of ten community banks from across the U.S. These interviews covered a wide range of topics, and the interviewers encouraged respondents to include a large amount of detail in their answers. These interviews represent the first stage of an ongoing Federal Reserve effort to better understand the business strategies community banks are implementing to remain viable in a changing banking environment and to determine what community banks require from the payments system in order to survive in this environment. A secondary goal of the study is to stimulate research and public policy interest regarding the community bank sector.

The ten surveyed community banks were not selected using a statistically valid sampling technique, and in any event this sample of banks is too small to use for statistical inference testing. Rather, these banks were selected based on knowledge that Federal Reserve Business Development staff had accumulated about them over time. The ten banks share two important traits. First, each of their business models was based on the concept of community banking. Second, based on previous contact with these firms, Fed Business Development staff had reason to expect that the officers and employees of these organizations would answer the survey questions in an open and forthcoming manner. In addition, these ten banks were selected so that the sample, though small, was heterogeneous in terms of bank size, bank location, and other organizational characteristics.

The banks were selected from across the country, from urban, suburban, and rural areas, and from three *ad hoc* size tiers: less than $50 million in assets, between $50 million and $200 million in assets, and between $200 million and $1 billion in assets. Two of the banks are de novo (newly chartered) banks; two are minority-owned banks; one has a primarily commercial customer base (as opposed to the traditional community bank mix of commercial and retail customers); three have a bilingual/ethnic customer base; and three provide services to customers whose banking transactions sometimes involve foreign countries, including Canada, Mexico, and Pacific Rim countries. Table 1 summarizes the characteristics of the surveyed banks.

The major decision makers and policymakers at each bank participated in the interviews. This typically included the bank's chief executive officer (CEO), chief financial officer (CFO), chief operations officer (COO), and cashier, as well as a branch manager and a lending officer. Participants were asked a series of questions regarding their bank's business strategy, product offerings, operations, and purchases of payments and other financial services during the past three years, as well as projections for the next three years. Participants were specifically asked to discuss how their community bank was positioning itself to survive in a rapidly changing financial services environment. A representative list of questions is presented in box 1.

Below, we present a selection of responses from the community bankers that best reflect the challenges and issues facing the community banking sector. A full summary of the results can be read in the Federal Reserve System's (2002) *Community Bank Study*.

Mergers taketh away—but mergers giveth, too

As discussed earlier, the number of community banks in the U.S. has plummeted over the past two decades. This is partly because large banks gobbled up small banks in the process of building regional and national networks—but it is also because large

TABLE 1

Characteristics of community banks in the survey

Asset tier	Market type	Location	No. of branches	Other
3	Urban	Southeast	3	
3	Urban	Northwest	4	Minority owned and operated
3	Rural/small town	Midwest	0	A "bankers' bank"
3	Rural/small town	Mid-South	18	
2	Urban	South	7	Minority owned and operated
2	Suburban	West Coast	3	Recently chartered
2	Rural/small town	Midwest	2	
1	Suburban	East Coast	0	Savings and loan
1	Rural/small town	Midwest	0	Recently chartered
1	Rural/small town	Southwest	0	Serves a bilingual population

Note: Banks in asset tiers 1, 2, and 3, respectively, have less than $50 million in assets, between $50 and $200 million in assets, and between $200 million and $1 billion in assets.

BOX 1

Community bank survey topics

- What current and expected future strategic initiatives will position your institution for profitable growth?
- Does your institution face potential challenges in implementing these strategic initiatives?
- Which customer segments will you target with these initiatives?
- What is your current and expected future product mix?
- Please describe the relationship between strategic importance and ease of offering the various products and services mentioned above.
- Which customer segments are most profitable?
- Which profitable customer segments have you recently lost to competitors?
- Which of your customers' business concerns are not adequately addressed in the financial marketplace?
- What are the competitive factors that affect the community bank sector?
- Please forecast the potential impact of current or impending regulations on your institution.
- Do you use strategic alliances? If so, in what ways?
- Do you use third-party processors? If so, in what ways?
- Which payments system services do you use? Which services do you plan to use in the future?

TABLE 2

Option and swap positions at U.S. commercial banks, year-end 2001

	Banks with positions	% of banks with positions	% total underlying notional value[a]
		Options	
Small community banks	39	0.53	0.01
Large community banks	16	4.92	0.03
Mid-sized banks	42	13.46	0.21
Large banks	54	69.23	99.74
		Swaps	
Small community banks	48	0.65	0.00
Large community banks	19	5.85	0.00
Mid-sized banks	88	28.21	0.12
Large banks	67	85.90	99.87

[a]Percentage of the total notional value underlying the derivatives contracts held by commercial banks.
Source: Call reports.

community banks acquired small community banks, and because small community banks merged with each other. Still, community bankers tend to focus on the competitive threat posed by large, acquisitive, out-of-state banking companies:

- *"Community banks aren't necessarily stealing customers from other community banks; larger banks are stealing customers from community banks."*

There is certainly some truth to this "David versus Goliath" point of view. In some lines of business—like mortgage banking and credit card lending—large banks have increased their market share substantially at the expense of small banks. But community banks sometimes experience increased demand in other lines of business—like household deposits and small business relationships—after large banks enter the local market due to differences in service quality, as the following responses suggest:

- *"With all these mergers, the personal service level isn't what people in small towns are used to. Big banks [from out of state] buy small banks and sell them off, because bankers in Minnesota don't know what the economy is like in Texas."*
- *"Most of our competitors are so big—the First Unions, the Commerce Banks—they're offering services in a different (impersonal) way. They're driving their customers away, and we're more than happy to take care of them."*

There is plenty of anecdotal evidence that supports these statements.[14] The $9.5 billion Roslyn Savings Bank recently reported that 15 percent of its new deposits were coming from former depositors of Dime Savings Bank, who were unhappy about changes made to their passbook savings accounts after Dime was acquired by the $275 billion thrift Washington Mutual. In the 12 months after NationsBank acquired Boatmen's Bancshares in 1997, community bank Allegiant Bancorp of St. Louis grew by $100 million, nearly a 20 percent increase in assets. And in the wake of its merger with First Interstate Corp, Wells Fargo faced a 15.5 percent reduction in deposits. These anecdotes are consistent with recent studies of de novo bank entry, which tend to find that new commercial banks are more likely to start up in local markets that have recently experienced entry (via merger or acquisition) by a large, out-of-state banking company (Berger, Bonime, Goldberg, and White, 1999; Keeton, 2000). The presumption is that new banks are starting up in these markets because they contain a substantial number of disgruntled customers of the acquired bank who are shopping for a new banking relationship.

What is it that attracts these disgruntled customers to community banks? Nearly all of the surveyed bankers identify the local focus of community banks as an important competitive advantage:

- *"We can't out-research and develop them, and we can't out-produce them. But we can have more and better knowledge of the personal situations and financial problems that we're trying to solve."*
- *"We're known and we're local. If you have the local connection, and I think a local bank has that better than anybody, then you have a foot up. You're going to have more credibility with your local people."*

Strategies and production functions

The strategic analysis in figures 4 and 5 juxtaposed community banks and large banks in a number of ways: small versus large, personal versus impersonal, high cost versus low cost. The common thread that connects each of these juxtapositions is the bank production function—that is, the methods and techniques that banks use to produce financial products and services. According to the analysis, if a bank uses a production process that includes automated credit-scoring models, moving loans off its books via asset securitization, and a widespread distribution network (branch offices, ATMs, and Internet kiosks), it will likely become a large bank, operate with relatively low unit costs (due to scale economies), and produce relatively standardized financial products. In contrast, if a bank uses a production process that includes personal contact with customers, portfolio lending, and a local geographic focus, it will likely become a small bank, operate with relatively high unit costs, and produce more customized financial services.

The community bankers that participated in the survey did not make explicit references to production functions or related concepts. But implicit in many of their remarks was the understanding that there are differences between large and small bank production functions, and that these differences cause challenges for community banks. For example, one banker stressed that the size deficit between community banks and their larger competitors has important cost implications for the type of financial services he produces and the prices that he charges for them:

- *"It's a volume-driven business [offering residential loans], and we can't compete with the larger banks and mortgage companies, because volume drives rates down. We offer it as a customer service ... but these loans aren't a big part of our portfolio."*

Indeed, economic research confirms that automated mortgage underwriting and servicing procedures have generated huge cost reductions at specialized mortgage banks and have allowed them to quickly become some of the biggest players in home mortgage markets. Rossi (1998) reported that mortgage banks were originating over 50 percent of all one-to-four-family mortgages in the U.S. in 1994, a spectacular increase from the 20 percent market share that they held just five years earlier. Rossi also estimated a series of best-practices production (cost) functions for mortgage banks and used them to illustrate some clear links between bank size and bank costs: Unit costs equaled about 1 percent of assets for the smallest quartile of mortgage banks, but fell to just 0.25 percent of assets for the largest mortgage banks. Cost advantages like these allow large mortgage banks to price below small, full-service community banks, as this comment confirms:

- *"Regional banks came in priced about 150 basis points below our market for a 15-year fixed term loan—we did lose about $10 million for that. Our strategy as a bank is not to fix for 15 years. Five years is our threshold. We still remember the 1970s when the rates went up and banks got in trouble with fixed rates."*

How can large banks offer these loans at terms that community banks find unprofitable? Large banks can write mortgage loans and consumer loans in volumes large enough to exploit the scale economies associated with automated lending processes (that is, credit scoring and securitization). Some of these savings can be passed along to the consumer. Furthermore, large banks are better able to manage the interest rate risk associated with long-term, fixed rate loans by using financial derivatives contracts. For example, banks that issue fixed-rate loans for terms that exceed 15 years can hedge against the risk that rates will rise (squeezing their profit margins by increasing the cost of their short-term deposit funding) by entering into fixed or floating rate swaps. Similarly, to hedge against the risk that borrowers will prepay their fixed-rate mortgages when interest rates fall, banks can purchase interest rate puts or floors where the option pays the difference in yield between the floor rate and a reference rate such as the London Interbank Offered Rate (LIBOR).

Although community banks could theoretically use derivatives positions like these to hedge against interest rate risk, most community banks lack the sophistication to do so. As illustrated in table 2 on the previous page, over 99 percent of interest rate swap and derivative positions are held by banks with more than $10 billion in assets. During 2001, options and swaps positions were held by 69 percent and 86 percent, respectively, of banks with over $10 billion in assets. In comparison, less than 1 percent of small community banks (assets less than $500 million) held options or swaps positions during 2001.

Maximizing the return from customer relationships

While community bankers often speak to the importance of "serving the community," they cannot pursue this "chamber of commerce" motive for long without earning at least competitive returns. Community bankers that sacrifice earnings to pursue other objectives become targets for takeovers. So as competition

in banking markets has grown more intense, community banks have been looking for ways to enhance their earnings. Some community bankers have recognized that basic marketing strategies—like cross-selling products to existing customers and imposing higher switching costs on those customers—can play a key role in their bank's earnings profile:

- *"If I can get your residential loan, that's a very important key element, and your main checking account. Now I'm starting to tie you down because I have two of your most basic needs met."*
- *"When they're tied to us with that many services, it makes it harder to leave us."*

Another banker noted that even though his bank may sell off a customer's loan, it doesn't sell off the all-important customer relationship:

- *"While we sell our loans on the secondary market, we're retaining the servicing. Customers deal with us, not an 800 number for [a credit company] in Colorado or California."*

These observations are consistent with recent research studies. Based on a survey of 500 U.S. households, Kiser (2002) found that switching costs are more severe for households with high income and education, which suggests that banks may be strategically targeting these lucrative customers. Hunter (2001) lays out a competitive strategy—which is based on the existence of switching costs—that a community bank can use to retain these high-value customers while it is converting its high-cost, brick-and-mortar distribution system over to an Internet-based distribution system.

When determining which customers are worth retaining and which are not, community banks have traditionally focused on the following banking truism: "80 percent of our profits are generated from just 20 percent of our customers." As a result, bankers have attempted (if only by benign neglect) to cull the less profitable 80 percent of their customers. But the Fed survey suggests that community bankers have started to look at customer profitability issues a bit differently:

- *"The irony is that 10 to 15 years ago, you wanted to get rid of that [frequent overdraft] account. Now, all of a sudden, everyone woke up and figured out that these are the most profitable accounts."*
- *"Our industry hasn't addressed the blue-collar segment of the market. One of the most profitable segments [due to fee income] is the blue-collar worker who goes from paycheck to paycheck. Those individuals are left behind in the industry. We [have tended] to focus our marketing efforts, our product development, toward the wealthier customer."*
- *"The most lucrative product is the checking account with an NSF [non-sufficient-funds] fee ... we used to close those accounts, but now we're letting those customers stay, and our fee income has doubled since last year."*
- *"A regulator told us, 'You've got a few of these people who pay late, you need some more of them.' You don't want the guy who is 30 days late, but 15 days late is okay. You get a nice return on someone who pays late a few times."*

High tech, low tech, or no tech?

Another issue that community banks are grappling with is whether, how quickly, and to what extent they should compete with the new technologies being rolled out by larger banks. Adding a new technology can range from installing individual applications (like account aggregation, automated credit analysis, or telephone banking) to purchasing entire established firms to provide products for on-line sales (like insurance or brokerage products). In either case, adding a new technology may be prohibitively expensive for a community bank:

- *"When the management of a community bank sits down to plan their budget for the next operating year, or for a horizon of three years, they've got one shot to get it right. They might be investing $300,000 or $500,000, which for a community bank might be an entire year's earnings or more. If they get it wrong, they've wiped out their bank for three years."*
- *"I don't think community banks have a more difficult time or are less flexible in their ability to deploy technology. I think we're more flexible than our larger competitors. We're able to roll out faster and more efficiently in a general sense. However, we don't typically have a large say in the design structure itself of the technology that becomes deployed—it's typically engineered by larger institutions."*

Furthermore, there is no guarantee that installing the new technology will add to the bank's bottom line. However, *not* installing certain applications may have even worse consequences, as these responses suggest:

- *"It would make us vulnerable [against the competition] if we didn't have it."*

- *"You're not going to get us to be the first bank in the country to claim that [Internet banking] is going to be a significant profit generator. It will be a means to protect the Gen Xers and Gen Yers and the Net generation, instead of finding another bank because their father's or grandfather's bank doesn't do anything."*

Given this uncertainty, it is paramount that community banks carefully choose only those applications that match their business strategies and serve the needs of their customers. But this is only half the battle. After the bank has chosen and installed the new applications, it must manage those applications efficiently. In a recent study of the Internet-only business model, DeYoung (2001a, 2001b) finds that the most successful Internet-only banks and thrifts are those that follow fundamental, low-tech management practices like controlling their costs. Here is the experience of one community banker with a new technology:

- *"We used to average 225 transactions per teller per day, and that average is down to 180 [because of telephone banking]."*

Because community banks are often too small to profitably deploy certain applications themselves, they may decide to form alliances with other financial services providers to give their customers access to brokerage services, insurance products, or even credit cards. However, a community bank that strikes up one or more strategic alliances must be careful to maintain its role as the primary customer contact, or risk losing customer relationships to the allied financial service providers. In fact, one community banker worries that her all-important customer relationships may be vulnerable to high-tech intrusions—in this case, account aggregation—even if she doesn't engage in strategic alliances:

- *"The rule is, he who aggregates first, wins. It's going to kill the ... community banks out there, because the large banks are going to cherry-pick the cream of the crop of your customers. They'll see what accounts your customers have, then offer them their teaser rates and the customers will take it. So, who's going to use aggregation services? The wealthier clients who are on the road and want to see all of their accounts in one place."*

Identity crisis: Banker or financial services provider?

Deregulation has removed most of the traditional boundaries that separated commercial banks from other financial services providers like insurance companies, brokerage firms, investment banks, and venture capital firms. Commercial banking companies can now offer virtually any of the financial products and services previously available only from those more specialized firms. Should community banks take advantage of this new freedom and broaden their product offerings? Or should community banks stick to a "pure banking" strategy? Some bankers wish they didn't have to make such choices:

- *"I think if we stuck with what we are best at, we would be a lot better off. If bankers stuck with banking, and let the insurance guys stick with insurance instead of them trying to write car loans, do IRAs, and write residential mortgages that they know squat about, and us trying to write homeowner's and life insurance and write trusts, we'd all live a better life."*

A narrowly focused, pure banking strategy may prove to be profitable for some community banks—but a focused strategy will not shield community banks from competition from nonbank financial firms. A number of the bankers that we surveyed used brokerage firms as examples of the threats, pitfalls, and opportunities facing the community banking sector in the newly deregulated financial services world:

- *"The competition isn't commercial banks anymore, it's brokerage companies. You have [national insurance company] offering car loans. Your broker is giving you investments, selling you credit cards, giving you a second mortgage on your house, giving you a line of credit, giving you interest on your checking account, on your idle funds."*
- *"It's difficult to offer [financial planning] and make money through a third-party. You have to contract because you need brokerage licenses, and most banks don't have staff that are licensed. So, you have to have a partner that can do it, and the margins aren't very good."*
- *"I think that the general public really prefers the stereotype of the financial planners of the [national brokerage firms]. We have a person who is just as capable, but he focuses on things that are more profitable. Most financial planning is not profitable. There are software packages that for $40 can do what 80 percent of the people want."*

The playing field isn't level

Many of surveyed community bankers voiced strong concerns that the rules of competition worked against them—namely, that state and federal regulations

placed them at a disadvantage relative to their large bank and nonbank rivals. All commercial banks must comply with costly regulations, such as the requirements of the Community Reinvestment Act (CRA) and the costs related to periodic safety and soundness examinations. In some cases, the fixed costs of complying with these regulations may fall more heavily on community bankers. The Fed survey uncovered some differing points of view about the impact of these costs on community banks:

- *"We shouldn't minimize the significance of competition from our large bank counterparts, but at least they play by the same rules."*
- *"The new state laws tie our hands because of all the regulations that come with it. Out-of-state banks open branches here but are regulated by their own state's laws, while we are subject to the laws of this state, which mandate a lower loan to value ratio. It hurts us in our ability to do loans that they [the out-of-state competition] might be able to do."*

The surveyed bankers were more uniformly concerned about the regulatory advantages enjoyed by their nonbank competitors. While it is true that these nonbank competitors incur substantially fewer regulatory expenses, limitations, and intrusions, it is also the case that banks enjoy two regulatory advantages that are unavailable to many of their nonbank competitors: access to the payments system and the ability to issue insured deposits. On balance, it is not clear how the various costs and benefits of the financial regulatory environment net out, but community bankers nonetheless feel that they often come out on the short end:

- *"Farm Credit has an advantage in that they have no requirement to live up to CRA rules. They can cherry-pick. They don't have to provide funding to low and moderate groups."*
- *"Payday loan companies are driving bankers crazy because they're totally unregulated."*

The most frequent and vociferous complaints were reserved for credit unions—cooperatively owned depository institutions that are not subject to federal or state income taxes. Credit union members (that is, their owners) can consume the resulting tax savings in the form of lower interest rates on loans and/or higher interest rates on deposits. This tax advantage makes membership in a credit union an attractive alternative to depositing funds in a community bank.

TABLE 3

Trends at U.S. credit unions and community banks, 1997–2001

	Credit unions			
	Number	**Membership**	**Assets**	**Mean assets**
		(millions)	*($ billions)*	*($ millions)*
1997	11,238	71.4	351.2	31.25
1998	10,995	73.5	388.7	35.35
1999	10,628	75.4	411.4	38.71
2000	10,316	77.6	438.2	41.51
2001	9,984	79.4	501.6	50.24
% change	–11.2	+11.2	+42.8	+60.8

	Community banks[a]			
	Number	**Deposit accounts < $100,000**	**Assets**	**Mean assets**
		(millions)	*($ billions)*	*($ millions)*
1997	9,323	108.5	1,103.8	118.40
1998	8,946	106.8	1,132.7	126.62
1999	8,779	104.8	1,202.7	136.88
2000	8,524	103.0	1,247.7	146.38
2001	8,295	101.9	1,326.6	159.93
% change	–11.0	–6.1	+20.2	+35.1

[a]Community banks defined as insured commercial banks with assets less than $1 billion in 1997; after 1997 this threshold was adjusted upward for 12 percent annual industry growth.

Sources: National Credit Union Administration (2001) and Federal Deposit Insurance Corporation (1997–2001).

TABLE 4

Mean averages for selected financial ratios at large banks and community banks, 1996–2000

		Small community banks		Large community banks	
	Large banks	All banks	Best-practices banks	All banks	Best-practices banks
Return on equity	.1653	.1267***	.1748**	.1431***	.1832***
Loans to assets	.6469	.6207***	.6426	.6304*	.6342
Noninterest expense to net revenue	.6013	.6133	.5646***	.6040	.5776***
Core deposits to assets	.4749	.7286***	.7387***	.6785***	.7258***
Noninterest income to net revenue	.3967	.1684***	.1800***	.2192***	.2229***

Notes: Large banks have more than $10 billion in assets. Small community banks have less than $500 million in assets. Large community banks have between $500 billion and $1 billion in assets. Best practices banks are defined as having return on equity higher than the group median. Assets are in 1999 dollars. ***, **, or * indicate that the community bank mean is significantly different from the large bank mean at the 1 percent, 5 percent, or 10 percent level, respectively.
Source: DeYoung and Hunter (2003).

- *"It's not a fair playing field. Credit unions are not subject to taxation, so they can lend their money out at 38 percent less. Second, they don't have to spend their time on CRA and other regulations."*
- *"... Credit Unions ... I won't get started on that! We get hammered on the rates that we're able to pay on our deposits, whereas credit unions can offer lower rates on vehicles and higher rates on deposits, and they're not subject to tax."*

Although membership in a credit union is limited to people who share a "common bond"—such as a common employer or a common geographic neighborhood—recent federal legislation liberalizing the interpretation of "common bond" has allowed credit unions to expand their market share at the expense of community banks.[15] As illustrated in table 3, the numbers of credit unions and community banks in the U.S. have declined about equally over the past five years. But while the number of deposit accounts at community banks has declined over this period, the number of credit union members has increased. Furthermore, the assets of credit unions have grown much faster than the assets of community banks.[16]

Conclusion

The slide in the number of community banks over the past 20 years is undeniable. The implications of this slide for the future of the community banking sector are open to debate. What does the future hold for community banks?

Recent experience indicates that well-run community banks can earn high and sustained profits. Table 4 compares selected financial ratios from large banks, community banks, and "best-practices" community banks, defined here simply as the community banks that generated above median return on equity. The best-practices community banks generated significantly higher returns than the average large commercial bank. Furthermore, the table indicates that these well-run community banks used a business model that was clearly different from the one used by the average large bank. On average, these banks used higher amounts of core deposit funding (evidence of relationship banking), incurred lower levels of noninterest expenses (suggesting that well-managed community banks are more likely to survive the industry consolidation), and generated less noninterest income (indicating that high earnings are available to community banks even if they don't enter nontraditional lines of business).

All else equal, the recent past is generally a good predictor of the near future. But long-run predictions about the future of the community banking sector—like all other long-run economic predictions—are subject to a large degree of uncertainty. Ken Guenther, president of the Independent Community Bankers of America, recently issued a statement on this issue that echoed in many ways the sentiments of the community bankers that we have quoted anonymously above:[17]

- *"Pundits continue to mistakenly announce the demise of the community-based banking sector. Simply stated, increased prosperity for Americans means a greater demand for financial services, and community banks continue to provide the customized personal financial services that can*

compete effectively with other providers. Greater use of technology is in no way limited to the exclusive benefit of large financial conglomerates but is employed successfully by community banks to compete most effectively. Before discounting the future of our nation's community-based banks, one should bear in mind that small banks have always been more nimble and responsive than huge banks and have been able to position themselves much faster than the bureaucratic giants. Given their proven ability to adapt to change and their survival over the past century, we can be confident that community banks will remain a competitive force well into the future."

Despite Guenther's optimistic predictions, some would consider the disappearance of almost half of the nation's community banks over the past 15 years to be *prima facia* evidence that the community bank business model is losing its viability. However, others argue that the healthy competition introduced by the deregulation and consolidation of the U.S. banking sector merely exposed the inefficiently run community banks to the pressures of the marketplace, while at the same time providing increased opportunities for efficiently run, progressive community banks to flourish. Not surprisingly, the community bankers that we surveyed embrace the second of these two visions of the future of community banking.

NOTES

[1]There is no generally accepted definition of "community bank." For convenience, a size-based threshold of less than $1 billion in assets is used.

[2]Although economists continue to debate how large a bank must be before it fully exhausts all potential for scale economies, there is general agreement that small community banks have access to substantial economies of scale. For an in-depth review of scale economies in banking, see Berger, Demsetz, and Strahan (1999).

[3]See DeYoung, Hasan, and Kirchhoff (1998), Evanoff and Örs (2001), and Whalen (2001). One explanation for this phenomenon is that the acquiring bank makes numerous changes that intensify competitive rivalry in the local market—for example, underperforming managers are replaced, assets are reallocated to higher yielding investments, excess expenses are slashed, new products are introduced, fees are reduced, or deposit rates are increased. Local banks either respond in kind or lose market share.

[4]This deregulation does have some technical limits. For example, to engage in certain nonbanking financial activities (for example, insurance underwriting) a bank must adopt a new organizational structure called a financial holding company (FHC), in which commercial banking affiliates are capitalized separately from nonbanking affiliates.

[5]Federal Reserve System (1997, 1998, 1999).

[6]See Genay (2000) for details. Core deposits are typically defined as funds in transactions accounts plus funds in savings accounts under $100,000.

[7]There is evidence consistent with this in the Federal Reserve's *Survey of Retail Pricing and Fees* (1997, 1998, 1999), which reports that small banks tend to charge lower fees on deposit accounts.

[8]"Hard" information (for example, salary, wealth, debts) can be gleaned from a borrower's financial statements and credit reports. In contrast, accumulating "soft" information (for example, the borrower's character or her ability to run a business) requires the lender to have personal interactions with the borrower. See Stein (2002) for a detailed discussion.

[9]A study by Celent Communications found "negative returns" to Internet banking at banks with fewer than 10,000 customers. See article in *American Banker* (Thomson Corporation, 2000a). Consistent with these findings, DeYoung (2001a) finds that newly chartered Internet-only banks tend to exhibit deeper scale economies than newly chartered branching banks.

[10]DeYoung (2001b, p. 65) discusses these issues at greater length and provides some industry evidence.

[11]See Strahan and Weston (1998), Peek and Rosengren (1998), and DeYoung, Goldberg, and White (2000) for details on small business lending and the consolidation of the banking industry.

[12]There is an extensive literature on scale and scope economies in the commercial banking industry. See Hunter, Timme, and Yang (1990), Hunter and Timme (1991), Evanoff and Israilevich (1991), Berger and Mester (1997), and Hughes, Lang, Mester, and Moon (2000) for evidence. This evidence suggests that scale economies are modest for community banks under $1 billion, but that larger banks produce a different output mix using a different production technology that yields more substantial economies of scale.

[13]Note that large banks do personalize some of their financial services—for example, investment banking or merger finance to large wholesale clients—but their retail and small business strategies tend to be commodity-like compared with those delivered by small community banks.

[14]The three anecdotes that follow come from the following sources: Thomson Corporation (1999, 2002b) and Bank Administration Institute (1997).

[15]The Credit Union Membership Access Act of 1998 (P.L. 105-219) allows a federal credit union to accept as members groups of up to 3,000 individuals that are not related by a common bond to the current membership group.

[16]The comparatively low community bank asset growth rates are not due to our working definition of a community bank, which truncates the annual populations at $1 billion. The differences in growth rates were even larger when we used a $10 billion asset threshold. (Note that in both cases, we allowed the asset threshold to increase by 12 percent per year to account for average nominal industry growth rates.)

[17]The quoted material is condensed from Guenther (2002).

REFERENCES

Bank Administration Institute, 1997, "The devil in the details," *Banking Strategies*, November/December, p. 24.

Berger, Allen N., Seth D. Bonime, Lawrence G. Goldberg, and Lawrence J. White, 1999, "The dynamics of market entry: The effects of mergers and acquisitions on de novo entry and small business lending in the banking industry," Board of Governors of the Federal Reserve System, working paper, No. 1999-41.

Berger, Allen N., Rebecca S. Demsetz, and Philip E. Strahan, 1999, "The consolidation of the financial services industry: Causes, consequences, and implications for the future," *Journal of Banking and Finance*, Vol. 23, pp. 135–194.

Berger, Allen N., and Robert DeYoung, 2001, "The effects of geographic expansion on bank efficiency," *Journal of Financial Services Research*, Vol. 19, pp. 163–184.

Berger, Allen N., and Loretta J. Mester, 1997, "Inside the black box: What explains differences in efficiency of financial institutions?," *Journal of Banking and Finance*, Vol. 21, pp. 895–947.

DeYoung, Robert, 2001a, "Learning-by-doing, scale efficiencies, and financial performance at Internet-only banks," Federal Reserve Bank of Chicago, working paper, No. WP-2001-06.

_____**,** 2001b, "The financial performance of pure play Internet banks," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 25, First Quarter, pp. 60–75.

_____**,** 1999, "Mergers and the changing landscape of commercial banking (Part I)," *Chicago Fed Letter*, Federal Reserve Bank of Chicago, No. 145, September.

DeYoung, Robert, Lawrence G. Goldberg, and Lawrence J. White, 2000, "Youth, adolescence, and maturity of banks: Credit availability to small business in an era of banking consolidation," *Journal of Banking and Finance*, Vol. 23, pp. 463–492.

DeYoung, Robert, Iftekhar Hasan, and Bruce Kirchhoff, 1998, "The impact of out-of-state entry on the efficiency of local banks," *Journal of Economics and Business*, Vol. 50, pp. 191–204.

DeYoung, Robert, and William C. Hunter, 2003, "Deregulation, the Internet, and the competitive viability of large banks and community banks," in *The Future of Banking*, Benton Gup (ed.), Westport, CT: Quorom Books, pp. 173–201.

Evanoff, Douglas D., and Philip R. Israilevich, 1991, "Productive efficiency in banking: Econometric and linear programming evidence," *Economic Perspectives*, Federal Reserve Bank of Chicago, Vol. 15, No. 4, pp. 11–32.

Evanoff, Douglas D., and Evren Örs, 2001, "Local market consolidation and bank productive efficiency," Federal Reserve Bank of Chicago, manuscript.

Federal Deposit Insurance Corporation, 2001, *Reports of Condition and Income*, Washington, DC.

_____**,** 2000, *Reports of Condition and Income*, Washington, DC.

_____**,** 1999, *Reports of Condition and Income*, Washington, DC.

_____**,** 1998, *Reports of Condition and Income*, Washington, DC.

_____**,** 1997, *Reports of Condition and Income*, Washington, DC.

Federal Reserve System, 2002, *Community Bank Study*, Federal Reserve Bank of Chicago, March, available at: www.frbservices.org/Industry/pdf/CommunityBankStudy.pdf.

_____**,** 1999, *Survey of Retail Pricing and Fees*, Federal Reserve Board, June, available at: www.federalreserve.gov/boarddocs/rptcongress/1999fees.pdf.

_____**,** 1998, *Survey of Retail Pricing and Fees*, Federal Reserve Board, June, available at: www.federalreserve.gov/boarddocs/rptcongress/1998fees.pdf.

_____, 1997, *Survey of Retail Pricing and Fees*, Federal Reserve Board, June, available at: www.federalreserve.gov/boarddocs/rptcongress/feesindex.htm.

Genay, Hesna, 2000, "Recent trends in deposit and loan growth: Implications for small and large banks," *Chicago Fed Letter*, Federal Reserve Bank of Chicago, No. 160, December.

Guenther, Kenneth, 2002, "Community banks have a solid place in the future," *Chicago Tribune*, letter to the editor, June 17.

Hughes, Joseph P., William Lang, Loretta J. Mester, and Choon-Geol Moon, 2000, "Recovering risky technologies using the almost ideal demand system: An application to U.S. banking," Federal Reserve Bank of Philadelphia, working paper, No. 00-5.

Hunter, William C., 2001, "The Internet and the commercial banking industry: Strategic implications from a U.S. perspective," in *Financial Intermediation in the 21st Century*, Zuhayr Mikdashi (ed.), Basingstoke, Hampshire, UK: Palgrave, pp. 17–28.

Hunter, William C., and Stephen Timme, 1991, "Technological change in large U.S. commercial banks," *Journal of Business*, Vol. 64, pp. 206–245.

Hunter, William C., Stephen G. Timme, and Won Keun Yang, 1990, "An examination of cost subadditivity and multiproduct production in large U.S. banks," *Journal of Money, Credit, and Banking*, Vol. 22, pp. 504–525.

Keeton, William R., 2000, "Are mergers responsible for the surge in new bank charters?," *Economic Review*, Federal Reserve Bank of Kansas City, First Quarter, pp. 21–41.

Kiser, Elizabeth K., 2002, "Predicting household switching behavior and switching costs at depository institutions," *Review of Industrial Organization*, Vol. 20, pp. 349–365.

National Credit Union Administration, 2001, *Year-end Statistics for Federally Insured Credit Unions*, Washington, DC.

Peek, Joe, and Eric S. Rosengren, 1998, "Bank consolidation and small business lending: It's not just bank size that matters," *Journal of Banking and Finance*, Vol. 22, pp. 799–819.

Rossi, Clifford V., 1998, "Mortgage banking cost structure: Resolving an enigma," *Journal of Economics and Business*, Vol. 50, pp. 219–234.

Stein, Jeremy C., 2002, "Information production and capital allocation: Decentralized versus hierarchical firms," *Journal of Finance*, forthcoming.

Strahan, Philip E., and James P. Weston, 1998, "Small business lending and the changing structure of the banking industry," *Journal of Banking and Finance*, Vol. 22, pp. 821–845.

Thomson Corporation, The, 2000a, "Internet banking profit seen harder for small banks," *American Banker*, November 3.

_____, 2002b, "Roslyn claims gains from Wamu runoff," *American Banker*, June 19, p. 1.

_____, 1999, "Mercantile customer exodus unlikely as Firstar takes over," *American Banker*, May 5, p. 8.

Whalen, Gary, 2001, "The impact of the growth of large, multistate banking organizations on community bank profitability," Office of the Comptroller of the Currency, working paper, No. 2001-5.

Entry and competition in highly concentrated banking markets

Nicola Cetorelli

Introduction and summary

What determines the number of banks operating in a market? What is the relationship between the number of banks in a market and competitive conduct? These are important questions, whose answers define the industrial organization characteristics of a banking market. They are also questions of fundamental policy relevance for antitrust regulation.

In this article, I address these questions by focusing specifically on very highly concentrated banking markets. I focus on these markets because this is where we would expect to observe the least competitive conditions. Indeed, if there is any likelihood of establishing and maintaining a *cartel,* where firms explicitly or tacitly collude in order to behave as one monopolist, it will be in markets with the fewest firms. It is in these markets, therefore, that firms should be able to impose the highest mark-ups; and, by definition, these markets should raise special antitrust concerns in the event of a merger application. How anticompetitive are highly concentrated banking markets? Is there any evidence of actual collusive behavior? Also, how quickly do markets approach a competitive benchmark, that is, how many additional entrants does it take before we observe higher degrees of competition?

Answers to these questions contribute to the policy debate on competitive conditions in the banking industry and provide information on the current practice for assessing market competition in merger analysis. As is widely known, the procedures to evaluate the competitive impact of merger proposals require an evaluation of the concentration of deposit market shares held by banks operating in the market affected by the merger. According to the so-called structure–conduct–performance paradigm (Bain, 1951), one would expect to observe increasingly anticompetitive conduct where market shares are more concentrated. Market concentration is commonly measured by the Herfindahl-Hirschman Index (HHI), which is defined as the sum of the squared market shares of all banks in the market. The HHI index is bounded from below at zero in the (hypothetical) case of a very large number of extremely small banks and bounded from above in the other extreme case of a monopolist, where the index would then be equal to $100^2 = 10{,}000$. According to the current guidelines for antitrust analysis in banking, if a merger brings a market HHI above the value of 1,800, it has the potential for anticompetitive consequences, thus triggering further analysis before approval. In other words, any market with an HHI above 1,800 is considered highly concentrated and, therefore, more likely to be characterized by anticompetitive conduct. To have a better idea of how an HHI around 1,800 translates in reality, consider that a market with five banks, each controlling an equal share of the deposits market, has an HHI equal to $20^2 + 20^2 + 20^2 + 20^2 + 20^2 = 2{,}000$. As I show below, the average HHI across all the markets I analyze in this article is about 4,000, and 90 percent have an HHI greater than 1,800. Hence, the focus of this article is exactly on the markets that raise special antitrust concerns.

How can we evaluate competitive conduct in such highly concentrated markets? What we would like to measure is what Sutton (1992) defines as the *toughness of price competition*, that is, by how much market prices vary as the number of competing firms increases. If it is really the case that incumbent firms collude and maximize joint monopoly profits, then the entry of an additional firm would not have *any* effect on prices. This extreme model features the *least intense*

Nicola Cetorelli is a senior economist in the Research Department of the Federal Reserve Bank of Chicago.
The author would like to thank Jeff Campbell for many insightful remarks. Shah Hussain provided excellent research assistance.

level of competition (really the lack thereof) and thus represents a good benchmark against which to compare actual market behavior. Any other model of competition will typically assume some price response by incumbents to the decision of an additional firm to enter the market. The general prediction of such models is that prices gradually decrease from the monopoly level as the number of firms increases, converging—at higher or slower speed—to marginal cost, the level predicted by the model of *perfect competition*.

The question then is: How quickly do prices drop from the monopoly level? Figure 1 depicts alternative paths for the price level as a function of the number of firms in the market for different competitive models. According to what I illustrated above, the joint monopoly model does not predict any change in prices as N increases. The other two paths (C1 and C2), from top to bottom, are for two alternative models with increasing intensity of competition.

Ideally, we would like to be able to estimate the empirical relationship between price and the number of firms. However, doing so requires accurate information on price and cost variables, information that is typically unavailable, especially at the required level of disaggregation (that is, focusing on local markets). The methodology I adopt here, proposed by Bresnahan and Reiss in a series of papers (1987, 1990, 1991), exploits the fact that there is a close association between the "price to number of firms" relationship (unobservable) and the relationship between the number of firms and the corresponding minimum market size needed to accommodate one firm, two firms, three firms, and so on. These levels of market size are defined as *entry thresholds*.[1]

In the following sections, I show that one can estimate entry thresholds and, therefore, that one can observe the relationship between the number of firms in a market and the entry thresholds. By analyzing this relationship, one can infer the characteristics of the relationship between the number of firms and the price. Estimating entry thresholds for a cross-section of U.S. local banking markets, I find no evidence consistent with collusive behavior leading to maximization of joint monopoly profits, even in those markets with only two or three banks in operation. Instead, the evidence shows substantial increases in the intensity of competition as markets see the entry of a third or fourth bank and gradual convergence toward more competitive behavior as more banks enter.

Description of the methodology

The following graphical illustrations are helpful in clarifying the concept of market-size entry threshold, its relationship with the number of competing firms, and how this relationship varies according to the underlying competitive behavior of market participants.

Consider an economy with identical firms facing the same cost structure and producing the same homogeneous good. Figure 2, panel A depicts the average cost function, AC, and the marginal cost function, MC, of a prospective entrant in a market with $N-1$ firms already in operation. The downward sloping lines D1 and D2 represent alternative levels of *residual demand*, that is, the demand schedule that the entrant would face given the price–quantity decisions of the $N-1$ incumbents (or, in other words, total market demand minus the total quantity produced by the incumbents). Assume that the existing firms maximize joint monopoly profits and that they would continue to do so after the Nth firm enters. I denote the equilibrium monopoly price as $p = p_m$. At that price, if the residual demand schedule is D1, the Nth firm could not enter and survive in the long run, since it would not be able to cover average costs (even though it could be making a handsome price–cost margin, as depicted by the vertical difference between price and the marginal cost function at $q = q_1$). However, at price $p = p_m$ and residual demand schedule D2, the firm could enter, produce q_m, and break even. Hence, given incumbent competitive behavior, if there is a sufficient *per firm* market size, expressed in terms of number of consumers generating a level of demand equal to q_m, then the Nth firm is able to enter the market and join

FIGURE 1

Relationship between price and number of firms

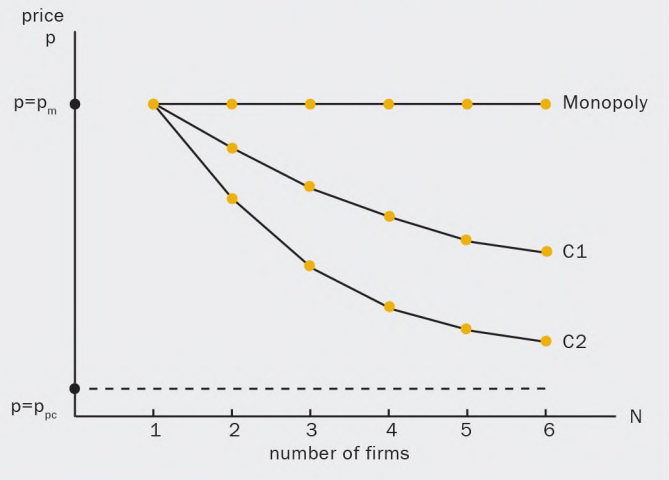
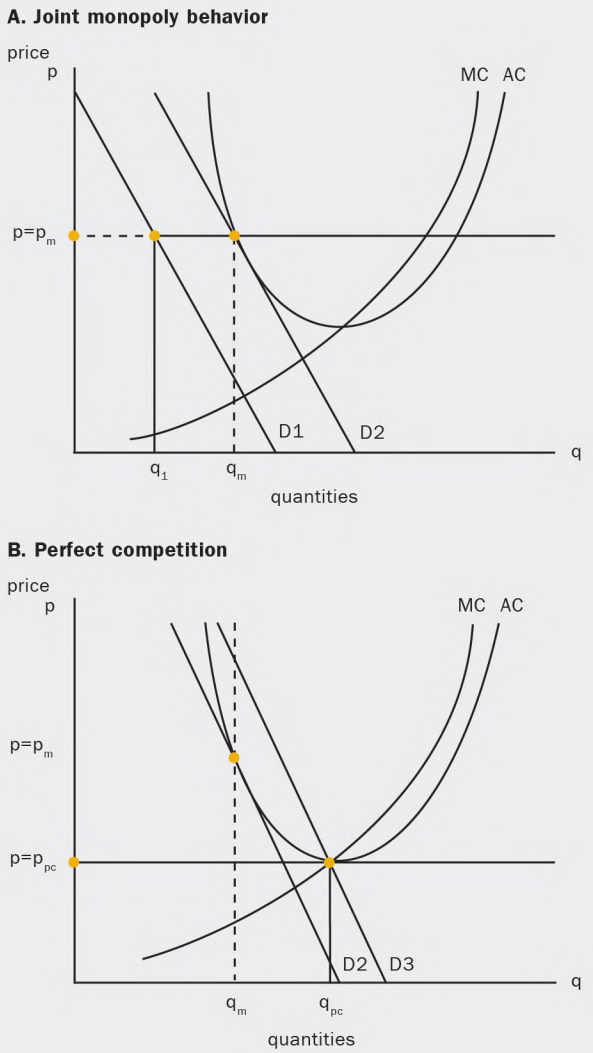


FIGURE 2

Entry conditions



the monopoly agreement. Such a minimum level of per firm market size, conditional on joint monopoly behavior, defines the entry threshold for the Nth firm, which we denote as $s_N(m)$ (where m indicates that this is the per firm entry threshold under joint monopoly behavior).

Consider now the opposite extreme scenario, where the Nth firm would face the most intense competitive response from the $N-1$ incumbents. Figure 2, panel B describes the cost functions of the Nth prospective entrant, its residual demand schedule, and the market price $p = p_{pc}$. This price, equal to the minimum of the average cost function, is the lowest possible that can be set in the industry while allowing firms to break even in the long run. This is the level of price predicted by the model of *perfect competition*. If the residual demand schedule is D2, at price p_{pc} the firm could not meet the long-run profitability condition. The firm could enter only if residual demand were high enough so that it could produce at least a quantity $q = q_{pc}$. As in the previous case, a corresponding per firm market-size entry threshold conditional on perfectly competitive behavior and denoted as $s_N(pc)$ generates the required quantity level.

As one can see from the two graphs, for a given number of market incumbents and a given cost structure, $s_N(pc) > s_N(m)$. This is no accident; it shows that a more intense level of competition necessarily corresponds to a larger per firm entry threshold. This observation is fundamental to learning how to draw an inference from the entry threshold–number of firms relationship to the price–number of firms relationship.

To explore this correspondence further, I use a model characterized by an "intermediate" degree of competitive behavior, the well-known Cournot model. Under Cournot behavior, prospective entrants know that incumbents will not modify their production levels as a consequence of their entry into the market. Hence, given a downward sloping market demand function, the post-entry equilibrium price will necessarily be lower than it was ex ante. Because prices fall as N increases, the Cournot model also predicts that profitability is decreasing in the number of competing firms. But if profitability is decreasing in N, it follows that each consecutive entrant will require an increasingly larger entry threshold in order to enter and survive in the long run.

For example, consider the case where identical firms have cost function $C = cq_n + F$, where cq_n is variable cost and F is a fixed cost component (start-up costs plus additional costs unrelated to the scale of production). Firms face a linear (inverse) demand function, $q(p) = (a - bp)\,S$, where q is total output, $(a - bp)$ is the demand of a representative consumer, and S is the total number of consumers.[2] Under Cournot behavior, each firm chooses the optimal level of production in order to maximize profitability, that is,

$$\underset{q_n}{Max}\ \pi_n = p(q)\,q_n - c\,q_n - F.$$

It can be easily shown[3] that equilibrium profit for each firm n in a market with N firms is

$$1) \quad \pi^*_n = \left(\frac{a-bc}{N+1}\right)^2 \frac{S}{b} - F.$$

As one can see, firms' profitability decreases in N. Therefore, for an "intermediate" model of competition, such as Cournot, the "price to number of firms" relationship follows a decreasing path, such as either C1 or C2 in figure 1. Equation 1 also indicates that, for a given N, profits are increasing in total market size, S.

At what point could the Nth firm enter? As stated above, entry is possible so long as the residual demand for the Nth firm is large enough for revenues to cover average cost. I can express this formally by saying that entry is granted if the following condition is met:

$$2) \quad \pi_N = p_N(a - bp_N)\frac{S}{N} - c(a - bp_N)\frac{S}{N} - F \geq 0,$$

where p_n is the resulting market price after entry of the Nth firm, $(a - bp_N)\frac{S}{N}$ is the quantity produced by firm N, and $\frac{S}{N}$ is the *per firm* market size.

Solving equation 2 in $\frac{S}{N}$ with an equality sign defines the per firm entry threshold:

$$3) \quad s_N = \frac{S}{N} = \frac{F}{[p_N - c](a - bp_N)} = \frac{F}{VP_N},$$

where VP_N denotes per customer variable profits.

Thus, the per firm entry threshold needs to be larger if fixed costs are higher or if variable profitability is lower.

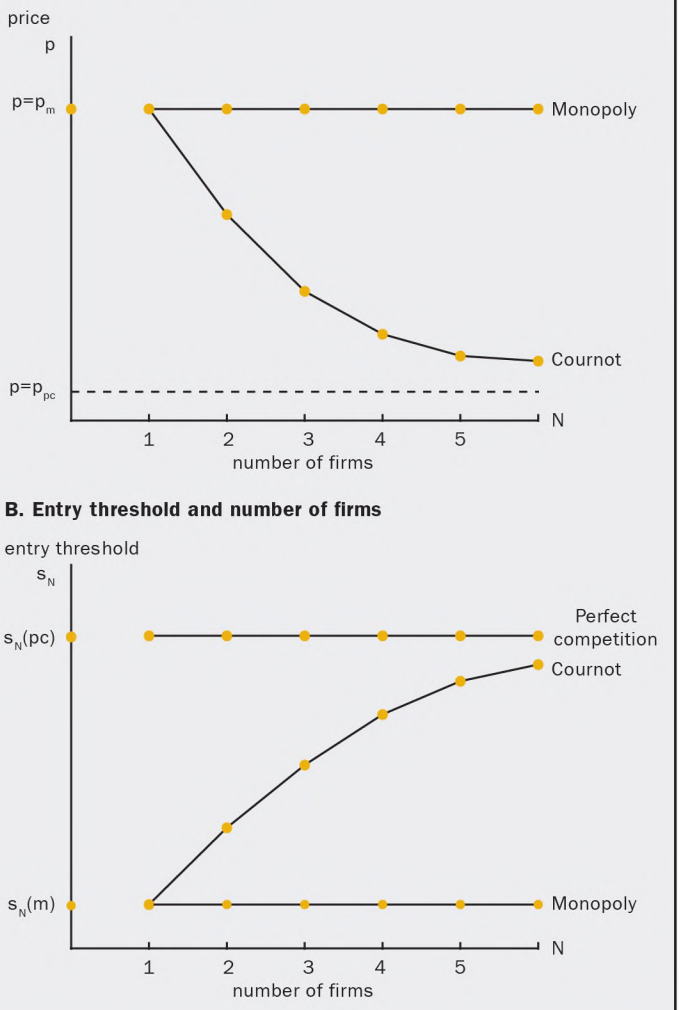
With this last piece of information, I am ready to establish my basic prediction regarding the relationship between entry thresholds and number of firms and in particular how this relationship varies as a function of the intensity of market competition. First, in the benchmark case of joint monopoly behavior, prices do not change with the entry of additional firms. Assuming that each firm has identical cost structure, it follows that under joint monopoly behavior variable profitability does not vary with entry. From equation 3, we see that so long as each firm faces the same cost function, under joint monopoly behavior per firm entry thresholds will be *constant* in the number of competing firms, that is,

$$s_1 = s_2 = s_3 = s_4 = \cdots.$$

For example, suppose that it takes $s_1 = 2{,}000$ consumers for the first firm to enter. Under joint monopoly behavior, the second firm will require an additional

FIGURE 3

Relationship of price, entry threshold to number of firms



2,000 customers before it can enter, and the same holds true for each additional firm.

Still observing equation 3, under Cournot behavior, because profitability decreases in N, per firm size thresholds will actually *increase* in N. In addition, recall that as N grows unbounded, the Cournot equilibrium converges to perfect competition. But from our previous graphical illustration, under perfect competitive conditions the per firm entry threshold is equal to s_{pc}. Therefore, under Cournot:

$$\lim_{N\to\infty} s_N = s_{pc},$$

and consequently,

$$s_1 < s_2 < s_3 \cdots < s_{pc}.$$

Figure 3 describes the predicted path of s_N as a function of N for alternative models of competition (panel B) and the direct correspondence with the "price to number of firms" relationship (panel A). Under Cournot, the path is increasing in N, but it converges to its upper bound s_{pc}. Actual market behavior may show more or less intensity of competition than Cournot; therefore, an actual path for s_N may lie above that for the Cournot economy or below it. The goal of this article is to estimate the empirical path for consecutive per firm threshold ratios and infer changes in competitive "toughness" as N increases.

Data and estimation details

The methodology adopted in this paper allows me to estimate consecutive entry thresholds in local banking markets using a very parsimonious dataset, allowing me to infer the intensity of competition facing new market entrants.

My empirical analysis is based on a cross-section of local U.S. markets, defined as rural counties. Rural counties and metropolitan statistical areas (MSAs) are typically considered reasonable approximations of local banking markets.[4] However, I exclude MSAs from the analysis because this methodology may not be appropriate for markets of relatively large size (see Campbell and Hopenhayn, 2002).[5]

I collected information for the year 1999 on the number of banks, both commercial banks and savings institutions, competing in each U.S. county, from the Summary of Deposits database and matched it with county-level demographic variables from the Regional Economic Information System (REIS) dataset of the Bureau of Economic Analysis. The Summary of Deposits dataset has information through 2001, but the REIS dataset only goes up to 1999. By focusing on a recent year, I have access to a cross-section of markets that have become more and more harmonized in terms of the regulatory playing field. Both intrastate and interstate restrictions to branching and to the creation of *de novo* banks existed to differing degrees in all U.S. states in previous decades. However, the relaxation of these restrictions, culminating in 1994 with the passage of the Riegle–Neal Interstate Banking and Branching Efficiency Act, has led to greater homogeneity of local banking markets across state borders. Hence, one should find more uniform entry conditions for the sample of markets in 1999 and need not be concerned with cross-state differences in the intensity of regulatory entry barriers.

I analyze the likelihood that there is only one bank in a market, two banks, three, four, five, and six or more. The dataset includes 2,257 rural counties. Table 1 illustrates the frequency of bank monopolies, duopolies, and other oligopolies across the total number of counties. In 1999, there were 147 markets with only one banking institution, 281 duopolies, 339 markets with three banks, 313 with four banks, 267 with five banks, and the residual 910 markets with six or more banks. The rural counties with the largest number of banking institutions were La Salle, Illinois, and Dodge, Wisconsin, with 23 banks each.

My emphasis here is on the number of banking institutions that have a presence in a market and not on the total number of bank offices that may be located in a certain market. Certainly the same institution may have multiple branches located in the same market, but my underlying assumption is that within

TABLE 1

Number of banks, markets, and average market size

Number of banks	Number of markets	Frequency	Cumulative percentage	Average market size
1	147	6.51	6.51	3,879
2	281	12.45	18.96	8,656
3	339	15.02	33.98	12,139
4	313	13.87	47.85	16,980
5	267	11.83	59.68	21,713
6+	910	40.31	100	26,429

Notes: Number of banks is the sum of commercial and savings banks in a market. Markets are defined as rural U.S. counties. Average market size is the average population across markets with the same number of banks. Data are for 1999.

the same local market, branches follow a homogeneous strategy vis-à-vis other competitors. Moreover, treating individual branches as independent competitors and estimating conditions of entry would imply that the decision to add an additional branch in a market would be based on competitive considerations against a bank's existing offices, which seems rather implausible.

As pointed out in the introduction, the average HHI across the markets under analysis is about 4,000; 90 percent of the markets have an HHI above 1,800, the level that, if reached as a consequence of a merger, would trigger special scrutiny by antitrust authorities. Hence, my presumption is that if there is any evidence of collusive behavior in banking, this is the sample of markets where it is most likely to show up.

TABLE 2

Estimated entry thresholds

Entry thresholds (000s)		Per bank entry thresholds (000s)		Per firm entry threshold ratios	
S_2	2.170	$S_2/2$	1.085		
S_3	5.782	$S_3/3$	1.927	s_3/s_2	1.776205
S_4	11.211	$S_4/4$	2.803	s_4/s_3	1.454294
S_5	17.091	$S_5/5$	3.418	s_5/s_4	1.219625
S_6	23.825	$S_6/6$	3.971	s_6/s_5	1.161692

Notes: Entry thresholds are obtained using formula 8 in the appendix. S_N denotes the minimum total market size necessary to accommodate N banks. $s_N = S_N/N$ is the per bank entry threshold. Figures are obtained using the maximum likelihood estimated coefficients from table A2 and the sample mean values of the regressors.

Empirical results

The details of the methodology and the econometric analysis are reported in the appendix. In this section, I focus directly on the end product, that is, the estimated entry thresholds reported in table 2.

The results rule out the extreme model of collusion leading to joint monopoly profit maximization. As the estimates indicate, the per bank entry thresholds display a clearly increasing path (see also figure 4). The results are consistent with the predictions of intermediate oligopolistic behavior, where the intensity of competition is sufficiently strong that the entry of each consecutive bank requires significant increases in per bank market size to achieve long-run profitability. More precisely, the entry of a third bank requires the per bank threshold to be about 78 percent higher than that needed in two-bank markets (I obtain this by computing the ratio s_3/s_2). Furthermore, the entry threshold for a fourth bank needs to be an additional 45 percent higher than that for three-bank markets (computed as s_4/s_3). As reported in the last column in table 2, these consecutive per firm entry threshold ratios indicate substantial changes in competitive conduct going from duopolistic market structures to markets with five or six banks. Indeed, the estimates suggest that the per bank entry threshold needed to accommodate a sixth bank is about four times as large as that needed for a duopolist (s_6/s_2, not reported in the table).

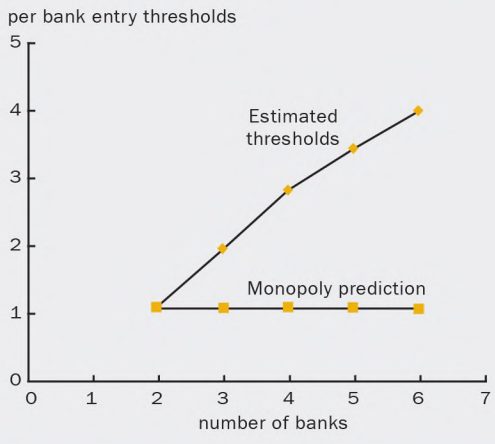
However, the results also suggest that much more of the action, in terms of competitive changes, occurs with the entry of a third or fourth bank than with the entry of a fifth or sixth bank (s_3/s_2 and s_4/s_3 are substantially larger than s_5/s_4 and s_6/s_5). This observation may actually reinforce the justification for setting the HHI threshold level at 1,800 for antitrust regulation: Recall that this number approximately refers to a market with five banks (each one with equal market share). These results suggest that, in fact, with five, six, or more banks, there is not much change in terms of competitive conditions; this implies that there may not be a need for regulatory action in those markets in the case of a merger request.

Conclusion

This article analyzes the conditions of entry and the competitive conduct in a cross-section of highly concentrated U.S. banking markets. The empirical

FIGURE 4

Estimated entry thresholds vs. those implied by joint monopoly behavior



results show, first of all, no evidence consistent with collusive behavior. Indeed, duopolist markets seem already sufficiently competitive. The continuous increase in per bank entry thresholds as additional banks access markets provides further evidence that entry, or the threat of it, improves market competition. By the time a sixth bank has entered, the per bank entry threshold is about two and a half times as high as that needed to accommodate a duopolist. My results, therefore, suggest that U.S. local banking markets have tended to approach fairly high competitive levels rather quickly in recent years, as the number of competing banks has increased. Presumably, by eliminating important barriers to entry, the process of deregulation in banking has enhanced the conditions for market competition.

NOTES

[1]The basic intuition behind this methodology can also be found in Sutton (1992), pp. 27–37.

[2]Bresnahan and Reiss (1991) use a demand function with such characteristics.

[3]See, for example, Mas-Colell, Whinston, and Green (1995), pp. 387–407.

[4]There is a broad list of empirical studies using MSAs and rural counties to define the geographical boundaries of banking markets. Rural counties can be defined as integrated local markets with respective county seats acting as focal points of economic activity. Metropolitan areas are defined as large population nuclei, with adjacent communities having a high degree of social and economic integration with the core. Metropolitan areas comprise one or more entire counties, except in New England, where cities and towns are the basic geographic units.

[5]The median MSA has a population of about 900,000, while the median rural county has a population of about 16,000.

APPENDIX: ESTIMATION OF THE ENTRY THRESHOLDS

The only industry information I need using the methodology proposed by Bresnahan and Reiss (1987, 1990, 1991) is the number of banks operating in each market. Suppose we observe that a market has only two banks in operation. Then they must both be profitable (or in any case the long-run profitability condition for entry for each one of them was met), but a third bank entering the market would have negative profits. More generally, if we observe N banks in a market, we assume their profitability but not that of a potential $N+1$st entrant.

Consequently, I can estimate the likelihood that a market had one bank, two banks, three banks, and so on as a function of a set of variables that should affect bank profitability. This observation suggests the use of a *qualitative response* model, where the dependent variable is the number of banks operating in each market (that is, it takes values 1, 2, 3, 4, 5, or 6, where 6 actually clusters all markets with six *or more* banks). The function to estimate is a profit function similar to equation 2, written in a more general form as

$$4) \quad \Pi_N = S_N V_N(X, \alpha, \beta) - F_N(W, \delta, \gamma) + \varepsilon = 0,$$

where $V_N(X, \alpha, \beta)$ is per customer variable profits for the Nth bank, and $F_N(W, \delta, \gamma)$ is fixed costs. X and W are vectors of market-specific variables affecting variable profits and costs, α, β, δ, and γ are profit function parameters to be estimated, and ε is an error term.

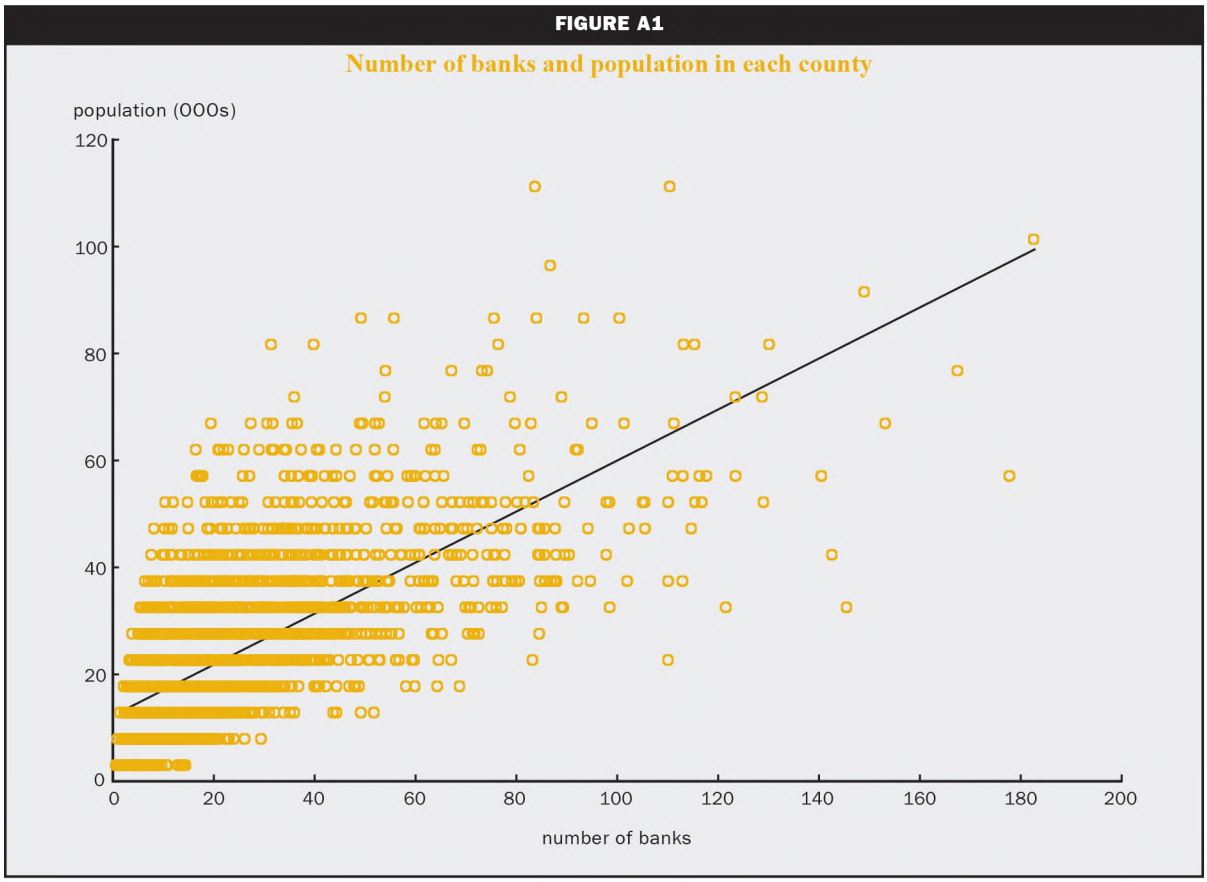
Market size, S_N, is proxied by county total population. Figure A1 shows a scatter plot of market population size and the corresponding number of banks in operation. As expected, we see a positive relationship between market population and number of banks in the market. Indeed, the simple correlation between the two variables is 0.69.

As proxies of demand conditions, I have included the levels of farm income per capita, nonfarm income per capita, and the employment rate. Since markets are represented by rural counties, I have included both farm and nonfarm income per capita as proxies of demand conditions. The prior is that markets with higher per capita income levels should be indicators of more prosperous local economies, which should be reflected in higher demand for banking products; this, in turn, enhances the likelihood of bank entry (for given market size). Similarly, I have also included the county employment rate as an indicator of overall economic activity, which should have the same prediction on the likelihood of entry of the income variables. In order to take into account cost characteristics, I have included a measure of the going wage rate in each county and a measure of land value in the state as indicators of input costs that a potential entrant would face in a particular market. My prediction is that the likelihood of bank entry should be lower in markets exhibiting higher wage rates or land value. Table A1 (on page 26) reports summary statistics for the main variables.

I model firms' variable profits as a linear function of the number of firms and economic variables:

$$5) \quad V_N = \alpha_1 + X\beta - \sum_{n=2}^{N} \alpha_n.$$

FIGURE A1



In particular, this expression allows for variable profitability to progressively decrease in the number of firms operating in the market. More precisely, the variable profits for a monopolist would be $V_1 = \alpha_1 + X\beta$; in the case of a duopolist market it would be $V_2 = \alpha_1 + X\beta - \alpha_2$; in a three-firm market, $V_3 = \alpha_1 + X\beta - \alpha_2 - \alpha_3$, and so on. The decrease in variable profitability could be the result of increased competition or lower efficiency of the subsequent entrants.

I also assume fixed costs are a linear function of the number of firms and of market variables and allow them to be progressively larger for subsequent entrants:

$$6) \quad F_N = \gamma_1 + W\delta + \sum_{n=2}^{N} \gamma_n,$$

so that, $F_1 = \gamma_1 + W\delta$, $F_N = \gamma_1 + W\delta + \gamma_2$, $F_N = \gamma_1 + W\delta + \gamma_2 + \delta_3$, and so on. The increase in fixed costs captures the possible presence of barriers to entry for an additional firm.

Assuming that the error term in equation 4 has a normal distribution, the likelihood to observe N banks in a market is estimated through an *ordered probit* model, where, as noted earlier, the categorical dependent variable is the number of banks reported in operation in each market, and the corresponding probabilities for each category are estimated maximizing a likelihood function whose arguments are those of the profit function in equation 4.

Note that estimating the probability of observing markets with only one bank in operation would require the observation of markets with *no* banks. Given our definition of local markets, there are no rural counties with a count of zero banks in them. Consequently, the first entry threshold that I can actually estimate is that for a second entrant.

With this consideration in mind, and using equations 5 and 6, the profit function to estimate is

$$\begin{aligned} 7) \quad \Pi_N = \; & S_N[\alpha_2 + \beta_1 \, Nonfarm\,Income + \\ & \beta_2 \, Farm\,Income\,Per\,Capita + \\ & \beta_3 \, Employment\,Rate - \sum_{n=3}^{N} \alpha_n] - \\ & [\gamma_2 + \delta_1 \, Market\,Wage\,Rate + \\ & \delta_2 \, Land\,Value + \sum_{n=3}^{N} \gamma_n] + \varepsilon. \end{aligned}$$

TABLE A1

Demographic variables

Variable	Observations	Mean	Standard deviation	Minimum	Maximum
Population	2,257	24.03209	22.9808	0.412	182.399
Nonfarm income	2,257	19.82249	3.854842	8.15167	65.64529
Farm income	2,231	0.870377	1.981094	-7.85946	30.1501
Employment rate	2,257	0.526713	0.151423	0.145826	2.487055
Wage rate	2,257	15.91700	4.486300	4.215000	59.87400
Land value	2,257	933.6611	540.5683	159	6,304

Notes: County population is in thousands. County nonfarm and farm personal income, in thousands of dollars, indicates income levels from nonfarm and farm activities per total county population, respectively. The employment rate is the ratio of total employment and total population in a county. The wage rate is the ratio of total wages, in thousands of dollars, and total employment in a county. Land value is an average across each state. All data are for 1999.

The subscripts for the αs and the γs indicate that the first coefficients to estimate, and the first threshold to calculate, are those for duopolist markets. In view of equation 5, we expect α_2 to be positive, α_i, ($i = 3, ..., 6$) to be negative, and the βs to be positive. In view of equation 6, we expect the γs and δs to be negative (there is a negative sign outside the second bracket in equation 7). Also, following Bresnahan and Reiss, since we allow for constant terms in the V_N function, the coefficient for market population is set equal to one. This is a normalization that expresses units of market demand into units of market population.

Table A2 shows the estimation results for the ordered probit regression model. As the table indicates, all the variables display the expected effect on the probability of bank entry. Entry is more likely in markets with higher levels of both farm and nonfarm income per capita and with higher employment rates, as denoted by the positive and significant coefficients of both income variables and the employment variable. Accordingly, entry is less likely in markets characterized by higher input costs, as indicated by the negative and significant coefficients for the two cost variables. Also, as expected, the variable profitability of each subsequent entrant is estimated to be progressively declining (the $\tilde{\alpha}_i$, $i = (3, ..., 6)$ are negative and significant). At the same time, additional entry is also associated with

TABLE A2

Estimation of the maximum likelihood function

Regressor	Coefficient	Standard error	Z-value	p > Z-value
Nonfarm income	0.00131	0.00048	2.740	0.006
Farm income	0.00730	0.00155	4.700	0.000
Employment rate	0.00019	0.00002	12.330	0.000
Wage rate	-56.27922	7.05996	-7.970	0.000
Land value	-0.00011	0.00006	-1.950	0.051
α_2	0.13115	0.01824	7.190	0.000
γ_2	-0.38973	0.12557	3.100	0.002
α_3	-0.11398	0.01711	6.660	0.000
γ_3	-0.28753	0.09289	3.100	0.002
α_4	-0.03500	0.00649	5.390	0.000
γ_4	-0.41147	0.06537	6.290	0.000
α_5	-0.02126	0.00384	5.530	0.000
γ_5	-0.30051	0.05483	5.480	0.000
α_6	-0.01589	0.00295	5.380	0.000
γ_6	-0.23454	0.05351	4.380	0.000
Observations	2,231			

Notes: This table reports the coefficient estimates of the profit function (equation 7). The model is an ordered probit, where the dependent variable takes values, 1, 2, 3, 4, 5, 6, for the number of banks in each cluster of markets. The last cluster groups markets with six or more banks. A p value below 0.05 expresses statistical significance at the 5 percent level or higher.

increasingly higher fixed costs (the $\hat{\gamma}$ coefficients are also negative and significant).

Once the ordered probit model is estimated, I calculate the entry thresholds using the following formula, obtained by rearranging terms in equation 7:

$$8) \quad S_N = \frac{\hat{\gamma}_2 + \bar{W}\hat{\delta} + \sum_{n=3}^{N} \hat{\gamma}_n}{\hat{\alpha}_2 + \bar{X}\hat{\beta} - \sum_{n=3}^{N} \hat{\alpha}_n},$$

where the circumflex indicates the maximum likelihood estimated coefficients and the upper bar indicates the sample mean values of the regressors in the ordered probit model.

So, for instance, using the actual numbers from the regression results in table A2, the entry threshold for duopolists is calculated as $S_2 = \frac{\hat{\gamma}_2 + \bar{W}\hat{\delta}}{\hat{\alpha}_2 + \bar{X}\hat{\beta}} = 2{,}170$. In per bank terms, $S_2/2 = 1{,}085$. Accordingly,

$S_3 = \frac{\hat{\gamma}_2 + \bar{W}\hat{\delta} + \hat{\gamma}_3}{\hat{\alpha}_2 + \bar{X}\hat{\beta} - \hat{\alpha}_3} = 5{,}782$, $S_3/3 = 1{,}927$, and so on.

REFERENCES

Bain, J., 1951, "Relation of the profit rate to industry concentration: American manufacturing, 1936–1940," *Quarterly Journal of Economics*, Vol. 65, pp. 293–324.

Bresnahan, T. F., and P. C. Reiss, 1991, "Entry and competition in concentrated markets," *Journal of Political Economy*, Vol. 99, No. 5, pp. 977–1009.

_____, 1990, "Entry in monopoly markets," *Review of Economic Studies*, Vol. 57, No. 4, pp. 531–553.

_____, 1987, "Do entry conditions vary across markets?," *Brookings Papers on Economic Activity*, Vol. 1987, No. 3, Special Issue on Microeconomics, pp. 833–871.

Campbell, J., and H. Hopenhayn, 2002, "Market size matters," University of Chicago, mimeo.

Mas-Colell, A., M. Whinston, and J. Green, 1995, *Microeconomic Theory*, Oxford: Oxford University Press.

Sutton, J., 1992, *Sunk Cost and Market Structure*, third printing, Cambridge, MA: The MIT Press (first printing, 1991).

CALL FOR PAPERS

May 7–9, 2003

39th ANNUAL CONFERENCE ON BANK STRUCTURE AND COMPETITION
FEDERAL RESERVE BANK OF CHICAGO

Corporate Governance: Implications for Financial Services Firms

The Federal Reserve Bank of Chicago invites the submission of research and policy-oriented papers for the 39th annual Conference on Bank Structure and Competition to be held May 7–9, 2003, at the Fairmont Hotel in Chicago. Since its inception, the conference has aimed to encourage an ongoing dialogue on current public policy issues affecting the financial services industry. *Although we are requesting, and will include, papers related to the conference theme, we are most interested in high-quality research addressing public policy issues affecting financial services and welcome submissions on all related topics.*

The theme of the 2003 conference will address issues related to corporate governance. In recent months, there have been a number of highly publicized incidents, most notably the Enron and WorldCom scandals, in which appropriate corporate governance may have been lacking. Deficiencies include inadequate oversight by boards of directors, misleading or fraudulent accounting practices, questionable audit arrangements, and various efforts to obfuscate the true financial condition of the firm. As a result, there has been a general rise in investor skepticism, leading to significant uncertainty in equity and credit markets.

These events have also affected the banking sector. First, a number of banks and other financial intermediaries were directly affected because they had large credit exposures to firms that followed questionable accounting practices and subsequently failed—the most obvious being the structured finance arrangements provided to special purpose entities associated with the failed firms. Second, the revelation of these problems has brought into question the efficacy of current mechanisms used to monitor and control firm behavior in order to prevent such problems. The appropriate role and effectiveness of boards of directors, shareholders, creditors (including banks), financial regulators, self-regulation, market regulation, accounting standards, and disclosure rules are all being challenged and modifications are being recommended.

Third, and perhaps more fundamentally, are the implications of these events for the very nature of the financial services business. The ability of financial firms, and financial markets more generally, to function is not based on trust as is sometimes argued, but on information. These recent events serve to highlight the fact that the quality of that information is all-important. Banks have long been recognized as "delegated monitors" for their ability to closely and accurately monitor the economic viability of their customers. It has been argued that this monitoring role is what makes banks "special" and gives them a unique role to play in the economy. Because of the general opaqueness of bank assets, the potential scope for agency problems may be greater here than in other industries.

These corporate governance concerns have raised a number of important public policy questions for the financial services industry. For example:

- How effective is corporate governance—or alternatively, how significant are agency problems—in the financial services industry? Has the potential for agency problems changed as a result of structural changes in the industry?
- What role, if any, did banks or bank regulation play in enabling firms to take advantage of questionable accounting practices? If there was a role, what can be done to prevent such practices in the future?
- The proposed bank capital requirements introduced in the new Basel Accord are highly dependent on accounting and market information. However, recent events bring into question the accuracy of the accounting information and the ability of markets to process that information. Can the Basel standards be successful without changes to accounting standards and/or disclosure requirements? Given the apparent lack of financial transparency at the recent well-publicized failures, to what degree can regulators rely on market discipline?
- Some observers argue that the intertwining of the auditing and consulting functions was a major cause of the recent problems. Are the increased linkages between investment and commercial banking, and between underwriting services and the provision of investment advice, precursors of similar problems for the financial services sector? Should the provisions of the Gramm–Leach–Bliley Act allowing financial holding companies to offer a broader array of services be reevaluated?
- Has the effectiveness of boards of directors to govern firm behavior deteriorated in recent years? Why? Should the liability of directors be changed in an attempt to improve their effectiveness? What are the implications of these questions for bankers who serve on corporate boards?
- Is there a need to overhaul accounting standards? To harmonize international standards? Should the United States consider moving away from a 'rule-based' accounting standard toward a 'principle-based' standard (common in Europe) in which there is an overriding requirement that the reported information fairly represent the true nature of the firm's assets and liabilities?

The 2003 conference will focus on these and related questions. Depending on paper submissions, there will also be a number of additional sessions on industry structure and regulation concerning topics such as:

- Financial market lessons learned from recent crises, particularly the Asian and Latin American crises and September 11th
- Bank capital standards (particularly the proposed Basel Accord)
- Credit access, fair lending issues, and predatory pricing issues
- Measuring and managing risk (particularly for transnational/global financial services companies)
- Financial industry merger activity
- The viability and role of community banks
- Deposit insurance reform
- Restructuring of financial regulatory agencies

Continuing the format of recent years, the final session of the conference will feature a panel of industry experts who will discuss the purpose, structure, problems, and proposed changes associated with an important and topical banking regulation. Past topics discussed at this session include bank antitrust analysis, capital regulation, the role of government-sponsored enterprises, optimal regulatory structures, the appropriate role of the lender-of-last-resort, and alternative means to resolve large complex financial organizations. Proposals for this session are also welcomed.

If you would like to present a paper at the conference, please submit four copies of the completed paper or a detailed abstract (the more complete the paper, the better) with your name, address, affiliation, telephone number, and e-mail address, and those of any coauthors, by December 27, 2002. Correspondence should be addressed to:

Conference on Bank Structure and Competition
Research Department
Federal Reserve Bank of Chicago
230 South LaSalle Street
Chicago, Illinois 60604-1413

For additional information contact:
Douglas Evanoff at 312-322-5814 (devanoff@frbchi.org), or
Regina Langston at 312-322-5641 (regina.langston@chi.frb.org).

Understanding U.S. regional cyclical comovement: How important are spillovers and common shocks?

Michael A. Kouparitsas

Introduction and summary

The holy grail of the study of business cycles is identifying the source of economic fluctuations that affect an economic region. For anyone participating in the quest, there are three paths. First, shocks might be region-specific, affecting only one region of a broader economy. An obvious example is a weather-related shock. Second, they might be common to all regions, such as a change in federal tax rates or monetary policy. Finally, they might initially be region-specific, originating in one region, but eventually spill over to another. The high level of business cycle comovement among U.S. regions suggests that region-specific shocks have a minor role in regional business cycles, leaving spillovers and common shocks playing the major parts in regional business cycles. Despite the growing literature on the subject of regional business cycles, the question of whether the high level of regional business cycle comovement is the outcome of spillovers of shocks from one region to another or common shocks remains unanswered.

The purpose of this article is to determine the extent to which fluctuations in regional economic activity are driven by common and region-specific shocks (including spillovers of shocks across regions). The scope of my analysis is limited to real quarterly per capita income data for the eight U.S. Department of Commerce, Bureau of Economic Analysis (BEA) regions,[1] covering the period from 1961:Q1 to 2000:Q4. I use these data to estimate a model of regional business cycles. This model allows me to decompose a region's cyclical innovations into a part that is common across regions and a residual component that is region-specific. At the same time, the model's structure is rich enough to allow me to formally test whether these region-specific shocks spill over to other regions with at least a lag of one quarter.

Using this framework, I find that spillovers of region-specific shocks across regions account for a statistically insignificant share of the business cycle variation of regional per capita income across the eight BEA regions, while common shocks account for a large and statistically significant share of the business cycle variation of regional income. Based on these findings, I conclude that the high degree of business cycle comovement across U.S. regions over the last 40 years reflects the fact that regions are influenced by common sources of disturbance, rather than any significant spillover of shocks across regions.

Given the different industry mix and strong interregional trade across U.S. regions, these results provide evidence against theories of the business cycle that suggest it owes to cyclical fluctuations being transmitted through trade or production linkages. At the same time, my findings support the notion that the U.S. is an optimum currency area, since they reveal that the BEA regions are largely subject to common sources of disturbance to which they have common responses, which suggests that a common monetary policy is the ideal choice for U.S. regions.

Business cycle properties of per capita U.S. regional income

The starting point for any business cycle analysis is the age-old problem of decomposing fluctuations of economic time series into trend and cycle components. There are many competing methods. I begin my analysis of regional cycles by applying a popular approach to trend/cycle decomposition known as a *band-pass filter*, which limits the cyclical component

Michael A. Kouparitsas is an economist at the Federal Reserve Bank of Chicago. *This article has benefited from discussions with William Testa, Thomas Klier, and David Marshall. The author would also like to thank Carrie Jankowski for outstanding research assistance on this project.*

to that part of the time series occurring at frequencies of 18 months to eight years to real per capita income of U.S. regions.[2] I concentrate on these frequencies of the data since they are arguably of most interest to policymakers (especially those charged with formulating monetary policy). I construct real regional per capital income using the BEA's eight-region nominal quarterly personal income from 1961:Q1 to 2000:Q4, divided by the size of the regional population and deflated by the national Consumer Price Index.[3] With these cyclical components in hand, I can make a preliminary assessment of sources of disturbance to U.S. regions by simply calculating the correlation between regional business cycles. A high correlation implies common sources of disturbances and similar responses to disturbances across U.S. regions, while a low correlation indicates differences in the sources of disturbances and/or different responses to disturbances across U.S. regions.

Estimates reported in table 1, panel A indicate a high level of comovement across U.S. regions, with the contemporaneous correlation between regional and aggregate U.S. income (last row of table 1, panel A) ranging from 0.77 for the Southwest to 0.97 for the Southeast. A similar picture emerges for the interregional correlation statistics. Regions that are geographically close tend to have higher correlation coefficients than other regions. For example, the correlation between New England and Mideast business cycle fluctuations is 0.91, while the correlation between New England and Southwest business cycle fluctuations is 0.51.

Panel B of table 1 reports the correlation coefficients for leads and lags of regional income. The results along the diagonal from the top left corner of the first row to the bottom right of the last row reveal the persistence of regional fluctuations. Coefficients close to one indicate highly persistent cyclical fluctuations, while coefficients close to zero indicate very little persistence in regional fluctuations. Regional cycles are roughly as persistent as the aggregate cycle, with own-lag-correlation coefficients of between 0.90 and 0.94. The off-diagonal cells of this panel, in contrast, highlight whether one region's business cycle leads (or lags) that of the other regions. For instance, if the lead/lag coefficient for regions i and j exceeds their corresponding contemporaneous correlation coefficient in panel A, this implies that i's business cycle leads j's business cycle. The coefficients reported in panels A and B of table 1 do not reveal a strong lead/lag relationship for U.S.

TABLE 1

Regional business cycle comovement and persistence

A. Contemporaneous correlation

	Income at time *t*								
Income at time *t*	**New England**	**Mideast**	**Great Lakes**	**Plains**	**Southeast**	**Southwest**	**Rocky Mt.**	**Far West**	**U.S.**
New England	1.00	0.91	0.76	0.61	0.83	0.51	0.54	0.80	0.85
Mideast	0.91	1.00	0.82	0.68	0.90	0.67	0.66	0.89	0.93
Great Lakes	0.76	0.82	1.00	0.84	0.92	0.65	0.72	0.82	0.94
Plains	0.61	0.68	0.84	1.00	0.82	0.64	0.80	0.68	0.84
Southeast	0.83	0.90	0.92	0.82	1.00	0.75	0.82	0.85	0.97
Southwest	0.51	0.67	0.65	0.64	0.75	1.00	0.77	0.71	0.77
Rocky Mountains	0.54	0.66	0.72	0.80	0.82	0.77	1.00	0.68	0.80
Far West	0.80	0.89	0.82	0.68	0.85	0.71	0.68	1.00	0.92
U.S.	0.85	0.93	0.94	0.84	0.97	0.77	0.80	0.92	1.00

B. Lead/lag correlation

	Income at time *t*+1								
Income at time *t*	**New England**	**Mideast**	**Great Lakes**	**Plains**	**Southeast**	**Southwest**	**Rocky Mt.**	**Far West**	**U.S.**
New England	0.94	0.84	0.73	0.58	0.77	0.40	0.43	0.71	0.78
Mideast	0.87	0.93	0.78	0.63	0.84	0.54	0.55	0.80	0.86
Great Lakes	0.70	0.75	0.94	0.75	0.84	0.52	0.60	0.71	0.85
Plains	0.56	0.65	0.80	0.90	0.75	0.52	0.70	0.59	0.77
Southeast	0.79	0.84	0.88	0.78	0.93	0.61	0.71	0.76	0.90
Southwest	0.54	0.68	0.68	0.65	0.76	0.92	0.72	0.71	0.78
Rocky Mountains	0.56	0.67	0.73	0.78	0.80	0.70	0.92	0.64	0.79
Far West	0.79	0.86	0.84	0.68	0.83	0.60	0.62	0.94	0.89
U.S.	0.82	0.88	0.91	0.79	0.91	0.65	0.70	0.83	0.93

Note: Regional and aggregate income data filtered using the quarterly business cycle band-pass filter described in Baxter and King (1999).
Source: Author's calculations using data from the BEA.

regional business cycles at one quarter, since there are only a couple of cases where a lead/lag correlation exceeds the corresponding contemporaneous correlation. The lead/lag relationship is somewhat weaker at longer horizons of two to four quarters. Overall, these results suggest that U.S. regions have common sources of innovation and similar responses to these disturbances or strong spillovers of shocks across regions that occur at business cycle frequencies. An obvious weakness of this simple approach is that it does not allow for a comparison of the sources of disturbances or responses to disturbances across regions.

A structural model of U.S. regional economic fluctuations

One way of overcoming the limitations of the simple correlation analysis is to use a structural model of the trend and cycle. With appropriate parameter restrictions, a structural model can identify common and region-specific sources of innovation, and identify the shape of responses to common shocks and region-specific shocks. I follow the *unobserved components* (UC) approach of Watson (1986) in decomposing U.S. regional per capita income fluctuations into their trend and cycle components. Unlike the band-pass filter, this approach requires assumptions about the data-generating process. For example, in his analysis of the cyclical characteristics of U.S. aggregate output, Watson modeled the trend of the log of output as a *random walk with drift* and the cyclical component as a *stationary second-order autoregression*. Watson's approach explicitly assumes that the current log of output depends on the most recent past observation plus some random component and a constant term. The constant term, typically called drift, measures the underlying trend growth rate. That is, in the absence of random fluctuations, trend output grows at a rate equal to the drift term. In contrast, positive random fluctuations lead to trend growth in excess of the drift, while negative random fluctuations cause the trend to grow by less than the drift. Using this method, Watson generated a cyclical component for U.S. aggregate output with peaks and troughs that closely match those reported by the National Bureau of Economic Research's (NBER) Business Cycle Dating Committee. Elsewhere, I have shown that this method generates a cyclical component for U.S. aggregate output that closely matches that generated by a band-pass filter.[4]

Unobserved components model

Following Watson's approach, I assume that log per capita income for region i at time t, y_{it}, is composed of a trend τ_{it} and cyclical c_{it} component,

$$1) \quad y_{it} = \tau_{it} + c_{it}, \qquad \text{for } i = 1, \ldots, 8.$$

The trend is assumed to be a unit root with drift,

$$2) \quad \tau_{it} = \delta_{it} + \tau_{it-1} + \mu_{it}, \qquad \text{for } i = 1, \ldots, 8,$$

where the drift term, δ_{it}, measures the trend growth rate of per capita income in region i at time t; μ_{it} is the innovation to the trend of region i's per capita income at time t, which is distributed as an independent normal random variable with mean zero and variance $\sigma^2_{\mu i}$; and the μ_{it}s are assumed to be orthogonal for all t. In this setting, trend output grows at the rate of the drift term in the absence of random fluctuations. Positive shocks lead to trend growth above the drift, and negative shocks lead to trend growth below the drift. Elsewhere, I have shown that the trend growth rate of U.S. aggregate output has changed over time, so I extend Watson's model by allowing the drift to vary over time according to predetermined break points. I consider three periods that are widely considered by empirical researchers, such as Gordon (2000), to be periods in which the trend growth rate of productivity changed significantly: the productivity slowdown era from 1972:Q3 to 1995:Q4; the new economy era from 1996:Q1 to 2000:Q4; and the pre-productivity slowdown era from 1961:Q1 to 1972:Q2.

I also build on Watson's approach by assuming the cyclical component is made up of two parts, a common cycle across regions, x_{nt}, and a regional cycle, x_{it}. I permit regions to have different sensitivity to the common component governed by a parameter γ_i:

$$3) \quad c_{it} = \gamma_i x_{nt} + x_{it}.$$

Under this assumption, regions that do not have a region-specific cycle x_{it} would have income y_{it} that was directly proportional to the common component x_{nt} and their business cycles would be perfectly correlated. The dynamics of the common component x_{nt} follow Watson's specification for the U.S. aggregate cycle of a stationary second-order autoregression:

$$4) \quad x_{nt} = \rho_1 x_{nt-1} + \rho_2 x_{nt-2} + \varepsilon_{nt},$$

where ρ_1 and ρ_2 are scalar coefficients and ε_{nt} is the innovation to the common cyclical component at time t, which is distributed as an independent normal random variable with mean zero and variance σ^2_n. For ease of exposition, I allow $X_t = [x_{1t}, x_{2t}, \ldots, x_{8t}]'$. I assume that the dynamics of the regional cycles follow a first-order vector autoregression:

$$5) \quad X_t = \Phi X_{t-1} + \varepsilon_t,$$

where Φ is an 8×8 matrix of coefficients and $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{8t}]'$ is the vector of innovations to the regional cycle, which is distributed as an independent normal random vector with a zero mean vector and covariance matrix Λ. I identify the region-specific cyclical innovations by limiting the analysis to the case where shocks to x_{it} do not affect x_{jt}, for all $i \neq j$, at time t. In other words, the covariance matrix of regional innovations Λ is assumed to be a diagonal. In this case, the extent of spillovers of cyclical shocks from one region to another is indicated by the off-diagonal elements of the coefficient matrix, Φ. Details of the estimation strategy are provided in box 1.

Results

With the estimated model in hand, I present two sets of results. The first set focuses on measures of U.S. and regional business cycles. The second set concentrates on answering the question of whether the strong pattern of regional cyclical comovement is due to common shocks or spillovers. For completeness, I report all model parameters in tables 2 to 5. I discuss previous approaches to modeling regional income fluctuations in box 2.

Measuring business cycles

The mainstream academic view of business cycles emphasizes that they consist of expansions at about the same time in many economic activities/regions, followed by similarly general contractions. In other words, the U.S. business cycle can be measured by common cyclical fluctuations in regional activity, while variation in regional activity that is not explained by the common cycle serves to highlight region-specific sources of disturbance.

U.S business cycle

Figure 1 (on page 35) plots the common cyclical component of per capita income across U.S. regions

BOX 1

Estimation strategy

The model described by equations 1 to 5 is a variant of Watson and Engle's (1983) general dynamic multiple indicator-multiple cause (DYMIMIC) model. This framework allows unobserved variables to be dynamic in nature, as well as being associated with observed variables. DYMIMIC models are typically estimated using maximum likelihood. In this setting, the likelihood function is evaluated using the Kalman filter on the model's state space representation.[1]

One of the requirements of maximum likelihood is that the data used in the estimation must be stationary. Augmented Dickey-Fuller unit root tests applied to the log-levels and log-first-differences of real per capita income for the eight BEA regions suggest that the null of a unit root cannot be rejected for any of the level data series at the 5 percent level of significance. However, the null of a unit root is rejected for the first-difference data at the same level of significance. In light of this, I specify and estimate the model using the log-first-differences of real per capita regional income.

Under this transformation of the data, the state space representation of the model is described by the following measurement equation:

$$\Delta Y_t = \begin{bmatrix} \delta_{61:1,72:2} & \delta_{72:3,95:4} & \delta_{96:1,02:4} \end{bmatrix} \begin{bmatrix} D_{61:1,72:2} \\ D_{72:3,95:4} \\ D_{96:1,02:4} \end{bmatrix} + \begin{bmatrix} \gamma & I_{8\times 8} \end{bmatrix} \begin{bmatrix} \Delta x_{nt} \\ \Delta X_t \end{bmatrix} + \mu_t ,$$

and transition equation:

$$\begin{bmatrix} x_{nt} \\ X_t \end{bmatrix} = \begin{bmatrix} \rho_1 & 0 \\ 0 & \Phi \end{bmatrix} \begin{bmatrix} x_{nt-1} \\ X_{t-1} \end{bmatrix} + \begin{bmatrix} \rho_2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_{nt-2} \\ X_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{nt} \\ \varepsilon_t \end{bmatrix},$$

where $Y_t = [y_{1t}, y_{2t}, \ldots, y_{8t}]'$; $\delta_{t1,t2} = [\delta_{1t1,t2}, \delta_{2t1,t2} \ldots, \delta_{8t1,t2}]'$; $D_{t1,t2}$ is one for $t1 \leq t \leq t2$ and zero for all other t; $\gamma = [\gamma_1, \gamma_2, \ldots, \gamma_8]'$; $\mu_t = [\mu_{1t}, \mu_{2t}, \ldots, \mu_{8t}]'$; $I_{8\times 8}$ is an 8×8 identity matrix and $\Delta z_t = z_t - z_{t-1}$.

Identification of the model's parameters requires two additional restrictions on the parameter space. First, the vector governing the sensitivity to the common income component γ is identified by normalizing one γ_i to unity. I use the Southeast as the benchmark region, largely because the volatility of fluctuations of the quarterly growth rates of Southeast income is the same as that of aggregate U.S. income. Second, all innovations are assumed to be orthogonal.

[1] I estimate my DYMIMIC model using the recursive EM algorithm described by Watson and Engle (1983). To avoid local optimization problems, I examined a wide range of starting values and imposed severe convergence criteria on the parameter space of 1×10^{-7}. Standard errors of the parameters are estimated using a standard gradient search algorithm to evaluate the matrix of second derivatives of the likelihood function at the EM parameter estimates.

BOX 2

Previous approaches to modeling regional income fluctuations

The most closely related study is Carlino and Defina (1995), hereafter CD.[1] They use a structural model to estimate the effects of region-specific spillovers of real per capita income of the eight BEA regions that is virtually identical to the one described by equations 1 to 5. However they deviate along one significant dimension, using observed data rather than unobserved components to decompose regional output into its trend and cycle parts. In particular, they assume that the common cyclical component of regional per capita income is proportional to log U.S. per capita income, which allows them to simply estimate the region-specific cycle as the difference between log per capita regional income and log U.S. per capita income. To see the implications of this assumption, it is important to note that log U.S. per capita income is well approximated by a weighted sum of the log per capita regional income, where the weights are equal to the share of regional per capita income in aggregate per capita income sy_i. In terms of my notation, CD assume:

$$x_t = \sum_i sy_i y_{it}.$$

In the context of both models this implies:

$$x_t = \sum_i sy_i(\gamma_i x_t + x_{it}).$$

CD also assume that regional sensitivity to the common cyclical component is the same across regions (that is, $\gamma_i = 1$ for all i), which according to my analysis cannot be rejected at typical levels of statistical significance (see table 2). However, this restriction implies that the share-weighted sum of the regional cyclical components at all dates is zero:

$$x_t = \sum_i sy_i(x_t + x_{it}) = x_t + \sum_i sy_i x_{it}\ , \text{ or}$$

$$\sum_i sy_i x_{it} = 0\ ,$$

which is clearly rejected by my and CD's analyses, since a failure to reject this assumption would mean that the regional cyclical components were collinear, thereby making it impossible to identify the spillover matrix Φ in equation 5. In other words, CD's model is misspecified, because their simplifying assumption that the common cycle is explained by observed fluctuations in aggregate income is not consistent with the rest of the model. My unobserved component model overcomes this weakness, since the common and region-specific components are by design consistent with all aspects of the model.

[1]See Carlino and DeFina (1998) for an extensive literature review of empirical studies of regional business cycles. From a methodological standpoint, Rissman (1999) is the most closely related study to mine. Her unobserved components model of regional fluctuations, which is estimated using regional employment data, differs significantly from the model of this article along a number of dimensions that make direct comparison of the estimated coefficients impossible. Despite these differences, her analysis delivers similar conclusions to this article with regard to the sources of innovation in regional activity. In particular, she finds, as I do, that fluctuations in regional activity are largely driven by common sources of innovation.

(expressed as a percentage deviation from the Southeast's trend), against the NBER's business cycle peaks and troughs. I find, just as Watson did with U.S. aggregate income, that the UC approach generates a measure of the U.S. business cycle that has turning points that closely match those of the NBER.

According to this figure, the U.S. economy has been operating below its trend for most of the 1990s, which on first glance is difficult to reconcile with the fact that U.S. output grew strongly in the mid- to late 1990s. This counterintuitive finding is resolved by the fact that the UC model attributes much of the strong growth in income over the second half of the 1990s to an increase in the trend growth rate of regional per capita income (see table 3). One interpretation of these results is that the U.S. experienced a permanent rather than a temporary increase in its productivity growth rate in the 1990s.

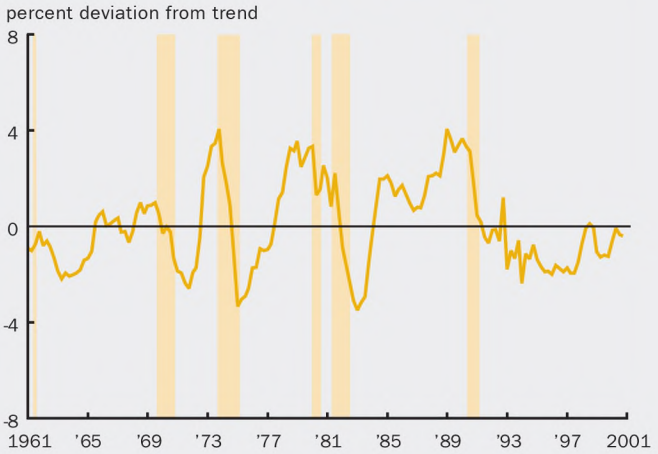
Table 2 reports the differences in regional sensitivity to the U.S. business cycle captured by the γ_is in equation 3. As discussed in box 1, γ_i for the Southeast is normalized to 1. The point estimates of these coefficients indicate that the Plains is the only region that is more sensitive than the Southeast. The Great Lakes has roughly the same sensitivity to the U.S. business cycle as the Southeast, while all the other regions are less sensitive than the Southeast. However, a formal statistical test cannot reject the hypothesis that the γ_i values are equal to one, suggesting that differences in regional sensitivity to the U.S. business cycle are not statistically significant.

Regional cycles

Figure 2, in contrast, highlights differences in the cyclical fluctuations of U.S. regions by plotting the region-specific cycles (expressed as a percentage

FIGURE 1

U.S. business cycle



Note: Shaded areas indicate recessions as defined by the National Bureau of Economic Research.
Source: Author's calculations based on BEA data.

deviation from the region's trend). According to this figure, the Southeast has the weakest region-specific cycle, suggesting that its cyclical behavior is largely explained by movements in the common cyclical component. This reflects the fact that the Southeast has an industrial structure that is virtually identical to that of total U.S. income (see table 6 on page 40). The remaining seven regions fall into two distinct groups.

The first comprises regions, the Southwest, Rocky Mountains, and Plains, that devote a disproportionate share of their industrial activity to the production of commodities. Region-specific cycles of this group are dominated by fluctuations in commodity prices that are to a large extent exogenous to the region. For example, the oil-intensive Southwest's idiosyncratic cycle clearly reflects the large oil price movements of the 1970s and early 1980s, while the mineral-intensive Rocky Mountains' region-specific fluctuations are influenced by movements in prices of oil substitutes over this same period. The idiosyncratic cycle of the Plains, in contrast, takes on the highly volatile pattern of agricultural prices, including the boom that occurred in 1973.

Region-specific cycles of the remaining regions appear to be heavily influenced by the creation and destruction of productive inputs in response to economic slowdowns, changes in defense spending, and technical innovation. Two examples clearly stand out in figure 2, the Rust Belt era of the Great Lakes and the Massachusetts Economic Miracle episode of New England.

The Great Lakes' Rust Belt era began with a strong downturn in regional activity in the late 1970s and ended with a regional recovery in the early 1990s. There is a widely held view that because it had developed much earlier than that of other regions, manufacturing technology in the Great Lakes was of an earlier vintage and relatively less efficient. As a result, the Great Lakes' manufacturing sector experienced a relatively larger decline in demand for its products following the economic slowdown

TABLE 2

Sensitivity to common cycle

Region	Coefficient (γ_i)	Standard error	t-statistic ($\gamma_i = 1$)
New England	0.93	0.16	-0.43
Mideast	0.90	0.11	-0.92
Great Lakes	0.99	0.16	-0.04
Plains	1.10	0.20	0.51
Southeast	1.00		
Southwest	0.81	0.16	-1.18
Rocky Mountains	0.82	0.12	-1.47
Far West	0.80	0.15	-1.32

Note: γ_i indicates the parameter for regional sensitivity.
Source: Author's calculations using data from the BEA.

TABLE 3

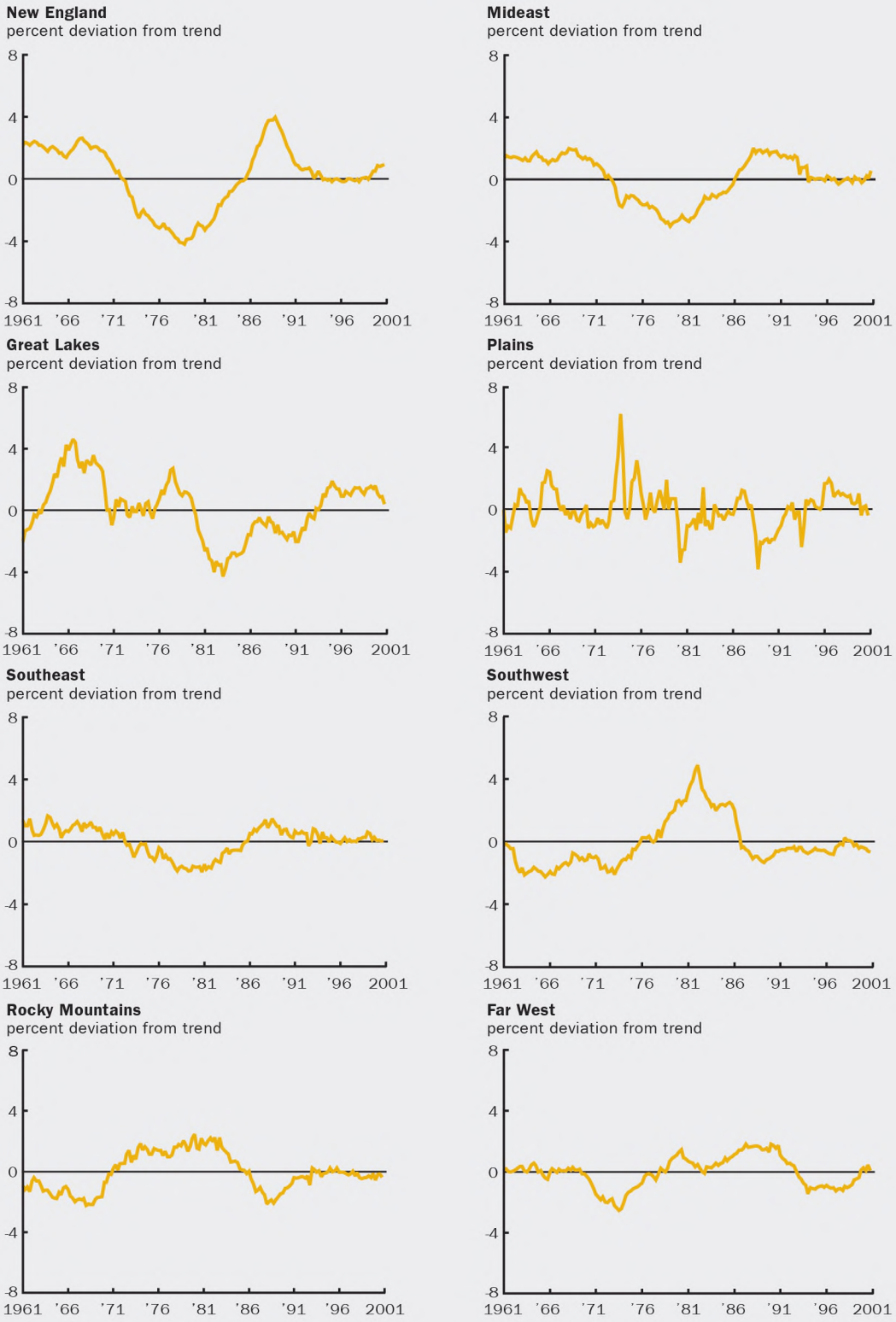
Trend parameters

Region	δ_{it} 1961–72	δ_{it} 1973–95	δ_{it} 1996–2001	$\sigma_{\mu i}$
New England	3.36	2.42	3.35	0.02
Mideast	3.21	2.16	2.84	0.01
Great Lakes	2.78	1.95	2.47	0.02
Plains	3.42	2.06	2.94	0.01
Southeast	4.46	2.43	2.27	0.00
Southwest	3.75	1.98	3.13	0.03
Rocky Mountains	2.80	2.03	3.49	0.01
Far West	3.01	1.65	2.98	0.05

Notes: δ_{it} is the drift term. $\sigma_{\mu i}$ is the standard deviation of the innovation to the regional trend.
Source: Author's calculations using data from the BEA.

FIGURE 2

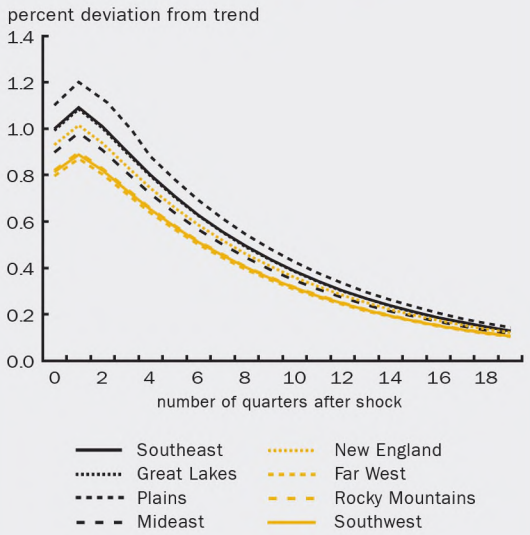
Region-specific cycles



Source: Author's calculations based on BEA data.

FIGURE 3

Response of regional income to common shock



caused by the oil price shocks, since a significant portion of its market share was lost to regions with newer plants. Ultimately, the downturn drove out a significant share of the older plants in the region and paved the way for plants with relatively more efficient technologies to gain market share during the recovery from the recession of the early 1990s.

The Massachusetts Economic Miracle describes the unexpected hi-tech boom of the late 1970s that more than offset the decline in activity brought about by the rapid erosion of New England's manufacturing sector that started in the early 1970s. The era came to an end in the 1980s as New England's hi-tech sector eventually lost its competitive advantage to other regions, such as the Far West, and the end of the Cold War brought about a dramatic decrease in demand for the region's defense-related products. The Far West's regional cycle shows that the region was affected by the same cuts in defense spending that led to the downturn in New England.

Finally, the Mideast's idiosyncratic cycle also reflects the erosion of its industrial sector that started in the early 1970s. In contrast, the Mideast's region-specific cycle improved because of a growing demand for financial services. That trend has persisted since the mid-1980s, leaving the Northeast overall with the largest regional share of activity in finance, insurance, and real estate (FIRE) in table 6. (For a more detailed discussion of these events, see Kouparitsas, 2002).

Common shocks versus spillovers

I assess the source of high comovement of U.S. regional business cycles along two dimensions. First, by studying the cyclical impulse response functions generated by the vector autoregression (VAR) described by equation 5, I assess whether cyclical shocks that originate in one region have a significant effect on the cycles of other regions and at what horizon. Second, I determine the importance of common and region-specific disturbances by decomposing the variance of regional output at business cycle frequencies by source of innovation.

Impulse response functions

Figure 3 describes in detail the way that the eight BEA regions respond over time to a common cyclical shock, normalized to 1 percent of Southeast per capita income. The response of the Southeast is dictated by the coefficients of the second-order autoregressive model reported in table 4. The responses of the other regions reflect differences in the regional sensitivity to common cyclical innovations as reported in table 2.

Figure 4 describe how the level of per capita income (expressed as a percentage deviation from trend) in all eight BEA regions responds over time to an innovation that originates in one of the regions. All shocks are normalized to 1 percent of the per capita income of the region in which the shock originates. For ease of exposition I do not report confidence intervals in this figure; instead I report in the text the few cases where the impulse response functions are significant.[5]

According to my parameter estimates, the Southeast is the only case where shocks that originate in that region have a statistically significant effect on the income of other regions, namely New England and the Mideast. Elsewhere, shocks that originate in one region have a significant positive effect on their own per capita income, but not on the income of other regions. These regions can be divided into two groups according to the persistence of the response to their region-specific income shocks. New England, Great Lakes,

TABLE 4

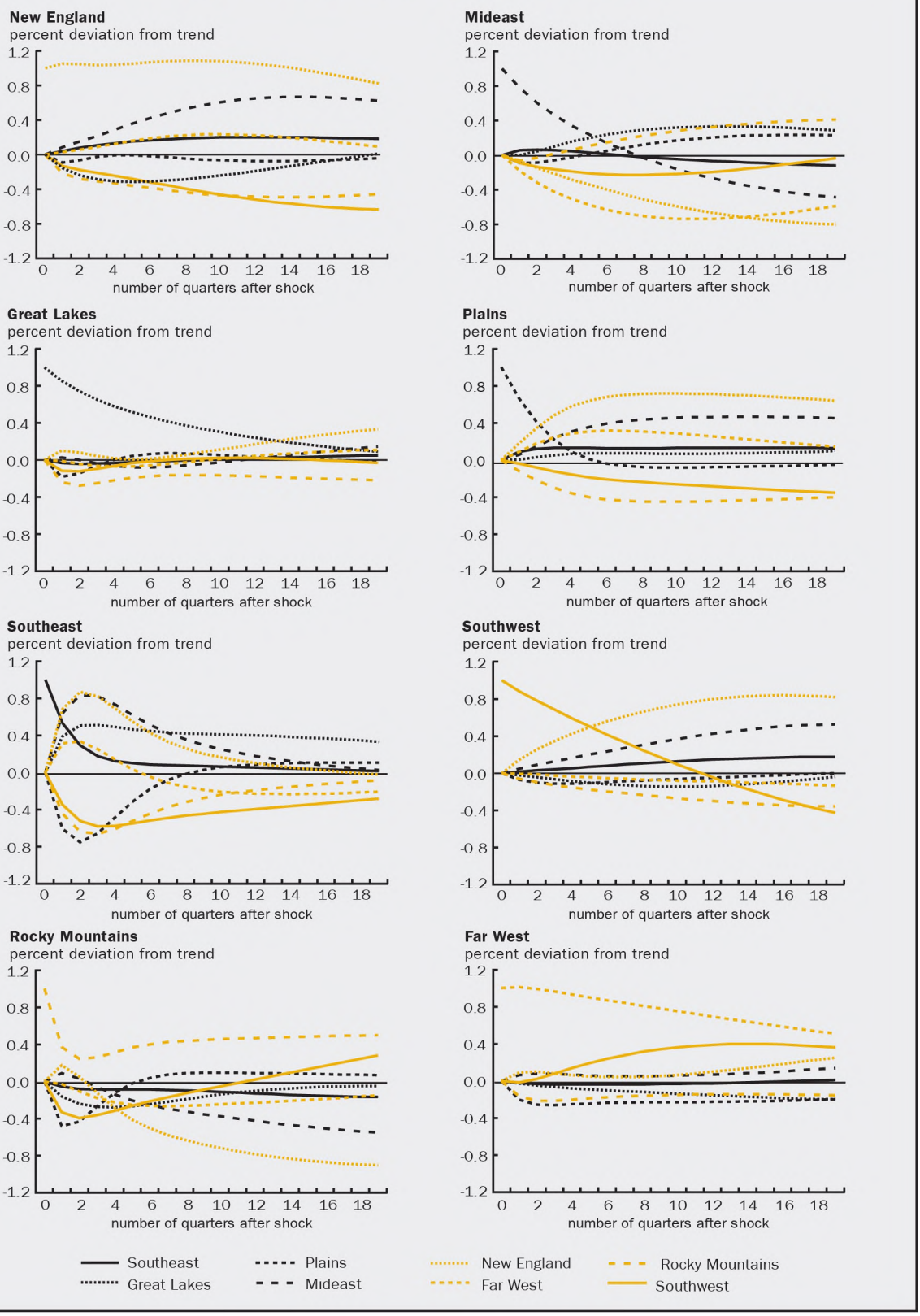
Common cycle parameters

Coefficient	Value
ρ_1	1.09
ρ_2	-0.18
σ_n	0.73

Notes: ρ_1 and ρ_2 are the autoregressive coefficients. σ_n is the standard deviation of the innovation of the common cycle.
Source: Author's calculations using data from the BEA.

FIGURE 4

Responses of regional income to region-specific shocks



Southwest, and Far West have persistent responses to their region-specific income shocks that are statistically significant five to seven quarters after the shock date, while shocks originating in the Mideast, Plains, and Rocky Mountains die out one to two quarters after the shock date.

Returning to the Southeast case in figure 4, note that the response functions of the Southeast and New England are statistically significant two quarters after the shock, while the Mideast response is significant for four quarters after the shock. According to this figure, a 1 percent shock to the Southeast's per capita income causes per capita income of New England and the Mideast to rise by 0.7 percent in the following quarter and an additional 0.2 percent in the subsequent quarter. The confidence interval surrounding these point estimates ranges from 0.2 percent to 1.5 percent, which implies that the spillovers from the Southeast are potentially significant from an economic standpoint. However, note that a typical Southeast shock from 1961:Q1 to 2000:Q4 had a standard deviation of 0.25 percent (see the column labeled $\sigma_{\varepsilon i}$ in table 5), which suggests that spillovers from the Southeast to the Northeast were probably not an economically significant source of innovation over this period.[6]

Variance decomposition at business cycle frequencies

Table 7 ties together the sources of, sizes of, and responses to disturbances by decomposing the variance of regional output at business cycle frequencies.[7] Each column breaks down the variance of regional income by source of shock. For example, the first number in the first column reveals that innovations to the common cyclical component account for a statistically significant 56 percent business cycle fluctuations in New England per capita income. The next number in that column reveals that 5 percent of New England's business cycle variation is explained by shocks that originate in New England, although this is not statistically different from zero at typical levels of significance. Moving down the column uncovers the influence of shocks that originate in other regions. In all cases, the estimates are not statistically different from zero. Overall, the results suggest that spillovers of shocks from other regions are not a statistically significant source of business cycle variation for the New England region.

The remaining columns tell a similar story for the other seven U.S. regions, with a large (statistically significant) share of their business cycle fluctuations explained by the common component. The only other statistically significant sources of business cycle variation in these regions are innovations that originate in the region. For example, region-specific shocks explain almost 30 percent of the business cycle variation of per capita income of the Plains and Southwest, which is not surprising given that they derive a disproportionately large share of their income from commodities, whose price fluctuations are largely exogenous to the U.S. On the other hand, region-specific shocks account for an insignificant share of the business cycle variation of per capita income in the Southeast, which reflects the fact that their industrial composition is virtually identical to that of aggregate U.S. income.

Conclusion

This article develops an empirical model to study the sources of business cycle variation of the eight U.S. BEA regions. Using unobserved component modeling techniques, I identify both common and region-specific sources of innovation in U.S. regional per capita income data. I show that spillovers of region-specific shocks to other regions account for a statistically insignificant share of the business cycle variation of

TABLE 5

Regional cycle parameters

	Φ								
Region	New England	Mideast	Great Lakes	Plains	Southeast	Southwest	Rocky Mountains	Far West	$\sigma_{\varepsilon i}$
New England	1.05	–0.07	0.10	0.18	0.68	0.14	0.18	0.09	0.22
Mideast	0.09	0.77	0.03	0.06	0.64	0.05	0.09	0.06	0.36
Great Lakes	–0.15	0.00	0.85	0.00	0.39	–0.02	–0.16	–0.03	0.43
Plains	–0.09	–0.10	–0.18	0.67	–0.61	–0.07	–0.48	–0.20	0.75
Southeast	0.04	0.05	–0.03	0.09	0.53	0.02	–0.06	–0.02	0.25
Southwest	–0.13	–0.10	–0.12	–0.04	–0.35	0.88	–0.33	–0.02	0.46
Rocky Mountains	–0.21	–0.06	–0.24	–0.12	–0.44	–0.06	0.37	–0.17	0.51
Far West	0.03	–0.19	–0.01	0.09	0.31	–0.01	–0.03	1.01	0.38

Notes: Φ indicates the 8 × 8 coefficient matrix. $\sigma_{\varepsilon i}$ is the standard deviation of the innovation to the region-specific cycle.
Source: Author's calculations using data from the BEA.

TABLE 6

Percent of regional gross state product accounted for by major industry

Region	Agriculture	Mining	Construction	Manufac-turing	Transport. & public util.	Trade	FIRE	Service	Govt.
New England	1.03	0.08	4.62	23.81	7.04	16.11	18.73	17.88	10.68
Mideast	0.77	0.35	4.20	17.66	9.06	15.92	20.41	18.48	13.14
Great Lakes	1.94	0.86	3.69	28.55	9.07	16.16	14.40	14.76	10.58
Plains	5.90	1.53	4.05	20.13	10.43	17.20	14.02	14.49	12.25
Southeast	2.15	4.09	4.80	19.73	9.54	16.96	14.35	13.71	14.67
Southwest	1.77	12.98	5.36	13.14	9.72	16.39	14.71	13.57	12.36
Rocky Mountains	2.88	8.07	5.48	11.91	11.09	15.81	15.12	14.49	15.16
Far West	2.29	2.79	4.63	15.37	8.25	16.75	18.48	17.79	13.65
U.S.	2.04	3.26	4.49	19.38	9.13	16.46	16.54	15.81	12.89

Note: FIRE is finance, insurance, and real estate.
Source: Author's calculations based on BEA data.

TABLE 7

Variance decomposition of U.S. regional income at business cycle frequencies

	Percentage of total variation due to innovation							
Source of innovation	New England	Mideast	Great Lakes	Plains	Southeast	Southwest	Rocky Mountains	Far West
Common	56*	66*	76*	62*	94*	55*	71*	60*
New England	5	1	1	0	0	0	1	0
Mideast	2	14*	1	0	0	1	1	6
Great Lakes	1	1	16*	1	0	1	2	0
Plains	18	7	0	28*	1	2	11	7
Southeast	5	6	2	3	4	3	5	2
Southwest	5	2	1	0	0	29*	1	0
Rocky Mountains	8	3	3	5	0	7	8	2
Far West	0	0	0	1	0	2	1	21*
Total, all shocks	100	100	100	100	100	100	100	100

Note: Numbers in columns may not total due to rounding. * indicates significance at the 5 percent level.
Source: Author's calculations using data from the BEA.

regional per capita income across the eight BEA regions, while common shocks account for a large and statistically significant share of the business cycle variation of regional income. Overall, these findings suggest that the high degree of business cycle comovement across U.S. regions reflects the fact that the regions are influenced by common sources of disturbance, rather than any significant spillover of shocks across regions. Given the different industry mix and strong interregional trade across U.S. regions, this is evidence against theories of the business cycle that suggest it owes to cyclical fluctuations being transmitted through trade or production linkages.

The findings of this article also have implications for the choice of regional monetary policy. In particular, the techniques developed here can be used to address the question of whether a set of regions (or countries) meets Mundell's (1961) criteria for an optimum currency area, by showing that the importance of common sources of innovation in the test region is the same as that of a well-functioning currency union, such as the U.S. For example, one could test whether the European Monetary Union (EMU) was an optimum currency area by repeating the analysis of this article for the EMU countries, then testing to see if the common component across EMU countries is as important a source of variation as it is for U.S. BEA regions.[8]

NOTES

[1]A complete listing of the regions is available at www.bea.gov/bea/regional/docs/regions.htm.

[2]See Baxter and King (1999) for details.

[3]Gross state product (GSP) is an alternative measure of regional activity. The main drawback of GSP is that it is collected annually, which makes it less able to pick business cycle turning points with any precision.

[4]See Kouparitsas (1999) for details.

[5]Confidence intervals are calculated by Monte Carlo methods. Following Hamilton (1994) section 11.7, I randomly draw from the estimated distribution of the model's parameters. For each draw of parameters I generate an impulse response function. I repeat this process 10,000 times. At each lag I calculate the 500th lowest and 9,500th highest value across all 10,000 simulated response functions. These boundaries form the 90 percent confidence interval. If the zero line lies within this interval the impulse response is deemed to be not significantly different from zero at that lag.

[6]I leave a careful examination of the other impulse responses to the reader.

[7]I do this by way of a linear filter that allows me to map from the covariance of the first-difference of regional per capita income to the covariance of the business cycle components of per capita regional income. The mapping is carried using standard spectral/Fourier analysis tools. While, the precise form of the liner filter is, $G(L) = \frac{BP_{6,32}(L)}{1-L}$, where $BP_{6,32}(L)$ is the Baxter-King approximate business cycle band-pass filter for quarterly data; and L is the lag operator (that is, $Lz_t = z_{t-1}$).

[8]See Kouparitsas (2001) for an extended discussion of regional business cycles in the context of optimum currency area criteria.

REFERENCES

Baxter, M., and R. G. King, 1999, "Measuring business cycles: Approximate band pass filters for economic time series," *Review of Economics and Statistics*, Vol. 81, pp. 575–593.

Carlino, G. A., and R. Defina, 1998, "The differential regional effects of monetary policy," *Review of Economics and Statistics*, Vol. 80, pp. 572–587.

_____, 1995, "Regional income dynamics," *Journal of Urban Economics*, Vol. 27, pp. 88–106.

Gordon, R. J., 2000, "Does the new economy measure up to the great inventions of the past?," National Bureau of Economic Research, working paper, No. 7833.

Hamilton, J. D., 1994, *Time Series Analysis*, Princeton, NJ: Princeton University Press.

Kouparitsas, M. A., 2002, "A regional perspective on the U.S. business cycle," *Chicago Fed Letter*, Federal Reserve Bank of Chicago, November, No. 183.

_____, 2001, "Is the United States an optimum currency area? An empirical analysis of regional business cycles," Federal Reserve Bank of Chicago, working paper, No. 01-22.

_____, 1999, "Is there evidence of the new economy in the data?," Federal Reserve Bank of Chicago, working paper, No. 99-22.

Mundell, R. A., 1961, "A theory of optimum currency areas," *American Economic Review*, Vol. 51, pp. 657–665.

Rissman, E. R., 1999, "Regional employment growth and the business cycle," *Economic Perspectives*, Federal Reserve Bank of Chicago, Fourth Quarter, pp. 21-39.

Watson, M. W., 1986, "Univariate detrending methods with stochastic trends," *Journal of Monetary Economics*, Vol. 18, pp. 49–75.

Watson, M. W., and R. F. Engle, 1983, "Alternative algorithms for the estimation of dynamic factor, MIMIC, and varying coefficient models," *Journal of Econometrics*, Vol. 23, pp. 385–400.

Sorting out Japan's financial crisis

Anil K Kashyap

Introduction and summary

Over the last decade, the Japanese economy has underperformed dramatically—growing an average of 1.1 percent per year versus 4.1 percent per year in the previous ten years. At the same time, the country's financial system has fallen into disarray. Recently, the debate over how to address the financial sector problems and the role that this should play in Japan's economic policy have come to the fore. For instance, the International Monetary Fund's (IMF) 2002 Japan country report proposes a four-part program to address the decade long economic slump and to end the current deflation. The first pillar of the program is to "deal decisively with financial sector weaknesses."

In September, the Bank of Japan (2002a) announced an unusual policy initiative, whereby it would begin buying equities that were held by banks. In announcing this decision, the bank pointed to the importance of resolving the nonperforming loan problem. It stated that "in order to resolve the overall problem, a comprehensive and tenacious approach is needed, centering on a more appropriate evaluation of nonperforming loans, promotion of their early disposal, and efforts towards higher profitability on the part of both firms and financial institutions."

The debate came to head when Prime Minister Junichiro Koizumi replaced his financial services minister and promised that he would deliver a plan for the accelerated disposal of banks' bad loans. The new financial services minister, Heizo Takenaka, promptly formed the Financial Sector Emergency Response Project Team to study the bad debt problem, with a promise that the task force would issue an interim report within several weeks and a full report within a month. Yet, when the interim report was circulated in advance of its formal release, the report's analysis and recommendations were heavily criticized by a number of politicians, and the release of the document had to be delayed multiple times.

In this article, I explain why a quick resolution to this problem has not been possible. My central theme is that the financial crisis is sufficiently broad and deep that the necessary institutional changes cannot be initiated or implemented immediately. Nonetheless, many of the ingredients of what will be required for a successful resolution of the problem are clear. The overarching principle is that Japan's banks, insurance companies, and government financial agencies all suffer from different problems and require different solutions. But all three sectors are connected, and a failure to tackle concurrently the problems of all three promises to doom any reform plan.

In the first section, I review the macroeconomic factors that have caused the problems that are now evident in the Japanese financial sector. Poor macroeconomic performance is central to the story, and in this environment some strains on the financial system were inevitable. But I show that macroeconomic conditions alone cannot account for the problems. Instead, one must also account for a host of sector-specific considerations. In the next three sections, I review the challenges facing the reform of the banks, the insurance companies, and the government financial institutions.

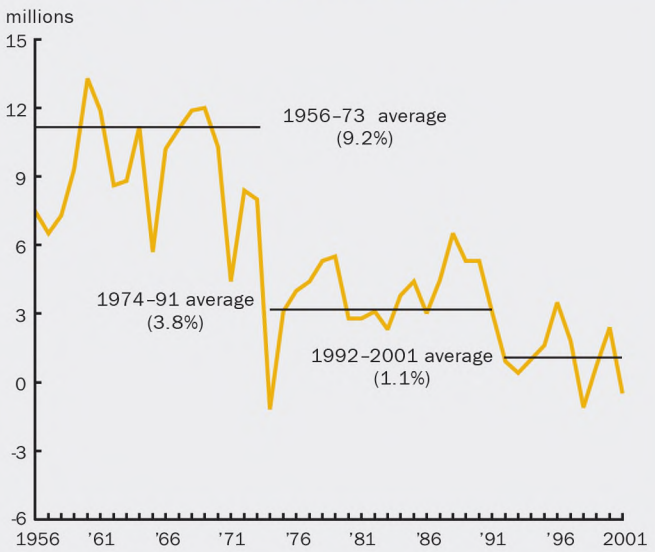
Anil K Kashyap is a professor in the Graduate School of Business at the University of Chicago, a consultant to the Federal Reserve Bank of Chicago, and a research fellow at the National Bureau of Economic Research. *Parts of this article draw heavily on Kashyap, 2002, "Japan's financial crisis: Priorities for policymakers" in* Fixing Japan's Economy *(Japan Information Access Project [JIAP]). The author would like to thank the JIAP for permission to reproduce that work. He is grateful to Robert DeYoung for providing the U.S. figures shown in table 1. He thanks David Atkinson, Tim Callen, Takeo Hoshi, Jakob Lund, Joe Peek, Paul Sheard, Reiko Toritani, and especially Mitsuhiro Fukao for helpful suggestions and comments on earlier drafts. The author also thanks the experts listed in table 2 for allowing him to publish their estimates.*

For each of the three, I provide some estimates on the size of the losses and then explain what will be required to stop them from continuing.

Two primary conclusions emerge. First, the likely cost of the financial problems to the taxpayer is huge: My rough estimate of the *lower bound* for the full cost is approximately 24 percent of Japanese gross domestic product (GDP). Second, the interaction of a number of factors contributes to delaying the resolution of the problems. This delay could easily raise the costs of resolution.

FIGURE 1

Japanese GDP growth 1956–2001



Role of macroeconomics in the financial crisis

The combination of slow growth and the decline of the aggregate price level have each contributed to Japan's financial crisis. The single most important problem for the financial sector has been the anemic growth of the Japanese economy over the last decade. Figure 1 shows GDP growth over the last 45 years to put recent performance in context. After averaging more than 3.8 percent between 1974 and 1991, growth dropped to 1.1 percent over the last decade. Obviously, if there had been more growth in the 1990s, the financial sector would be in better shape now.

The more challenging question is how much the financial sector problems themselves independently contributed to the growth slowdown. A full answer to this question is beyond the scope of this article, but even without resolving it, it is safe to conclude two things about the interplay between the financial sectors problems and growth.

First, it is implausible to argue that the decline in stock and land prices at the beginning of the 1990s can be blamed for the financial sector problems today. This simple explanation fails because the banks and government financial institutions continue to make losses on new loans today. Therefore, the crisis cannot accurately be described as merely delaying the recognition of bad news. While the asset price collapse may have triggered the problems, it cannot be blamed for their continuation at this point. This is an important conclusion because it suggests that ending the crisis will require substantial changes in the financial institutions' operating practices.

Second, recapitalizing the banks (and insurance companies) would not be a sufficient step to restore growth. The banking problems reflect the poor conditions of their borrowers. Putting capital into banks to make up for past losses would be pointless if the underlying corporate problems are not addressed. As Caballero, Hoshi, and Kashyap (2002) emphasize, the growth problems today cannot be due solely to a lack of solvent financial institutions. There have always been international banks (and insurance companies) operating in Japan, and the number rose substantially as a result of the so-called "Big Bang" deregulation that was completed in April of last year. These foreign firms are solvent but are choosing not to lend much in Japan. So the problem is not just that the domestic financial institutions are undercapitalized. This is important because it suggests that merely throwing money at the banks will not resolve the crisis.

Determining the appropriate policies to address the problems is difficult. Bank of Japan officials have often alleged that monetary policy is impotent because of the banking problems (see, for example, Hayami, 2002). While it is true that standard open market operations will not be stimulative if banks will not lend, this by no means impairs other types of monetary policy actions. For instance, the proposal by Svensson (2001) for the Bank of Japan to stimulate the economy through foreign exchange intervention is in no way compromised by the banking problems.

On the other hand, without a functioning system of financial intermediation, there are limits as to how successful nontraditional monetary policy actions will be. As growth resumes, government money will be

needed to combat some of the insolvencies that are hampering normal financial intermediation. There is a wide range of estimates of the degree of insolvency in the banking industry. But, even without settling the issue of how much it would cost to rehabilitate the banks, it is possible to identify many rescue arrangements that would be counterproductive in virtually all potential scenarios. Later in this article, I highlight many of these poor choices and give some loose bounds on the costs of better alternatives.

The second major macroeconomic problem has been the deflation that has accompanied the slow growth. As I explain below, the deflation has played a central role in the problems of the insurance companies. Besides this well-documented and widely discussed effect, the deflation independently has had three pernicious effects on the banking sector.

First, as stressed by Fukao (2003), the deflation squeezes banks profitability. Since nominal interest rates cannot go below zero, there is a floor on the cost of the banks' funds. Even with zero interest rates, depositors may be getting higher real returns than the banks would like to pay. But the banks face competition in lending and, consequently, limits on how much they can charge their customers. With falling prices, banks find it difficult to charge more than 1 percent or 2 percent interest on their loans (since the inflation-adjusted interest burden is much higher). With deflation, the gap between funding costs and lending rates is not sufficient for the banks to make money. If the inflation rate were positive, the banks would have more room to maneuver.

The low nominal rates that are charged to bank borrowers also complicate the problem of regulating the banks. With near-zero interest rates, almost all borrowers can make their required interest payments. Only when a loan matures, and the principal is due, can one gauge the health of the borrower. Since regulators are not necessarily privy to the negotiations that accompany a loan renewal, it can be difficult for them to spot the problem borrowers. Japanese lenders often allow borrowers with no hope of repayment to continue to operate (see Peek and Rosengren, 2002). If interest rates were 3 percent or 4 percent higher, then many of these "zombie" borrowers would soon be unable to service their debts. The regulators would then be able to easily spot the deadbeat borrowers and pressure the banks to cut them off, before more money is lost.

Finally, the deflation has meant that borrowers who took out long-term loans at historically low rates of interest (3 percent or 4 percent) have seen the inflation-adjusted burden of their debt grow. This is the converse of the more typical phenomenon, whereby borrowers benefit from unexpected jumps in inflation at the expense of lenders. One of the clear benefits from a more expansionary monetary policy would be to reverse the increasing debt burdens.

Banking sector problems

I begin the sector-specific analysis by analyzing the condition of the Japanese commercial banks. The most thorough, up-to-date analysis of banks available in English is Fukao (2003). Panel A in table 1 reproduces his key figures. As he stresses, the banking industry has not had a *net operating* profit since fiscal year 1993 (table 1, row G). Until late in the decade, the banks offset these losses by recognizing capital gains on long-held stocks and land. But at this point, there is little more that can be squeezed from these sources. As table 1, row I shows, since 1995 the banks have recorded net losses in more years than not. Cumulating the loan loss figures in table 1 (row F) shows that the banks have recorded losses of roughly ¥83 trillion (16.5 percent of current Japanese GDP) since 1992.[1] According to Japan's Financial Services Agency (FSA), this includes over ¥32 trillion in outright write-offs! Yet the losses are expected to continue for the foreseeable future.

As noted earlier, these losses are too large and persistent to be blamed solely on the rapid decline in asset prices at the beginning of the 1990s. Indeed, as the Bank of Japan (2002b) has pointed out, since 1990 the banks have disposed of more than ¥90 trillion, which amounts to 80 percent of the increase in loans between 1986 and 1990. Thus, it is implausible to suggest that the continued losses can be attributed to misguided lending decisions during the late 1980s. Rather, they are indicative of deeper underlying problems facing the financial services industry.

There are two complementary ways to analyze the banks' current problems that ultimately lead to similar solutions about what might be done to reverse their decline. One focuses on the banks' current costs and revenue structure, while the other looks at the economic forces operating in the industry.

Flow profitability problems

The first approach puts the emphasis on the failure of the banks to generate enough revenue on their loans and other assets to cover their funding and operating costs. To put this in perspective, the second panel of table 1 reports U.S. data that are roughly comparable to Fukao's data for Japan.[2] Despite the data limitations, the comparison clearly shows that the Japanese banks suffer from several structural problems. One is the lack of profitability of their lending operations. The Japanese banks' interest margin relative to assets has hovered around 120 basis points. The U.S. figures (which include both fees associated with the loans and

TABLE 1

Profitability of Japanese and U.S. banks

A. Japan (trillion yen, except last three rows)

	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001
A. Interest income – interest expense	7.1	8.9	9.8	9.2	9.7	10.8	10.7	10.0	9.6	9.7	9.4	9.8
B. Other revenue[a]	2.6	2.2	2.5	2.8	2.1	3.3	3.7	3.6	3.1	2.5	3.0	3.1
C. Operating costs	7.1	7.5	7.7	7.7	7.8	7.8	8.0	8.0	7.5	7.3	7.1	7.0
D. Salaries and wages	3.7	3.9	4.0	4.0	4.0	4.0	4.0	4.0	3.6	3.5	3.4	3.2
E. Gross profit = A + B – C	2.6	3.5	4.5	4.3	4.0	6.3	6.4	5.6	5.2	4.9	5.3	5.9
F. Loan losses	0.8	1.0	2.0	4.6	6.2	13.3	7.3	13.5	13.5	6.3	6.6	9.4
G. Net operating profit = E – F	1.8	2.5	2.5	-0.4	-2.2	-7.0	-1.0	-7.9	-8.3	-1.4	-1.3	-3.5
H. Realized capital gains[b]	2.0	0.7	0.0	2.0	3.2	4.4	1.2	3.6	1.4	3.8	1.4	-2.4
I. Net profit = G + H	3.8	3.3	2.5	1.7	1.0	-2.6	0.2	-4.2	-6.9	2.3	0.1	-5.9
J. Assets	927.6	914.4	859.5	849.8	845.0	848.2	856.0	848.0	759.7	737.2	804.3	772.0
Outstanding loans[c]	424.3	445.8	460.3	472.3	477.8	482.7	482.3	477.9	472.6	463.4	456.9	465.0
Return on assets (I/J)	0.0041	0.00360	0.0029	0.0020	0.0012	-0.0031	0.0002	-0.0050	-0.0090	0.0031	0.0001	-0.0076
Labor costs/operating costs (D/C)	0.52	0.52	0.52	0.52	0.51	0.51	0.50	0.50	0.48	0.48	0.48	0.46
(Interest income – interest expense)/ assets (A/J)	0.0076	0.0097	0.0114	0.0108	0.0115	0.0127	0.0125	0.0118	0.0126	0.0132	0.0117	0.0127
(Interest income – interest expense)/ total income = A/(A + B)	0.7320	0.8018	0.7967	0.7667	0.8220	0.7660	0.7406	0.7353	0.7559	0.8017	0.7581	0.7597

B. U.S. (millions of dollars, except last three rows)

A. Interest income – interest expense	114,948	121,288	132,872	138,785	145,999	153,483	161,172	172,667	180,601	189,655	200,814	210,801
B. Other revenue[a]	54,759	59,482	65,411	74,706	75,952	81,956	92,515	102,946	121,808	142,238	149,501	153,734
C. Operating costs	115,295	124,233	130,455	139,204	143,637	148,936	159,241	168,339	192,222	201,883	212,728	218,706
D. Salaries and benefits expenses	51,558	52,861	54,588	57,977	60,360	63,129	66,659	71,325	78,533	84,877	87,817	91,862
E. Gross profit = A + B – C	54,412	56,537	67,828	74,287	78,314	86,503	94,446	107,274	110,187	130,010	137,587	145,829
F. Loan losses (provisions)	31,953	34,158	26,061	16,753	10,892	12,411	15,483	18,913	21,218	20,758	27,796	41,008
G. Net operating profit = E – F	22,459	22,379	41,767	57,534	67,422	74,092	78,963	88,361	88,969	109,252	109,791	104,821
H. Realized capital gains	483	2,971	4,001	3,055	(558)	530	1,108	1,836	3,119	179	(2,293)	4,434
I. Net profit (before taxes) = G + H	22,942	25,350	45,768	60,589	66,864	74,622	80,071	90,197	92,088	109,431	107,498	109,255
J. Assets	3,378,859	3,420,481	3,496,120	3,695,838	3,999,354	4,299,278	4,554,234	4,989,642	5,410,923	5,690,193	6,152,551	6,454,543
Outstanding loans	2,045,822	1,989,229	1,969,920	2,088,626	2,296,944	2,539,682	2,736,615	2,895,082	3,156,861	3,398,030	3,704,686	3,591,147
Pre-tax return on assets (I/J)	0.0068	0.0074	0.0131	0.0164	0.0167	0.0174	0.0176	0.0181	0.0170	0.0192	0.0175	0.0169
Labor costs/operating costs (D/C)	0.4472	0.4255	0.4184	0.4165	0.4202	0.4239	0.4186	0.4237	0.4086	0.4204	0.4128	0.4200
(Interest income – interest expense)/ assets = A/J	0.0340	0.0355	0.0380	0.0376	0.0365	0.0357	0.0354	0.0346	0.0334	0.0333	0.0333	0.0326
(Interest income – interest expense)/ total income = A/(A + B)	0.6773	0.6710	0.6701	0.6501	0.6578	0.6519	0.6353	0.6265	0.5972	0.5714	0.5732	0.5783

[a]Includes all other profit, such as trading for own account and fees, but excludes capital gains realized from stock and real estate sales (which are in row H).

[b]From sale of stocks and real estate.

[c]Domestic banks only.

Notes: Financial statements of all commercial banks. Data are for fiscal years, which end in March of following calendar year.

Sources: Panel A, Fukao (2003); panel B, call reports, and author's calculations.

interest revenue) are roughly three times as high—far too big a difference to be attributable to the differences in measurement.

Fukao notes that if the deflation were to stop, then the banks could raise nominal interest rates without raising the real interest burden for borrowers. But there are limits to how much this could be expected to help. For instance, assuming, optimistically, that when the deflation ends the banks could raise their interest margin by 1 percentage point (say by increasing lending rates by 2 percent and deposit rates by 1 percent), this would add only another ¥5 trillion in interest margin. While this might be enough to stop the banks' losses, they would still be far less profitable than their U.S. counterparts.

Another problem is the Japanese banks' high labor costs. The banks have made some progress in reducing salary and wage expenses from roughly 52 percent of operating costs to 46 percent. Although anecdotal reports of overpaid and underutilized bank staff still abound, the Japanese banks likely will have to increase pay to some workers if they want to upgrade competency levels in order to increase fee- and commission-based income. It is doubtful therefore that the Japanese banks can push their salary expenses all the way down to the U.S. level of about 42 percent of operating costs.

Finally, while not evident in the table, it is also well known that the Japanese banks underinvested in technology during the last half of the 1990s. This has long been recognized as a problem. For example, although a condition of government-provided funds offered to the banks in 1999 was that the banks had to improve efficiency and reduce costs, the general cutbacks in investment were not to be extended to investment in computing and automation. Still, more than three years later, the concerns about poor computing operations persist.

The failure of the Mizuho Group's computers that occurred on the first day that the bank began operating could hardly have been more symbolic. Due to poor integration of the three merging banks' antiquated systems, the new bank's computers failed. As a result, the ATM network was unavailable, a number of automatic payments were not made, and many customers were double-billed for credit card transactions.[3] The Bank of Japan subsequently had to order Mizuho to upgrade its computing systems.

Fukao reports that the main funds payment system used by Japanese banks (*zengin*) is unable to handle two-byte codes, and hence cannot transmit customer names and messages in *kanji* (characters). This is one of the reasons why convenience stores (that have typically installed more sophisticated technology) have won customers that would like to make an occasional electronic payment at the banks' expense. The *zengin* system is scheduled for an upgrade in April 2003, but the banks will have to do much more if they want to match the technological efficiency of many of their global competitors.

Japanese banks' limited comparative advantage

The alternative (and complementary) approach to analyzing the banks' profitability problems is to look at their product mix and ask which lines of business can be expected to earn normal rates of return? The Japanese banks are among the largest in the world in terms of assets. For instance, Agosta (2002) reports that the Mizuho Group and Mitsubishi Tokyo Financial Group are the first and third largest in the world, respectively, and that 19 of the largest 100 banks in the world are Japanese. Yet, there are few if any product lines for which the Japanese banks are world leaders. I find no examples where Japanese banks and their global rivals have competed for business on a level playing field and the Japanese banks have emerged as market leaders. Instead, the recurring pattern is that Japanese banks are later to enter markets or offer new products and, consequently, their profitability lags.

The low levels of fee income alluded to earlier are a particularly important reflection of this problem. As Hoshi and Kashyap (2001) note, for the Japanese banks in aggregate, fee and commission income as a percentage of total income was essentially identical in 1976 and 1996; the U.S. banks during this period increased their percentage of fee and commission income by two-and-a-half times. This disparity partially was attributable to regulation that handicapped the Japanese banks. For instance, until 1998 the banks were simply barred from many activities, such as over-the-counter derivatives transactions, brokerage activities, and underwriting.

But even after the full deregulation that was completed on April 1, 2001, the gap persists. The last row in panel A of table 1 shows that the Japanese banks continue to make roughly 76 percent of their income (excluding capital gains) from their lending operations. In contrast, U.S. banks make only 58 percent of their income from this low margin activity; instead they bring in a much higher percentage of high margin fee and commission businesses.[4] Since these nontraditional products and the associated revenue streams are central to the business strategies of most global banks, this deficiency is a huge problem for the Japanese banks. Without making comparable profits in these areas, it is hard to see how the Japanese banks could ever reach the same rates of return as their competitors.

One way to address this problem would be for Japan's major banks to scale back on their operations

and to focus on niche needs of Japanese customers (mostly small and medium-sized businesses). Japanese banks might arguably be better than foreign banks operating in Japan in this product line. The loan demand of these customers, however, is much lower than the assets of the current banking system; therefore, shifting in this direction in order to raise profitability would imply considerable downsizing. But downsizing would involve the release of many mid-career and upper level managers, who might face significant hurdles in becoming reemployed.

A final, further impediment to the banks' profitability is the difficult competition that they face from subsidized government financial institutions. The postal savings system poses a particularly big problem. As Fukao asks, how can the private banks make profits when Japan's government-sponsored postal savings system has 40 times the number of offices of the largest banking group, pays roughly the same rate on deposits as the banks, and charges no maintenance fees? The extra convenience of the postal accounts, combined with the government guarantee of deposits, represents a major challenge for the banks.

The government-subsidized Housing Loan Corporation (HLC) also compromises the banks' ability to make money through home mortgage lending. The HLC receives subsidies (as described below) from the government and passes these savings on to their customers. The HLC makes about 40 percent of all home mortgages. Fukao (2003, table 8) shows that the HLC loans have rates that are substantially lower than those offered by private banks, despite typically having longer maturities. Moreover, the HLC loans come with no prepayment penalties (unlike typical Japanese bank mortgages).

These kinds of government-sponsored financial institutions will have to be reined in if Japanese banks are to regain profitability. This is widely recognized outside Japan. For instance, the Bank for International Settlements in its 2001 annual report (2002, p. 133) notes that one of the contributing factors to the banks' profitability problems is the "strong competition from government sponsored financial institutions." The IMF 2002 country report goes further and says that the (p. 3) "exit of non-viable banks and a scaled down role of government financial intermediation are necessary to improve bank profitability."

Yet, despite making it a priority to privatize the postal savings system and otherwise reform many government agencies, the Japanese government has encountered strong resistance to its efforts to address this problem. The postal savings system and the government's home lending program are popular with the public. Furthermore, the public has not been convinced that these programs in fact are contributing to the banking troubles. Given the public support, and the role that the postal savings system plays in the Fiscal Investment Loan Program (described below), it is not too surprising that many politicians have fought the Koizumi administration's reform efforts, delaying a full-fledged attack on the banking problems. However, without some adjustments to these reform programs, the banks' problems are likely to reappear even if they were to regain solvency.

How much would recapitalization cost?

Assuming that the banks could figure out how to resume making profits, the next obvious question is how much would it cost to make the banks solvent? I review first the three main problems that plague attempts to arrive at an estimate, before reporting the range of estimates currently made by market participants.

The first problem with this type of exercise is determining the current level of losses associated with existing loans. The banks in Japan are known for their propensity to under-reserve against *recognized* bad loans. For instance, they have set aside reserves sufficient to cover between 40 percent and 60 percent of bad loans over the last few years, whereas U.S. banks tend to hold closer to 160 percent in reserves (Fukao, 2003). By Fukao's estimate, the banks are currently short at least ¥7 trillion in loan loss provisions.

Then there is the larger problem of deciding how many additional loans are in fact already bad, but have not been revealed as such. Almost all analysts agree that there are many more bad loans than the banks have acknowledged. But there is considerable disagreement over the size of the under-reporting. For instance, Credit Suisse First Boston analysts estimate the ratio of problems loans for the seven major banks to total loans to be just about 27 percent, roughly four times the disclosed figure.[5] Meanwhile, Goldman Sachs estimates that all bad debts (for the entire banking system) are three times as high, ¥236.6 trillion (38.1 percent of all system loans)![6]

Translating the figures on nonperforming loans into estimates of taxpayer exposure requires a further step of netting out collateral and other bank reserves. But carrying out this netting is challenging when the underlying environment is still unstable and the reported levels of problem loans keep rising. The ratio of nonperforming to total loans has been steadily rising among the smaller banks in Japan. A first sign that disclosed losses are catching up to actual bad loans will be when the ratio of nonperforming loans to total loans levels off. In the meantime, much of the discrepancies in estimates across analysts arise because of different assumptions about under-reporting (and the methods used to net the losses against other assets).

A second issue is the quality of the other parts of the banks' balance sheets. Two items in particular are treated in ways that overstate the apparent health of the banks. One is that bank *capital* is permitted to include tax credits against future profits. The figures for the largest banks suggest that about 35 percent of shareholders' equity is made up of these deferred tax credits for loan losses.[7] But for these credits to be of any value, the banks must quickly regain profitability once the loan losses are recognized: Tax loss credits expire five years after the bad loans are actually worked out, so many of the existing credits will expire before they can be claimed.[8]

Another problem is that the banks hold a significant amount of insurance company debt (usually in the form of subordinated loans or surplus notes). As I discuss in the next section, the life insurance companies also tend to hold large amounts of subordinated bank debt and stock. Many of the life insurance companies are also in a very precarious financial position. This "double gearing" makes the banks and the insurance companies each look to be better capitalized than is in fact the case.[9]

The ownership of life insurance securities is also part of the broader tendency for Japanese banks to own corporate equities. Fukao estimates that as of March 2002, the banks held equities worth roughly ¥34.4 trillion, which was substantially larger than their true capital (by a factor of seven if one accepts Fukao's adjustments to correct for the overstated value of the deferred tax credits, the under-reserving of bad loans, and the preferred shares loans from the last public injection of capital in March 1999). Thus, the banking sector's value is quite sensitive to changes in share prices; based on Fukao's figures, the decline of the Nikkei from 11,025 on March 31 to 9,383 at the close of September 30 would have wiped out all the banks' private equity (assuming they had not bought or sold any in the interim). Accordingly, the size of the mismatch between the value of banks' assets and liabilities at any point in time depends importantly on the level of stock prices at the time.

The Bank of Japan recently announced that it was prepared to buy securities from the banks at market prices. If the banks do accept this offer and sell at prevailing market prices, this policy would have very little short-run impact. The banks would still have to accept any losses that were embedded in their portfolios and doing so would erase their capital by the amount of the losses. The only advantage for the banks would be that they could opt to significantly reduce their equity holdings without necessarily pushing prices down.

Conversely, if the Bank of Japan were to pay a premium for any securities bought from the banks, then the premium would increase the banks' capital. But, the outline for the stock purchasing plan announced by the Bank of Japan (2002c) states that the prices will be at the market prices defined as "the lesser of the volume-weighted average price or the day's closing price." More importantly, the total amount purchased will be limited to ¥2 trillion. Therefore, even if the prices were substantially above the market price, the potential transfer to the banks would be quite limited.

Finally, the amount of the funding needed to eliminate the banking sector's insolvency will depend on the macroeconomy (for all the reasons discussed earlier). The cost to the U.S. taxpayers of the U.S. banking crisis in the early 1990s turned out to be well less than 1 percent of (then current) GDP, because of the phenomenal growth of the economy over the 1990s. It is very unlikely that Japan will experience anything like that during its recovery, but differences of opinion over the likely path of the economy over the near term further contribute to the dispersion of estimates.

Given all these caveats, it should come as no surprise that different observers reach fairly different assessments about the amount of funding that would be required to make the banks solvent. Table 2 shows the estimates of many of the leading economists and bank analysts as of August 2002, collected by direct correspondence with the experts. They were each asked to report their estimates of the difference in the market value of assets and liabilities of the Japanese banks (as of August 1, 2002); as indicated in the table, several of the responses cited previously published estimates of slightly different quantities (for example, the value of all problem loans or losses at major banks only).

The most optimistic figure would suggest losses of less than ¥12 trillion (2.4 percent of Japanese GDP); this would be the case if the baseline ING Securities estimate were adjusted to take care of the phantom tax credits that overstate capital by roughly 35 percent. The Goldman Sachs and Lehman Brothers estimates suggest losses that are roughly three times as high. Regardless of which numbers one believes, it is clear that the losses for the taxpayers will be substantial.

These figures and the foregoing discussion also explain why Takenaka's Financial Sector Emergency Response Project Team posed such a threat to the opponents of reform. The task force's initial recommendations were reported to have centered on reducing the length of time that could be used to claim tax credits as part of banks' capital, tightening loan assessment standards, and forcing increased provisioning for losses. Significant changes in any of these directions would severely impact banks' capital and likely could push some (or nearly all) of the major banks below the mandated

TABLE 2

Experts' estimates of the insolvency of the Japanese banking system

Analyst	Firm (date of estimate)	Estimate	Comments
David Atkinson	Goldman Sachs (October 31, 2001)	¥70 trillion of net loan losses based on March 2001 loans (¥18.7 trillion for the major banks)	Large bank losses represent 161% of capital adjusted for tax loss carry forwards and public money.
Robert Feldman	Morgan Stanley (August 2002)	¥22 trillion	Intended to be a lower bound for additional taxpayer exposure.
James Fiorillo	ING Securities (Japan) (August 2002)	¥19.9 trillion in net loan losses, –¥2 trillion in unrealized capital gains	Capital (as reported without adjustments) ¥l6.2 trillion
Yukiko Ohara	Credit Suisse First Boston Securities (Japan) Limited (July 2002)	¥21.8 trillion in required credit costs for the major banks	Estimated non-performing loans for the major banks ¥121.9 trillion
Paul Sheard	Lehman Brothers (August 2002)	"To restore the balance sheet health and credibility of the banking system would probably require ¥30 to ¥50 trillion."	Notes that the deposit insurance fund has ¥49 trillion of untapped capacity. Thus, infrastructure and budgeting are in place to act if there were political will.
Reiko Toritani	Fitch Ratings (August 2002)	¥23 trillion for the major banks	Adjusting the stated value of equity for the major banks as of March 2002 to account for fictitious tax credits, public funds, and unrealized gains implies a market value of essentially zero.

level of capital. It is not surprising, therefore, that this possibility triggered intense criticism of Mr. Takenaka and his plan.

But, as the estimates in table 2 show, regardless of whether the capital deficiency is recognized by the regulators and acknowledged to the public, the private sector analysts are unanimously of the view that the banks are bankrupt—by a significant amount. This suggests that barring a miraculous economic recovery that no one is forecasting at this time, the banks will eventually be forced either to close or to raise more capital.

This conclusion leads to two criteria that can be used to judge policy proposals regarding bank recapitalization. First, it may be helpful to distinguish between proposals that do and do not facilitate the downsizing and consolidation of the banking sector. If one accepts the earlier analysis, it is quite likely that the road to profitability will come through focusing on more profitable activities and shedding assets. Under this view, the total level of capital to be committed to the industry should be determined by the level needed to support the long-run size of the industry, not necessarily its current size.

Second, since money to bail out the banks is limited, any refinancing proposed should be done in a focused fashion. In particular, if exit of some banks is inevitable, then it is poor policy to prop up banks that will soon go out of business. Past recapitalizations in Japan did not adhere to this rule, but featured across-the-board rescues, whereby some of the money was wasted on dying banks.

These mistakes could be avoided if more market signals were used to decide which banks merited funding. Banks that cannot attract private financing as part of their recapitalization might be given lower priority than those that can. This type of selective rehabilitation would lead to the best banks being rebuilt. The resulting banking sector would be more efficient at directing funds to deserving borrowers.

Problems with the life insurance sector

The life insurance companies comprise the second largest part of the financial system. As of March 2002, the ten major private insurance companies had assets of roughly ¥150 trillion (30 percent of GDP). Most insurers are mutual companies so that their shares are not traded on exchanges, but as explained earlier their financial linkages with the rest of the financial system are extensive. For instance, at least 10 percent of the equity of each of the major Japanese "city banks" (that is, those that are large and globally active) is owned

by insurance companies; as of March 2001, insurance companies owned ¥5.4 trillion of bank equity and ¥5.1 trillion of subordinated bank debt. Thus, it is necessary to recognize that the health of the insurers is intimately connected with that of the banks.

Similarities with the banks

The problems of the life insurance companies resemble those of the banks in three respects. First, they too have made bad loans. However, the scale of the insurers' lending mistakes is quite different. As of March 2002, the ten majors had disclosed ¥568 billion in loans to distressed firms. This amounts to less than 2 percent of their total loans. Moreover, they had reserves against these loans of over 70 percent.[10] Thus, even if there is substantial under-reporting of the problem loans, the bottom line of the insurers is much less likely to be affected.

Second, the insurers have very significant exposure to the changes in the aggregate stock market. The Daiwa Institute of Research (DIR) produces company by company estimates of the levels of the stock market at which unrealized gains on securities disappear.[11] The critical value of the Nikkei 225 stock index for the different firms is between 8,400 and 12,500, with an average of 10,880. DIR estimates that as of March 31, 2002, when the Nikkei was at 11,024.94, the aggregate unrealized gain on stocks was approximately ¥1.88 trillion. Fitch makes a similar calculation using cutoff values for the Tokyo Stock Price Index (TOPIX). With the TOPIX at 903 (as of October 1), the Fitch estimates imply that nine of the ten major insurers would have unrealized losses in their equity portfolios.[12]

Third, the insurance companies also face acute competition from the government-sponsored financial institutions. In their case, the key competitor is the postal life insurance program. The postal insurance program sells about one-third of the life insurance in Japan. While the same convenience advantage accrues to the postal insurance program as to the postal savings program, the pricing of the insurance does not seem to be as distortionary. However, the premiums paid into the postal insurance accounts are largely recycled through the fiscal investment loan program, as described below.

Excessively optimistic estimates of returns

Despite these similarities to the challenges facing the banks, the fundamental profitability problem for the private insurers is unique and largely self-induced. Primarily, they have been crippled by their overly optimistic assessment of anticipated investment returns. For instance, as of 1992, the life insurance companies were all selling lifelong annuities that promised to pay a return of 5.5 percent. As interest rates fell, a gap opened between what the insurers had promised to pay and what they could expect to earn. This difference is referred to in the insurance industry as the "negative carry" (or "spread").

As of March 2002, the insurers had a disclosed negative carry of ¥1.25 trillion. This flow loss can be compared with the profits of roughly ¥3.3 trillion from the other parts of their business. Because the disclosed carry omits unrealized capital losses, these figures are likely to provide an overly optimistic reading of the firms' health.

The regulatory assessment of the industry is based on the concept of a solvency ratio, which is intended to measure the extra capital that insurers are expected to hold in order to make good on their promised payouts. The formula for calculating the margin is complicated and involves estimating the risks from insurance underwriting, interest rates, asset management, and business administration and then comparing the risk with the insurer's ability to pay (based on the quality of its assets).[13] Insurance companies around the world are measured by this yardstick, and since 1999 Japanese insurers have been subject to prompt corrective action whenever their solvency margin fell below 200 percent. The ten major insurers all reported solvency margins in excess of 400 percent as of March 2002.

As with the banks, there are dramatic differences between the officially reported solvency margins and more realistic estimates. Fukao (2003) highlights three problems with the standards used in calculating Japanese solvency margins (compared with practices in the U.S.). First, the Japanese supervisors use lower risk weights than in the U.S. Second, the ability to pay is inflated by including assets that have no liquidation value. Finally, the ability to pay ignores unrealized capital gains and losses.

Fukao finds that making these corrections to move the Japanese figures toward the U.S. standard dramatically lowers the margins. Using March 2001 data, when the Nikkei 225 average was roughly 13,000, his estimates show that three companies' ratings (Mitsui Life, Asahi Life, and Sumitomo Life) were all below the critical level of 200. As reference, the official ratios for all three were in excess of 490.

Ex post, Fukao's adjustments seem to do a good job of predicting which firms will fail. All the major life insurers that have been distressed since mid-2000 (three that went bankrupt and one that required a significant equity injection by a foreign partner) showed similarly low adjusted-solvency margins. With the Nikkei now markedly lower than its March 2001 level, it is likely that the next weakest surviving firms (Yasuda, Meiji, and Daiichi) are also near the threshold.

Bailing out the life insurance companies

With the banks there is a strong presumption that depositors must be protected (to prevent runs and contagion). No similar argument can be made for insurance companies. The insurers would be viable if they simply had more realistic promised payout rates. The obvious solution is to have the companies declare bankruptcy and force the policyholders to take a reduced rate on their investments.

Political economy concerns regarding an insurance company write-down should be less of a problem than if a similar remedy were proposed for the banks (that is, forcing the bank depositors to absorb the banks' losses), the important difference being that the policyholders taking a haircut would be the ones that had benefited from the overly generous payments. For the banks, the depositors generally did not receive the loans that are now unrecoverable. This difference probably explains why so many prominent insurance companies have actually declared bankruptcy: eight major ones between April 1997 and April 2002. In contrast, only three of the large banks failed during this time.

The fact that the fundamental problem is so easily diagnosed is further confirmed by the behavior of foreign competitors. Since the Big Bang deregulation, foreign firms have raced in to partner with the insurers. For instance, Hoshi and Kashyap (1999) note that seven of the life insurance companies entered into significant alliances with foreign companies in 1998 and early 1999. This outnumbered the number of deals by the banks (six), despite the fact that there were vastly more banks looking for partners. Instead, the banks mostly worked out deals with other Japanese firms. One interpretation of this contrast is that the insurers are judged by potential investors as having much more underlying value than the banks.

Collectively, these observations can be summarized in three propositions. First, the continued functioning of the insurance markets does not require that the government put money into the sector. Second, there do not seem to be overwhelming political constraints that prevent market solutions (that is, bankruptcy) from working. Third, foreign firms see some underlying value in the sector, so that perhaps entry or acquisitions can be expected if the bankruptcies continue.

Thus, viewed in isolation, there seems to be no reason to provide public money to rescue the insurers. The weaker ones would be expected to fail, but the fallout from this would be limited. The fact that the Life Insurance Policyholders' Protection Corporation of Japan (the government bailout fund) is broke from past payouts reinforces this possibility.

However, the banks and insurance companies are linked through their double-gearing. Fukao estimates that the banks hold roughly ¥2.2 trillion of the surplus notes and subordinated debt of the insurers. Whenever the insurance losses are recognized, the banks will have to take their share of these losses. Unless the government were to purchase these securities as part of the bank clean-up, the banks would be at risk for requiring more capital. At the very least, planning how to decouple the two should begin and the banks and insurance companies should be encouraged to work to sever their linkages.

Most outside analysts take it for granted that the double-gearing is dangerous. For example, the Bank for International Settlements' 2001 annual report strongly criticized this practice (2002, p. 135), saying "these interlinkages increase systemic risk, particularly considering the weaknesses in the Japanese insurance sector."

However, Japanese government policymakers do not seem to recognize these risks. For instance, Shokichi Takagi, Commissioner of the FSA, responded to the BIS criticism saying that "the nature of the risks is different between insurance companies and banks. … As far as the conventional approach is concerned, the nature of risks is different and there[fore] the cross-holding of equity is not a big deal. So-called double-gearing is not excluded at this point, as the nature of the risks is different."[14] Thus, one big impediment to addressing this issue is the regulatory stance of the FSA.

Government-sponsored agencies

Finally, I consider the impact of the government financial institutions. These organizations engage in a host of activities, ranging from offering home mortgages and providing life insurance and savings accounts to financing highway development. They are relevant to our discussion for two different reasons. The first is that many of these agencies are losing money and will ultimately require a taxpayer bailout. The money that will be spent here constrains the funding that is available for the insurance and banking restructuring. Second, these agencies' losses are often related to operating practices that limit the viability of complementary private sector firms. Thus, one important public policy issue is whether these government agencies should continue to compete with the private sector firms.

Gauging the size of taxpayer exposure is very complicated, since financial disclosure is poor and many of the assets of these institutions are obligations of other government institutions. A typical transaction starts with a home loan extended by the HLC. The HLC would raise the money to provide the loan by issuing debt that is bought by the public and other government

financial institutions. Thus, determining the full taxpayer exposure will potentially involve looking at the financial condition of several organizations.

To do this systematically, I rely on the recent work of Doi and Hoshi (2003). They focus on the financial condition of the fiscal investment loan program (FILP). The FILP is often called the Japanese government's second general account budget. Historically, most of the money in this program was collected from people's deposits in the postal savings program. The ubiquitous branching system of the Post Office, combined with branching and other restrictions that prevailed until recently for commercial banks, led many Japanese to keep their wealth in postal savings accounts. The money in these accounts was then turned over to the Trust Fund Bureau of the Ministry of Finance (MOF) and loaned out as MOF officials saw fit through the FILP. The ability to direct funds to favored projects, which are not easily monitored, makes this process very convenient for political purposes.

Thus, one can think of the FILP money as funding the financial institutions, as well as providing significant money to local governments and many other programs. Importantly, these programs are not integrated with the central government's budget, so that the obligations for these programs are not part of the government's gross debt. In total, roughly ¥418 trillion (84 percent of GDP) flowed through the system during the fiscal year ending in March 2001. By assessing the health of the FILP-dependent borrowers, we can not only learn about the condition of key government sponsored financial agencies, but also about the other hidden losses that may be handed to the taxpayers.

In parsing the figures, it is instructive to separate the condition of the financial institutions and other special purpose agencies (that I collectively refer to as the FILP agencies) from those of the local governments. The two differ both in the nature of the accounting information that is available and in the role that they play in the economy. This leads to different levels of confidence in our estimates of losses and potentially different public policy implications. Thus, I follow Doi and Hoshi and report separate estimates.

Quantifying losses for the FILP agencies

To see the problems for FILP agencies, we can revisit the HLC example described above. If all the underlying assets are solid (in this case the assets associated with the property loan), then the intermediate transactions are irrelevant. The HLC debt can be repaid using the proceeds from the loan and this means that the government financial institutions that bought the debt can pay back their depositors. In other words, the relatively low net position of the government is what matters.

If the HLC loan is not performing, then the situation becomes more complicated. In this case, the HLC debt will not be fully paid with the proceeds from the loan. But it is unlikely that the government will default on the HLC bonds, so new funds must be raised to pay the bond holders (and ultimately the depositors of financial institutions that bought the debt). Effectively this means that the gross amount of debt (that owed by the HLC and the government financial institutions) is the relevant figure for determining the government's obligations.

The quantitative gap between the gross and net figures for Japanese government debt is huge. Japan has the highest level of gross debt relative to GDP of the G-10 (Group of Ten) countries and the lowest level of net debt. Thus, one's perspective on the quantitative importance of any FILP losses (which are outside of the official debt calculation) requires further judgment about the quality of the central government's assets. A full analysis of the entire budget is beyond the scope of this article, so I will tackle the narrower question of the FILP losses, which turns on the asset quality in the FILP transactions.

Doi and Hoshi point to three recurring problems that suggest asset quality is low. First, there are three cases (most notably the HLC) where past losses are recorded on the agency books as an asset. The agencies rationalize this by arguing that the losses were sudden and it would be misleading to immediately recognize them; instead they are expected to slowly eliminate these losses by reducing their capital. To correct for this inaccurate reporting, the first step in the analysis is to immediately count the losses. By doing this, Doi and Hoshi write down the capital of these three agencies by ¥0.5186 trillion.

A second more widespread problem is that many agencies acknowledge that their loan losses exceed their reserves. Doi and Hoshi estimate that this practice is employed by 22 of the 58 recipients of FILP funds. In total, they estimate ¥8.2 trillion in recognized bad loans have yet to be provisioned for. Of course, there is the additional problem that there are likely to be many more bad loans that have yet to be uncovered.

A third pervasive problem is the overvaluation of physical assets. The Public Highway Corporation and several other agencies do not properly account for depreciation. Instead, depreciation of assets is only recorded when operational revenues are high enough to count the depreciation and still show "profits" on the financial statements. A related problem is that the value of long-term assets is generally based on the historical acquisition costs. For land purchased in the 1980s, this will greatly overstate the current market value.

Doi and Hoshi attempt to correct for market value changes and depreciation of the 12 FILP agencies that have a high percentage of physical assets relative to total assets. (These turn out to be agencies that are involved in urban development or infrastructure provision.) It appears that losses of roughly ¥11.4 trillion are uncovered once these corrections are made.

Finally, there is the problem that many FILP agencies are making flow losses that need to be covered by taxpayers. Since fiscal year 1999, the agencies have been required to make a discounted present value calculation of the gap between their revenues and costs. Of the 33 agencies that report the figures for March 2001, 28 expected costs to exceed revenues. Moreover, Kikkawa et al. (2000) find that agencies have been extremely optimistic in their revenue forecasts. The March 2001 estimates suggest that net losses will total ¥11.7 trillion, and this is certainly a lower bound on the likely losses.

Doi and Hoshi do a careful agency by agency calculation of how all of the aforementioned problems will affect taxpayers. By disaggregating in this way, they can allocate any insolvency that is present in the agencies to the government and any other stakeholders. Moreover, they compare losses to the amount of capital that is already on the books to figure out how much more money will have to be provided. They arrive at a (intentionally conservative) cumulative estimate of ¥35.8 trillion (7 percent of GDP) for the taxpayer exposure from the operations of the FILP agencies.

Other FILP losses

However, the full taxpayer bill also depends on the other non-agency loans. As of March 2001, about ¥87 trillion of FILP funding was steered to local governments. Assessing the quality of these loans is difficult since local governments are not required to produce balance sheets or other financial statements that would allow a direct estimate of the quality. However, the fact that many local governments have substantial debts and are running very small surpluses (or outright deficits) suggests that some default on the debt is possible.

Doi and Hoshi run a variety of simulations to assess the local governments' ability to pay versus their debt levels. The simulations differ according to the assumptions that are made about the growth rates of future deficits and tax revenues. The locals had FILP obligations of ¥125.5 trillion as of March 2001. The resulting estimates of the size of the losses borne by taxpayers cluster between ¥30 trillion and ¥40 trillion—importantly, this accounts for the fact that the FILP is not the only creditor of the bankrupt governments and nets out all collateral that is available.

Implications for government financial institutions

Combining all the estimated FILP losses, Doi and Hoshi's preferred estimate of likely FILP losses is ¥78.3 trillion (just over 15 percent of GDP). The sheer size of these potential losses no doubt makes politicians hesitant to publicly acknowledge them. However, without building some public recognition of the losses, it will be difficult to undertake fully the necessary reforms. In the meantime there are several intermediate steps that would be useful.

One goal would be to stem taxpayer losses by reducing the flow of FILP money to insolvent borrowers. A FILP reform was enacted in April 2001 that *could* lead to this outcome. As part of the reform, government agencies were supposed to increase their funding through public bond issuance rather than relying on captive FILP financing. The reform, however, was inadequate in two respects. First, it provided a generous transition period during which money could continue to flow as it had in the past. Second, it did not contain any provisions for shutting down money-losing public corporations. Without such provisions, market discipline cannot take hold. Indeed, Doi and Hoshi find that the flow of funds through the FILP has not changed much.

Another goal is to limit the distortions for the private sector associated with the continued operation of the money-losing government-sponsored agencies. For instance, the pricing of government loans and deposits could be set to match the rates charged by the private firms. A current proposal to charge for deposit insurance on postal savings accounts would be a useful move in this direction. Another pro-competitive move would be to add prepayment penalties for government-agency loans. The general principle should be that if these kinds of agencies are to continue to operate, they should do so on a level playing field with the private sector.

Conclusion

There are different reasons for the sizable losses lurking in Japan's banking, insurance, and government agency sectors. Yet, the problems in these sectors are inter-related. The banking problems that attract so much attention will persist until the troubles in the other two sectors are also addressed. A satisfactory resolution requires recognition of the different driving factors behind the problems in all three sectors and would include measures that address all at the same time.

The combined effect of all the problems is huge. Representative estimates for the banking problems are roughly ¥40 trillion. I have argued that most of the losses for the insurance companies will not be borne by the taxpayers, but the FILP losses look to be at least

¥78.3 trillion. Thus, Japanese taxpayers are likely on the hook for roughly ¥120 trillion (24 percent of GDP)!

A variety of factors have contributed to the delay in confronting the problems. One huge problem is the government's unwillingness to force the restructuring that will be necessary to create a profitable banking sector. The restructuring will lead to business closures and job losses in the banking sector. Another serious problem is the lack of political will to shut down or restructure the popular, but unprofitable government-sponsored financial agencies. These organizations are especially problematic since they further impair the competitiveness of the private sector.

The recent bail out of Daiei Inc., a large bankrupt grocery store chain, shows how difficult this will be. Daiei had ¥420 billion of its debt restructured in January 2002 by its three major lenders. However, it was soon clear that the restructuring plan was insufficient (since Daiei still had ¥1.7 trillion in debt) and that the banks would need to accept more losses. In October, the Japan Development Bank came forward and offered ¥10 billion as part of a second restructuring plan (that included another ¥50 billion from the private banks). The move was hailed by the government as helping to protect the 96,000 Daiei employees as the restructuring continued.

But this tack is likely to be counterproductive in several respects. One problem is that it sets a bad precedent for future cases. The banks are already routinely rolling over loans rather then pulling the plug on bankrupt firms, because if the banks did recognize the losses they would be at risk for having too little capital and being shut down. Banks will have even less reason to recognize losses and take them when there is the chance that government assistance will be offered.

More importantly, the bailouts (and routine rollovers) that keep the deadbeat borrowers in business also distort competition. Other firms that could enter an industry or gain market share are held back. As Caballero, Hoshi, and Kashyap (2002) explain, suppressing the normal process of creative destruction leaves all banks with fewer good borrowers to lend to. Absent good borrowers, the banks have an even greater incentive to roll over loans to deadbeat borrowers. As the cycle progresses, the firms continue to lose money and increase the banks' losses.

Ironically, therefore, keeping the deadbeats alive likely raises the final costs to the taxpayers. In essence, continuing the lending to firms like Daiei amounts to a covert unemployment compensation program. But continuing to funnel the money through the banks creates other costly distortions. Because this stifles the creation of new jobs, there will be fewer alternatives for the displaced workers and less tax revenue accumulated to cushion the blow when the firm finally fails or is significantly downsized. It would be cheaper and more efficient to end this cycle promptly with a large-scale, comprehensive intervention.

NOTES

[1]Because exchanges rates have varied substantially over the last few years, I have opted not to convert figures into foreign currencies. Japanese nominal GDP has been roughly constant at ¥500 trillion for the last few years so I have normalized other figures relative to this benchmark.

[2]I thank Robert DeYoung for calculating these figures from the U.S. call reports.

[3]On April 1, 2002, the Industrial Bank of Japan, Dai-Ichi Kangyo Bank, and Fuji Bank consummated their merger to form Mizuho Bank, the largest bank in the world. See Associated Press Newswires (2002).

[4]Some of the gap is attributable to the slow development of the syndicated lending market in Japan, since loan syndications move revenue from the form of interest payments to fees.

[5]See Credit Suisse First Boston (2002), figure 8.

[6]See Goldman Sachs (2001), p. 77.

[7]ING Barings (2002).

[8]In principle, a slow winding down of the loan problems would give the banks more time to take advantage of the tax credits. But, as I explain below, stretching out the resolution of the problem is likely to lead to more losses. See Goldman Sachs (2002).

[9]The banks issue securities that are bought by the life insurance companies, which effectively buy the securities by turning over their own securities. The net effect is that reported capital may be increased but the amount of real money raised is greatly overstated.

[10]See table 11 of Merrill Lynch (2002).

[11]See table 6 of Daiwa Institute of Research (2002).

[12]See table 3 of Fitch Ratings (2002).

[13]The exact definition is 200 × (net assets/risk), where net assets are defined as the sum of capital, risk reserves, general loan loss reserves, excess reserves over the surrender value of policies, future profits, subordinated debt (and loans), and a correction for deferred taxes. The risk is the sum of business management risk and the square root of squared insurance risk plus squared interest rate plus asset management risk.

[14]Tagaki (2002).

REFERENCES

Agosta, Veronica, 2002, "Japan's banks are again among world's biggest," *American Banker*, June 18, p. 2.

Associated Press Newswires, 2002, "Mizuho, world's biggest bank, makes rough debut as ATMs crash, customers get double billed," New York, April 5.

Bank for International Settlements, 2002, *BIS 72nd Annual Report*, Basel, Switzerland, July 8.

Bank of Japan, 2002a, "New initiative toward financial system stability," public statement, Tokyo, September 18, available at: www.boj.or.jp/en/seisaku/02/sei0208.htm.

_____**,** 2002b, "Japan's nonperforming loan problem," public statement, Tokyo, October 11, available at www.boj.or.jp/en/seisaku/02/sei0210.htm.

_____**,** 2002c, "(Reference) The outline of the stock purchasing plan," public statement, Tokyo, October 11, available at www.boj.or.jp/en/seisaku/02/sei0209b.htm.

Bernanke, Ben, 2000, "Japanese monetary policy: A case of self-induced paralysis?," in *Japan's Financial Crisis and its Parallels to U.S. Experience*, Ryoichi Mikitani and Adam S. Posen (eds.), Washington, DC: Institute for International Economics.

Caballero, Ricardo J., Takeo Hoshi, and Anil K Kashyap, 2002, "Zombies," University of Chicago, work in progress.

Credit Suisse First Boston, 2002, "Japanese banks," New York, report, May 20.

Daiwa Institute of Research, 2002, "FY01 results of 10 major life insurers," Tokyo, report, June 10.

Doi, Takero, and Takeo Hoshi, 2003, "Paying for the FILP?," in *Structural Impediments to Growth in Japan*, Magnus Blomstrom, Jenny Corbett, Fumio Hayashi, and Anil Kashyap (eds.), Chicago: University of Chicago Press, forthcoming.

Fitch Ratings, 2002, "Japanese life insurers: Price volatility hitting solvency," New York, report, August 27.

Fukao, Mitsuhiro, 2003, "Financial sector profitability and double gearing," in *Structural Impediments to Growth in Japan*, Magnus Blomstrom, Jenny Corbett, Fumio Hayashi, and Anil Kashyap (eds.), Chicago: University of Chicago Press, forthcoming.

Goldman Sachs, 2002, "Bank value depends on slow structural reform," New York, report, April 23.

_____**,** 2001, "Japanese bank asset quality," New York, report, October 31.

Hayami, Masuru, 2002, "The challenges facing Japan's economy," speech given at the Naigai Josei Chousa Kai (The Research Institute of Japan), Tokyo, July 24.

Hoshi, Takeo, and Anil K Kashyap, 2001, *Corporate Financing and Governance in Japan: The Road to the Future*, Cambridge, MA: MIT Press.

_____**,** 1999, "The Japanese banking crisis: Where did it come from and how will it end?," National Bureau of Economic Research, working paper, No. W7250, July.

ING Barings, 2002, "Banks: What happens to capital if deferred tax assets are excluded?," Tokyo, report, April 26.

International Monetary Fund, 2002, "Japan: Staff report for 2002 Article IV consultation," Washington, DC.

Kikkawa, Masahiro, Takeshi Sakai, and Hiroyuki Miyagawa, 2000, "Soundness of the fiscal investment and loan program," in *Structural Problems of Japanese Financial System*, Mitsuhiro Fukao (ed.), Tokyo: Japan Center for Economic Research, pp. 41–59.

Merrill Lynch, 2002, "Life insurance industry," New York, report, June 5.

Peek, Joe, and Eric S. Rosengren, 2002, "Corporate affiliations and the (mis)allocation of credit," University of Kentucky, working paper.

Svensson, Lars, 2001, "The zero bound in an open-economy: A foolproof way of escaping from a liquidity trap," *Monetary and Economic Studies*, Vol. 19, February, pp. 277–312.

Takagi, Shokichi, 2002, press conference, Financial Services Agency, Tokyo, July 15, available at: www.fsa.go.jp/gaiyoue/tyoukan/e20020715.html.

Index for 2002

To order copies of any of these issues, or to receive a list of other publications, please telephone (312)322-5111 or write to: Federal Reserve Bank of Chicago, Public Information Center, P.O. Box 834, Chicago, IL 60690-0834. The articles are also available to download in PDF format from the Bank's website at www.chicagofed.org/publications/economicperspectives/index.cfm.