Federal Reserve Bank of Chicago

Announcement

2000 Conference on Bank Structure and Competition

First Quarter 2000

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.

ISSN 0164-0682

Contents

First Quarter 2000, Volume XXIV, Issue 1

The price of bank mergers in the 1990s

Elijah Brewer III, William E. Jackson III, Julapa A. Jagtiani, and Thong Nguyen

Introduction and summary

The last decade has witnessed an unprecedented pace of bank mergers and acquisitions. Between 1990 and 1998, the number averaged about 510 per year compared with 345 per year over the 1980–89 period. As a result of this activity, the number of banks operating in the U.S. has declined about 30 percent since 1990. In this article, we examine the primary motivations for this massive wave of bank mergers during the 1990s by analyzing the market prices of these mergers. A better understanding of the factors that determine market prices for bank mergers will shed some light on the implications of continuing mergers and acquisitions in the banking industry. We recognize that rapidly changing supply and demand conditions are fundamental to understanding what drives bank merger markets. For example, bank mergers may be driven by a desire to reduce overall risk by diversifying into new geographic or product markets. Additionally, bank mergers may be motivated by a strategic decision to exploit economies of scale, or to cut overhead and eliminate duplication by closing branches, or to achieve synergies through economies of scope. Of course, bank mergers may also be an attempt by banks to simply increase their market power or to quickly grow into superregional or money center banks.

To some extent, each of these motivations, and resultant strategies, became more feasible in the 1990s with the relaxing of state and federal restrictions on banks' activities. For example, the Riegle–Neal Interstate Banking and Branching Efficiency Act of 1994 allowed banks to branch interstate by consolidating existing out-of-state bank subsidiaries or by acquiring banks or individual branches through mergers and acquisitions. Prior to the Riegle–Neal Act, federal and state laws prevented banks from expanding across state lines (with some exceptions).[1] The Riegle–Neal Act allowed bank holding companies to acquire banks in any state, effective September 29, 1995, and allowed mergers between banks located in different states beginning June 1, 1997.[2]

On November 12, 1999, President Clinton signed the Financial Services Modernization Act (Gramm–Leach–Bliley Act of 1999), allowing banks to merge with securities firms and insurance companies within financial holding companies. This will further expand the merger opportunities for banking organizations and may lead to a new wave of consolidation in banking and other sectors of the financial services industry.

Another potential regulatory effect on bank merger trends is the Federal Deposit Insurance Corporation Improvement Act (FDICIA) of 1991. FDICIA introduced mandatory procedures called prompt corrective actions (PCA), which require regulators to promptly close depository institutions when their capital falls below predetermined quantitative standards, thus eliminating the possibility of regulators providing special consideration to large banks because of the possible systemic impact of large bank failure. Therefore, the notion of "too-big-to-fail" should be less relevant since FDICIA. However, an increase of megamergers has been noticeable in the mid- to late-1990s.

Elijah Brewer III is a senior economist and economic adviser in the Research Department at the Federal Reserve Bank of Chicago. William E. Jackson III is an associate professor of finance and economics at the Kenan–Flagler Business School of the University of North Carolina, Chapel Hill. Julapa A. Jagtiani is a senior financial economist in the Supervision and Regulation Department at the Federal Reserve Bank of Chicago. Thong Nguyen is a Ph.D. student in the Department of Economics at Yale University. *The research assistance of Gene Amromin, Sumeet Sawhrey, Susan Yuska, Oscar Cerda, Betsy Dale, and Erin Davis is greatly appreciated.*

Some research suggests that too-big-to-fail may have been one of the reasons for the rise in megamergers in the 1990s (see Kane, 1998). This is an important policy issue because previous research raises the possibility that banking organizations seek to become larger to increase the probability that the FDIC will cover 100 percent of their deposits. While most of the mergers between large publicly traded banks in the early and mid-1980s were not due to attempts to exploit deposit insurance, the too-big-to-fail consideration may have been important in megamergers of the 1990s (Benston, Hunter, and Wall, 1995; Hunter and Wall, 1989; and Boyd and Graham, 1991).[3] In this article, we examine the distinguishing characteristics of these megamergers.

Obviously, merger prices play an important role in the rise in bank mergers in the last decade. We analyze two types of prices commonly employed by both regulators and analysts: the merger bid premiums offered for a target bank, defined as the ratio of the market price offered for the target to the book value of equity of the target bank, and the excess stock returns earned by shareholders of the target bank around the merger announcement date.

We examine whether prices offered to target banks have been increasing over time. Increased prices would tend to make bank owners more willing to sell. There are theoretical reasons why prices must either increase or decrease as restrictions on expansion are reduced (Adkisson and Fraser, 1990). First, prior to Riegle–Neal, the number of potential bidders for a given target bank was limited by laws governing intrastate and interstate acquisitions. The removal of these restrictions should increase the demand for target banks as the number of potential bidders increases, resulting in higher acquisition prices. Thus, higher prices should be observed in the post-Riegle–Neal environment.

Alternatively, acquisition prices could be lower when restrictions are removed. Restrictions on geographical expansion form a barrier to entry that provides a bank with a protected niche and permits it to earn excess profits. These excess profits become part of the price in merger negotiations. Decreasing the barriers to entry reduces the excess profits and thereby lowers merger prices. By ensuring that they earn only normal profits, lowering the barriers to entry may increase substitutability among target banks, enlarging—from the acquirer's perspective—the effective supply of alternatives. Under the barriers to entry hypothesis, lower prices should be observed in the post-Riegle–Neal environment.

During the late 1970s and 1980s, individual states took steps, as permitted by the Douglas Amendment to the Bank Holding Company Act of 1956, to allow acquisition of banks in their states by bank holding companies headquartered in other states. The Bank Holding Company Act permitted multibank holding companies to acquire bank subsidiaries only to the extent allowed by the laws of the state in which the proposed target bank resided. Many states allowed acquisition by holding companies headquartered in only a limited number of states. Other states allowed entry from all states.

States in several regions developed formal compacts or treaties to allow entry from states in the region. The states in the Southeast formed the most cohesive unit, generally allowing entry from other states in the region and excluding entry from states outside the region (Savage, 1993). We use a Southeast indicator to test whether target banks in the Southeast received higher bid premiums than banks in other parts of the country.

In addition to examining how bank merger prices have changed over the 1990s and whether target banks in the Southeast receive higher bid premiums than other banks, we determine how prices are correlated with the financial characteristics of target banks and their market structure. As with any investment, the target bank's value to the acquiring bank should reflect its present discounted value of future net cash flows. At a minimum, the bid price should reflect the stand-alone value of the net assets of the target bank and the net cash flows from higher-valued deposit insurance as a result of the proposed merger.

Market structure, consisting of the number, size distribution, and market share of banks, influences the degree of competition and, thus, determines a bank's profitability. An often used measure of the degree of competition in banking markets is the Herfindahl-Hirschman Index (HHI), which is calculated by adding together the squares of the deposit shares of participants in a banking market and multiplying by 10,000. This index equals 10,000 for a monopoly market, and takes on lower values as more banks enter the market. For example, if there are five firms in a market and their deposit shares are 20 percent each, the HHI would be computed as follows: $[(.2)^2 + (.2)^2 + (.2)^2 + (.2)^2 + (.2)^2] \times 10{,}000 = 2{,}000$. Antitrust regulators use this measure to screen bank merger applications for potential anticompetitive effects.

In theory, target banks in markets with relatively high HHIs and, thus, operating in less competitive markets tend to receive high bid offers. A factor counterbalancing this tendency is the bank merger review process enforcing the U.S. Department of Justice's (DOJ) merger guidelines. This review

process could result in divestitures of banking offices as a condition of approval. Thus, the price an acquirer is willing to pay for a target bank should reflect the probable reaction of the antitrust authorities. If the merger review process works well, participating in a merger does not give banks greater market power. One way to analyze the effectiveness of the process is to determine the effects of market concentration on bid premiums. Bid premiums should be higher for targets in more concentrated markets.

We find a variety of interesting and important results. We find that higher performing targets, as measured by both return on equity and return on assets, receive higher bids. We also find that the lower the capital-to-deposit ratio, the larger the bid the acquiring bank is willing to offer. This may be because the target bank is funding its assets with relatively cheap funds. Additionally, we find that larger targets' loan-to-assets ratios are correlated with larger bid premiums, although this effect is not statistically significant. Bank size is positively related to bid premiums. Market concentration is not significantly correlated with bid premiums, reflecting the difficulties of applying our measure of concentration to banking organizations whose geographic scope and product mix may be broader than the local market area.

Prior to Riegle–Neal, prices paid for target banks in the Southeast regional compact tended to be higher than in other areas, perhaps reflecting the barrier to entry that provides the target in that region with a protected niche and permits it to earn excess profits. After Riegle–Neal, the Southeast effect was not significant. Overall, however, the price for target banks tends to be larger during the post-Riegle–Neal period, possibly because of the increase in the actual or potential number of bidders. Because Riegle–Neal provides increased interstate branching and banking opportunities, the demand for targets should increase as the universe of bidders increases, resulting in higher acquisition prices. Thus, we observe higher prices in the post-Riegle–Neal environment.

To get a better sense of how bank mergers are priced, we use daily stock return data to examine the stock market reaction to news of an announced merger. Results of this test are especially useful to interpret the wave of large bank mergers. If banks are using their increased freedom to merge in a way intended to increase the value of their deposit insurance, then megamergers should generate high bid premiums and, thus, greater than expected stock market returns than other types of mergers. However, bank managers may also pursue mergers to enhance their salary, perquisites, and personal prestige. As a result, high bid premiums for large banking organizations may be related to several different motivations, many of which will tend to lead to high merger prices.

A countervailing factor in large bank mergers, however, is the difficulty of merging two large banking organization or two organizations of equal size. According to organization theorists, melding cultures in a merger is more difficult and costly when the target is closer in size to the acquirer. If the short-run costs are a positive function of size and these costs outweigh the value of increased access to deposit insurance, then we would expect to see an inverse relationship between size and merger prices (Benston, Hunter, and Wall, 1995). Our stock return results suggest that the stock market views large bank mergers more favorably than small bank mergers. The unexpected stock returns for large bank targets are more than double those of small targets.

Our results suggest that changes in state and federal banking regulations have a significant impact on bank merger activity in general, and bank merger prices in particular. Furthermore, by restricting the types of merger transactions that can take place, state and federal interstate and intrastate banking laws may have had unintended consequences. Because restrictions on geographical expansion form a barrier to entry that provides a bank with a protected market and permits it to earn excess profits, we observe higher bid premiums in Southeast compact states relative to other parts of the country. Once these restrictions were removed with the passage of Riegle–Neal in 1994, bid premiums were no longer higher in the Southeast states than in other states. However, they rose overall relative to the pre-Riegle–Neal period. Thus, our results show how federal and state regulatory policies that restrict interstate branching and banking may produce very different (and distorted) merger prices relative to policies that are less restrictive and market driven.

Finally, our results provide empirical evidence that when target banks are large, but not megamergers of equals, there is a greater stock market reaction to the merger announcement than for other target banks. This is consistent with the notion that large banks are using their increased freedom to merge in a way intended to increase the value of their deposit insurance. A partial explanation for the recent wave of mergers, especially megamergers, may be the desire of merging institutions to obtain a size level sufficient to place them in the too-big-to-fail category. This is an important issue for policymakers, who are concerned about controlling bank risk-taking propensities and minimizing the loss exposure of the federal deposit insurance funds. Thus, size is important in merger decisions

because larger institutions may have increased access to deposit insurance. This greater access tends to be reflected in the stock market reactions to merger announcements.

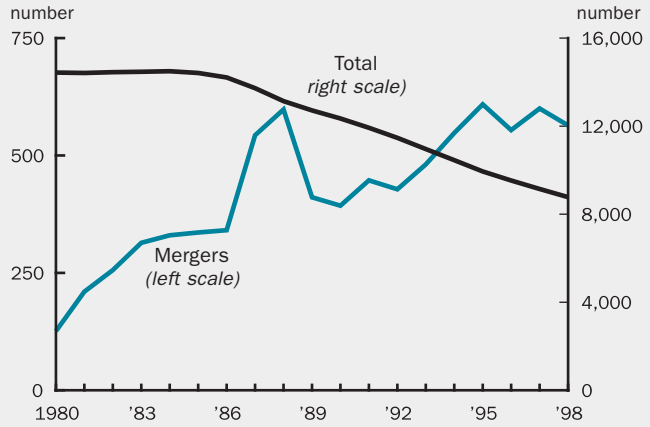
Regulatory background

Opportunity for nationwide branch banking

Figure 1 shows that the number of banking organizations in the U.S. has decreased by about 40 percent since 1980. This decline is related to the surge in the number of bank mergers—225 per year during the early 1980s compared with 580 per year during the late 1990s. The share of domestic deposits held by the nation's ten largest commercial banks nearly doubled from about 19 percent in 1980 to 37 percent in 1998 (DeYoung, 1999). Table 1 provides further evidence of the consolidation trend in banking, which has occurred contemporaneously with the reduction in restrictions on interstate banking and branching.

FIGURE 1

Growth in commercial banks and bank mergers



Source: Authors' calculations based on data from the Federal Deposit Insurance Corporation.

As mentioned earlier, under Riegle–Neal, banks have been permitted to engage in nationwide branch banking since June 1, 1997. This liberalization made

TABLE 1

Size distribution of commercial banks

Asset size *($ millions)*	Number of banks	Percent of banks	Cumulative percent	Total assets	
				Percent	Cumulative
A. December 31, 1980					
Less than 25	7,233	49	49	5	4
25–50	3,566	24	73	6	11
50–100	2,048	14	87	7	18
100–500	1,496	10	97	15	33
500–1,000	195	1	98	7	40
1,000–5,000	192	1	99	19	59
5,000–10,000	21	–	99	7	66
10,000 or more	18	–	100	34	100
Total	**14,769**	**100**			
B. December 31, 1998[a]					
Less than 25	3,156	36	36	3	3
25–50	2,261	26	62	5	8
50–100	1,700	19	81	7	15
100–500	1,279	15	96	14	29
500–1,000	149	2	98	6	35
1,000–5,000	114	1	99	15	50
5,000–10,000	20	–	99	8	58
10,000 or more	25	–	100	42	100
Total	**8,704**	**100**			

[a]Adjusted for inflation using the Consumer Price Index (all items, 1982–84 = 100).
Note: Numbers are rounded to the nearest whole number.
Source: Authors' calculations based on data from *Quarterly Reports of Condition and Income.*

possible the merger of large banking organizations to create true nationwide banking in the U.S.

However, Riegle–Neal was only the final piece of legislation in a long line of banking deregulation at the state level. Historically, restrictions on banks' ability to expand geographically have been among the primary determinants of the structure of commercial banking in the U.S (Frieder, 1988; and Cornett and De, 1991a). Concerns about undue concentration of banking resources and that banks might exercise their market power by setting high prices and restricting service led to the imposition of restrictions at both the state and national levels. The McFadden Act of 1927 restricted nationally chartered banks' branching ability to the same extent allowed to state-chartered banks. The Bank Holding Company Act of 1956 prevented multibank holding companies (MBHCs) from acquiring existing banks or chartering new banks in states other than their home state. The Douglas Amendment of the 1956 act allowed MBHCs to acquire banks only to the extent permitted by the laws of the state of the target bank. Even the Riegle–Neal Act limits the market share that a banking organization can hold nationwide or in any given state. The act established a 10 percent nationwide deposit concentration limit on organizations making interstate acquisitions and a uniform 30 percent statewide limit (unless a state chooses a different limit).

The first state statutes permitting entry to out-of-state MBHCs in accordance with the Douglas Amendment were enacted in 1975 and 1982 by Maine and Alaska, respectively. By the late 1980s, 41 states and the District of Columbia had passed similar laws (Amel, 1986; Frieder, 1988; and Cornett and De, 1991a). Moreover, several states formed reciprocal regional banking pacts to allow banks in pact states to acquire targets in other pact states. For example, prior to the Riegle–Neal Act, Wisconsin's regional reciprocal law allowed entry by acquisition for banking organizations from Illinois, Iowa, Indiana, Kentucky, Michigan, Minnesota, Missouri, and Ohio as long as those states allowed acquisitions by Wisconsin banks in their markets (Saunders, 1997).

Antitrust statutes and authorities

The federal statutes that govern bank mergers are the Bank Holding Company Act, the Bank Merger Act of 1960, and section 7 of the Clayton Act of 1914. The DOJ has general enforcement authority over all merger and acquisition activities and has established basic guidelines to cover the evaluation of competitive issues (Jackson, 1992; and Kwast, Starr-McCluer, and Wolken, 1997). Box 1 provides a discussion of the antitrust legal standards.

Under the Bank Merger Act the three federal regulatory agencies—the Board of Governors of the Federal Reserve System, the office of the Comptroller of the Currency (OCC), and the Federal Deposit Insurance Corporation (FDIC)—are required to take into account the competitive effects of a proposed merger. The agency to which a merger application should be submitted depends on the "resultant banking organization." If the resultant banking organization is a nonmember federally insured bank, the application needs to be made to the FDIC. If the resultant banking organization is a state member bank, the application needs to be made to the appropriate Federal Reserve Bank, and if it is a national bank, the merger application should be made to the OCC. In addition to the powers provided in the Bank Merger Act, the Federal Reserve System derives its legal authority over bank mergers from the Bank Holding Company Act, which prohibits a bank holding company from acquiring a bank unless the bank holding company has received prior approval from the Federal Reserve System. The DOJ may prevent consummation of the merger within 30 days of the approval from the relevant federal agency. After this 30-day period, the merger is immune from the DOJ and other private party litigation.

Antitrust guidelines

The regulators have adopted the DOJ's numerical criteria for assessing the impact of a merger or acquisition on competition. These criteria, first issued in 1968, were updated in June 1982 based on the HHI. In evaluating a merger application, antitrust authorities consider both the level of post-merger HHI and the change in the HHI resulting from the proposed transaction (see table 2).[4] If the post-merger market HHI is lower than 1,800 points, or the increase in the index from the pre-merger situation is less than 200 points (or 50 points in industries other than banking),[5] the merger is presumed to have no anticompetitive effects and is generally approved by regulators. The Federal Reserve uses the acquiring firm's market share as an additional merger screen. A merger is likely to raise concerns if the acquirer's pro forma market share exceeds 35 percent.

When a merger application violates the guidelines, regulators consider *mitigating factors* that would offset the anticompetitive effects of the proposed transaction. These factors include competitive viability of the target, presence of active competition from thrifts and other financial institutions in the market, competition from out-of-market financial institutions, and market attractiveness. These factors are weighted against the increase in concentration. If the increase in concentration is too large to be justified by mitigating

BOX 1

Antitrust legal standards

The major antitrust concern of each of the federal bank regulatory authorities is the competitive effect of mergers and acquisitions. An examination of this issue requires a clear and concise definition of the product and geographical markets in which competition takes place and a standard to measure the competitive effects of each merger (see Jackson, 1992). This framework was not specifically stated in the federal statutes that govern mergers. Instead, it has evolved from three Supreme Court decisions in the 1960s and 1970s: United States vs. Philadelphia National Bank (1963), United States vs. Phillipsburg National Bank and Trust (1970), and United States vs. Connecticut National Bank (1974).[1]

In the Philadelphia National Bank case, the Supreme Court:

- provided the principles by which product and geographic markets should be defined to assess the probable competitive effects of a bank merger or acquisition, and
- noted that commercial banks are unique among financial institutions (including thrifts) in that they alone are permitted by law to accept demand deposits and operate with the benefit of federal deposit insurance.

The court ruled that the relevant product market was the "cluster of commercial banking services differentiating commercial banking as a unique line of business." Thus, only competing commercial banks were included in the framework for the purposes of analyzing a proposed bank merger under the Clayton Act. The exact definition of cluster was not specifically stated in the court's decision. However, antitrust regulators have used total deposits as a proxy for the ability of commercial banking organizations to provide the cluster of banking services to both businesses and households in a given local banking market (Rhoades, 1987).

In the Philadelphia National Bank case, the court noted that the appropriate geographical market for competitive analysis does not depend on where the parties to a merger do business or compete. Instead, it depends on the geographical structure of the supplier–customer relationships and where a purchaser of products and services can practicably turn for alternative banking services. The court found that convenience of location is essential to effective competition, suggesting that geographical markets for commercial banking are generally considered to be local, for example, within counties or metropolitan statistical areas (Holder, 1993).

In the Phillipsburg National Bank and Trust case (1970), the Supreme Court held that:

- for the purposes of analyzing a proposed merger under the Clayton Act, regulators should consider both the level of concentration and the change in concentration of firms in the appropriate geographical market, and
- a merger application may be accepted if it can be shown that the transaction provides substantial public benefits even though it may violate antitrust guidelines.

The structure–conduct–performance paradigm suggests that market concentration beyond a certain point will likely lead to collusive or monopolistic behavior by banks, a direct violation of the Clayton Act. Banking regulators have thus focused on the anticompetitive issues of bank mergers and acquisitions in terms of the resultant effects on market concentration (Rhoades, 1987).

In the Connecticut National Bank case, the Supreme Court:

- revisited the geographical market definition and ruled that the relevant banking market is not a state but rather a segmented group of bank office areas where a bank would seek business and, as a practical matter, most of its customers would do their banking, and
- concluded that thrift institutions should not be factored into antitrust analysis, but acknowledged that they may be included "when and if saving banks become significant participants in the marketing of bank services to commercial enterprises."

These three court decisions provide the fundamental concepts for analyzing competitive effects of bank mergers and acquisitions. In particular, they hold that 1) the "cluster" of bank products is the relevant product line for competitive analysis; 2) this cluster is typically viewed as being consumed in geographically local banking markets; and 3) market structure is a key determinant of the degree of competition (Kwast, Starr-McCluer, and Wolken, 1997; and Jackson, 1992).

---

[1]United States vs. Connecticut National Bank, 418 U.S. 656 (1974); United States vs. Phillipsburg National Bank and Trust Company, 399 U.S. 350 (1970), and United States vs. Philadelphia National Bank, 374 U.S. 321 (1963).

TABLE 2

1982 Department of Justice horizontal merger guidelines

Post-merger market concentration	Level of HHI	Post-merger change in HHI and likelihood of challenge
Highly concentrated	Greater than 1,800	Greater than 100—Challenge likely
		50 to 100—Depends on other factors[a]
		Less than 50—Challenge unlikely
Moderately concentrated	1,000 to 1,800	Greater than 100—Challenge likely; other factors considered[a]
		Less than or equal to 100—Challenge unlikely
Unconcentrated	Less than 1,000	Any increase—Challenge unlikely

[a]Lead firm provision—A merger is likely to be challenged if the merger is between the lead firm and a firm with a market share of 1 percent or more provided that the lead firm has a market share of 35 percent or more and is approximately twice the size of the second largest firm in the market. These so-called other factors are often related to ease and profitability of collusion. In banking, they are often referred to as mitigating factors and include competitive viability of the target, presence of active competition from thrifts and other financial institutions in the market, competition from out-of-market financial institutions, and market attractiveness.

Note: When released on June 14, 1982, the guidelines in this table applied to all U.S. industries. In 1985, the U.S. Department of Justice modified the 1,800/50 rule for bank mergers to 1,800/200 to recognize the impact of competition from limited purpose lenders and other nondepository financial institutions.

Source: U.S. Department of Justice, 1982, press release, June 14.

factors, divestiture of some branches and offices may bring the concentration indicator close to or below the DOJ guidelines. Consequently, very few bank mergers are denied due to antitrust concerns. However, the official statistics do not include applications that are voluntarily withdrawn when consultation with regulatory agencies indicates they would be found to be anticompetitive.

Effects of geographical deregulation on bank acquisition prices

The literature suggests two competing hypotheses to explain how geographic deregulation might affect the prices paid for bank acquisitions (Adkisson and Fraser, 1990). Under the *excess demand* theory, prices of acquisitions should increase as restrictions on expansion are reduced. Prior to Riegle–Neal, the number of potential bidders for a target bank was limited by state law governing intrastate and interstate acquisitions.

As noted above, during the late 1970 and 1980s, some states formed regional banking pacts to allow banks to merge with or acquire targets in pact states (see details in table 3). Other states allowed nationwide entry with reciprocal arrangements. As these restrictions are removed, the demand for targets should increase as the universe of bidders increases, resulting in higher acquisition prices. Thus, all else being equal, higher prices should be observed as states liberalize their interstate banking laws and in the post-Riegle–Neal environment.

Conversely, the barrier to entry theory predicts that merger prices will be lower when bank acquisition laws are more liberal. Geographical expansion restrictions form a barrier to entry that provides the target with a protected niche and permits it to earn excess profits. Decreasing the barriers to entry reduces the excess profits and thereby lowers merger prices. By ensuring that they earn only normal profits, lowering the barriers to entry may increase substitutability among target banks, enlarging (from the acquirer's perspective) the effective supply of alternatives. Thus, lower prices should be observed in regional compact states and in the post-Riegle–Neal environment.

There are at least two approaches to calculating the price offered by acquirers for targets. One approach measures the size of the merger premium (or bid premium). More attractive targets receive higher bid premiums. The second approach uses stock return data and is usually called the event study approach. Under this approach, excess returns (or abnormal returns) are computed around the merger announcement date (see box 2).

Literature review and our contribution

Previous studies on mergers and acquisitions of nonfinancial firms have produced mixed results about the determinants of merger premiums. It is even more complicated to identify the determinants of these premiums in the banking industry due to the high level of governmental regulations and monitoring. In addition

TABLE 3

Interstate banking laws prior to Riegle–Neal Act

State	Area covered and reciprocity
Alabama	Reciprocal, 13 states (AR, FL, GA, KY, LA, MD, MS, NC, SC, TN, TX, VA, WV)
Alaska	National, no reciprocity
Arizona	National, no reciprocity
Arkansas	Reciprocal, 16 states (AL, FL, GA, KY, LA, MD, MO, MS, NC, NE, OK, SC, TN, TX, VA, WV) and DC
California	National, reciprocal
Colorado	National, no reciprocity
Connecticut	National, reciprocal
Delaware	National, reciprocal
District of Columbia	Reciprocal, 11 states (AL, FL, GA, LA, MD, MS, NC, SC, TN, VA, WV)
Florida	Reciprocal, 11 states (AL, AR, GA, LA, MD, MS, NC, SC, TN, VA, WV) and DC
Georgia	Reciprocal, 11 states (AL, FL, KY, LA, MD, MS, NC, SC, TN, VA, WV) and DC
Idaho	National, no reciprocity
Illinois	National, reciprocal
Indiana	National, reciprocal
Iowa	Reciprocal, 6 states (IL, MN, MO, NE, SD, WI)
Kansas	Reciprocal, 6 states (AR, CO, IA, MO, NE, OK)
Kentucky	National, reciprocal
Louisiana	National, reciprocal
Maine	National, no reciprocity
Maryland	Reciprocal, 14 states (AL, AR, DE, FL, GA, KY, LA, MS, NC, PA, SC, TN, VA, WV) and DC
Massachusetts	National, reciprocal
Michigan	National, reciprocal
Minnesota	Reciprocal, 16 states (CO, IA, ID, IL, IN, KS, MI, MO, MT, ND, NE, OH, SD, WA, WI, WY)
Mississippi	Reciprocal, 13 states (AL, AR, FL, GA, KY, LA, MO, NC, SC, TN, TX, VA, WV)
Missouri	Reciprocal, 8 states (AR, IA, IL, KS, KY, NE, OK, TN)
Montana	Reciprocal, 7 states (CO, ID, MN, ND, SD, WI, WY)
Nebraska	National, reciprocal
Nevada	National, no reciprocity
New Hampshire	National, no reciprocity
New Jersey	National, reciprocal
New Mexico	National, no reciprocity
New York	National, reciprocal
North Carolina	Reciprocal, 13 states (AL, AR, FL, GA, KY, LA, MD, MS, SC, TN, TX, VA, WV) and DC
North Dakota	National, reciprocal
Ohio	National, reciprocal
Oklahoma	National, no reciprocity for initial entry; after initial entry, bank holding company must be from state offering reciprocity or wait 4 years to expand
Oregon	National, no reciprocity
Pennsylvania	National, reciprocal
South Carolina	Reciprocal, 12 states (AL, AR, FL, GA, KY, LA, MD, MS, NC, TN, VA, WV) and DC
Rhode Island	National, reciprocal
South Dakota	National, reciprocal
Tennessee	National, reciprocal
Texas	National, no reciprocity
Utah	National, no reciprocity
Vermont	National, reciprocal
Virginia	Reciprocal, 12 states (AL, AR, FL, GA, KY, LA, MD, MS, NC, SC, TN, WV) and DC
Washington	National, reciprocal
West Virginia	National, reciprocal
Wisconsin	Reciprocal, 8 states (IA, IL, IN, KY, MI, MN, MO, OH)
Wyoming	National, no reciprocity

Note: Hawaii did not enact interstate bank holding company legislation.
Source: Savage (1993).

BOX 2

Estimating the stock price impact of mergers

We estimate the stock price impact of each of these merger announcements by employing a multivariate regression model (MVRM), similar to those used in Schipper and Thompson (1983), Binder (1988), and Cornett and Tehranian (1989). In the MVRM, abnormal returns are obtained by adding a (0,1) binary variable to the right-hand side of the traditional market model to capture the impact of the announcement or "event" date. The model takes the following form:

$$R_{j,t} = \alpha_j + \beta_{j,1}R_{M,t-2} + \beta_{j,2}R_{M,t-1} + \beta_{j,3}R_{M,t} + \beta_{j,4}R_{M,t+1} + \beta_{j,5}R_{M,t-2} + \sum_{s=0}^{+1}\gamma_{j,s}D_s + \varepsilon_{j,t},$$

where $R_{j,t}$ is the return on firm j on day t; $R_{M,t}$ is the return on the market portfolio; α_j is an intercept coefficient for bank j; $\beta_{j,1...j,5}$ are risk coefficients for the jth bank; γ_j is the effect of the merger announcement event on the jth firm; D_s is an event binary variable which equals 1 on day s (s = 0 to +1) in the event window, and 0 otherwise; and $\varepsilon_{j,t}$ is a random error term which is assumed to be identically distributed normally, independent of the return on the market and the binary variables. We specify the market return at several leads and lags as an explanatory variable to correct for the possibility of nonsynchronous trading, especially of some of the smaller banks (Scholes and Williams, 1977).

With this specification, the estimated parameters γ_j measure the daily abnormal returns associated with a merger announcement. We are testing for daily intercept shifts in the interval day 0 to day +1. Since this interval is "dummied out," the observations in the day 0 to day +1 interval do not influence the estimate of the intercept. Only those observations without dummies determine the value of the intercept.

We estimate the target bank's cumulative abnormal stock market returns over the two trading day period that includes the announcement date and the day after. The two-day cumulative abnormal returns (CAR) around the merger announcement date (t = 0 and +1) are then calculated by adding $\gamma_{j,0}$ and $\gamma_{j,1}$. The standardized cumulative abnormal returns are computed using a procedure reported in Bradley, Desai, and Kim (1988), and Stultz, Walking, and Song (1990). First, the standardized abnormal return to the jth security on day t ($SAR_{j,t}$) is computed using the following equation:

$$SAR_{j,t} = AR_{j,t} \Big/ \left[\sigma_j \left(1 + \frac{1}{T_j} + \frac{(R_{m,t} - \overline{R_m})^2}{\sum_{\tau=1}^{T_j}(R_{m,\tau} - \overline{R_m})^2}\right)^{1/2}\right],$$

where $AR_{j,t}$ is the abnormal return to the jth security on day t, S_j is the standard deviation of the residuals in the market model estimation period, T_j is the number of days in the estimation period, $R_{m,t}$ is the return on the market portfolio on day t, and $\overline{R_m}$ is the mean return on the market portfolio over the estimation period.

The $SAR_{j,t}$ is then used to obtain the standardized cumulative abnormal returns over the two event days:

$$SAR_{j,t} = \left[\sum_{t=1}^{2} SAR_{j,t}\right] / \sqrt{2}.$$

to characteristics of the deal, the target, and the acquiring banks, regulatory environments in both acquiring and target bank states tend to affect the bid premiums (see Palia, 1993). The analysis of bank merger premiums is further complicated by regulatory uncertainty (see Desai and Stover, 1985). All bank mergers require time-consuming regulatory approval, making hostile takeovers extremely difficult to execute.

Previous bank studies (Beatty, Santomero, and Smirlock, 1987; Cheng, Gup, and Wall, 1989; Fraser and Kolari, 1988; Rogowski and Simonson, 1989; and Rose, 1991), find that asset size, profitability, management, leverage, means of payment, and whether the mergers are interstate or intrastate are significant in determining the bid premiums or explaining the stock market's reaction to bank merger announcements.

The literature suggests that size is important in determining the bid premiums offered to the target, but less important in determining the abnormal returns. For example, Desai and Stover (1985) find that the relative size of target and acquiring banks has no significant impact on the abnormal returns around the announcement date. However, Shawky, Kilb, and Staas (1996) find that smaller targets tend to be offered a larger bid premium, and Palia (1993) finds that the

relative size of targets and acquiring banks is important in explaining the variation in the bid premiums.

With regard to profitability and capital, Shawky, Kilb, and Staas (1996) find that higher bid premiums tend to be offered to target banks with larger returns on equity and those with higher leverage. The latter result suggests that higher leverage may be associated with more efficient use of capital.

Whether the merger deals are stock exchange offers or cash offers may also affect the abnormal returns and the bid premiums—because of the differential tax implications associated with these offers. The market may view a cash offer positively for the acquirer, because it allows the acquiring bank to increase the depreciation tax shield as the depreciation basis of the acquired assets rises to the market value. However, the market may view it negatively for the target, because it imposes a greater immediate tax burden on target shareholders. However, the market may view a cash offer negatively for the acquirer and positively for the target if the acquirer's share price is relatively overvalued. Shawky, Kilb, and Staas (1996) find that the bid premiums offered to target banks are larger for stock deals (rather than cash payment), supporting the acquirer's overvalued stock hypothesis. However, Cornett and De (1991b) find that mergers financed with only stock or only cash produce higher abnormal returns to target shareholders than those financed with combinations of stock and cash. Interestingly, previous studies on nonbank mergers find medium of payment to be unimportant (Eckbo and Langohr, 1989; and Travlos, 1987).

Rhoades (1987) suggests that geographical expansion may be a primary motivation for bank mergers. Palia (1993) and Shawky, Kilb, and Staas (1996) find that the bid premiums offered to target banks are larger for out-of-state mergers than intrastate mergers. Using interstate bank mergers, Cornett and De (1991a) find significant positive announcement period abnormal returns for both target and acquiring banks. Again, this contrasts with findings for nonbank mergers, suggesting that bank mergers are different and, thus, the results for nonbank mergers cannot be generalized to the banking sector.

Examining the postmerger performance of large bank mergers between 1982 and 1987, Cornett and Tehranian (1992) find that merged banks tend to perform better than the banking industry overall. This superior performance resulted from improvements in the merged banks' ability to attract loans and deposits, employee productivity, and asset growth. (For a recent literature review on bank merger performance, see Berger, DeYoung, Genay, and Udell, 2000.)

Cornett and Tehranian (1992) examine operating cash flows as well as several accounting variables of the merged banks for one to three years after the mergers. Recognizing that accounting data are not perfect measures of economic performance, they utilize both accounting and market data to determine whether stock price gains associated with mergers announcement (short run) are the result of real economic gains (long run). Interestingly, they find a significant correlation between announcement-period abnormal stock returns and the various long-term performance measures, and conclude that market participants are able to identify in advance the improved performance associated with bank acquisitions.[6] We focus on short-term performance, using market data, rather than testing whether mergers will result in efficiency gains or improved long-run performance.

Overall, the empirical results presented in the previous studies have been mixed and largely depend on the sample period, sample observations, and methodology. We reexamine this issue using more recent and more complete data on bank mergers. Our results are more applicable to current policy issues than previous studies, given the rapidly evolving environment the banking industry faces.

The data

We obtained details of all bank mergers and acquisitions from 1990 to mid-1998 from the Security Data Corporation (SDC). To be included in our sample, both the target and bidding banks must be publicly traded.[7] We obtained financial data from the quarterly call reports and bank holding company Y9 reports, as of yearend prior to the merger announcement date. Stock market returns for target banks and the stock market index are from the Center for Research in Security Prices data tape. The merger announcement date, target name, acquirer name, value of the deal, bid premium, and other characteristics of the merger announcement are from the SDC database. We obtained the HHIs for various banking markets from the Board of Governors of the Federal Reserve System.

Banking agencies consider a local, economically integrated area to be a banking market. In practice, this usually means a city, a metropolitan statistical area (MSA), or a rural county. We matched the target bank's headquarters with an MSA or county. This does not allow us to consider a target bank present in several different market areas and points to the difficulty of using local market concentration measures for banking organizations that have broad geographical scope and product mix. The price that a bank offers (or accepts) reflects the activities of the entire organization.

The methodology

We use a *regression model* that relates a target bank bid premium to profitability, asset size, financial leverage, loan quality, Southeast indicator variable, Riegle–Neal indicator variable, and a concentration measure. A formal discussion of the model is presented in box 3.

To capture the profitability of a target banking organization, we include the return on equity and the return on assets in the year before the merger announcement date. We expect the sign on profitability to be positive, as higher profits are more attractive.

We include a variable that measures the size of each target. Bank size, as measured by the natural logarithm of total assets, may be either positively or negatively associated with the attractiveness of a given target. A positive coefficient for this variable would be consistent with the hypothesis that potential bidders look for significant targets that participate in significant markets. A negative coefficient, if found, may reflect the cost of melding the culture of a large target bank with that of the acquirer.

Banking organizations are required by regulation to meet minimum capital requirements. This regulation is aimed at reducing the risk-taking propensities of bank shareholders. That is, capital acts as a form of co-insurance with federal deposit insurance. We include the leverage ratio, defined as the capital-to-deposit ratio, in the year before the merger announcement date. We expect the sign on the leverage ratio to be negative. A high capital-to-deposit ratio may be an indication that the target banking organization is using capital inefficiently. This argument is consistent with Beatty, Santomero, and Smirlock (1987), Fraser and Kolari (1988), and Rogowski and Simonson (1989). A well-capitalized acquirer seeks target banks that offer an ample inexpensive source of funds.

Because loans are usually the most illiquid and subject to the greatest default risk of all bank assets, a bank's risk is greatly influenced by the quality of its loan portfolio. The ratio of loans to total assets measures the potential effects of loan losses on assets and equity and the illiquidity of assets. According to our hypothesis, the greater the proportion of loans to

BOX 3

The model

The following basic specification is used to examine the factors that are correlated with the bid premium offered for the target (*BVPREM*):

$$
\begin{aligned}
1)\quad BVPREM_{j,t} = {} & \alpha_0 + \alpha_1 PROFIT_{j,t} + \alpha_2 LEV_{j,t} \\
& + \alpha_3 SIZE_{j,t} + \alpha_4 LOAN_{j,t} \\
& + \alpha_5 CHARGE_{j,t} + \alpha_6 MEQUAL_{j,t} \\
& + \alpha_7 THRIFT_{j,t} + \alpha_8 SEAST_{j,t} \\
& + \alpha_9 RNEAL_{j,t} + \varepsilon_{j,t},
\end{aligned}
$$

where *PROFIT* is a measure of profitability of the target one-year before the merger announcement date; *LEV* is the capital-to-deposit ratio one year before the merger announcement date; *SIZE* is the natural logarithm of total assets of the target banking organization; *LOAN* is ratio of total loans to total assets; *CHARGE* is the ratio of net chargeoffs- to-loans; *MEQUAL* is an indicator variable that is equal to one if the target and the acquirer are of equal asset size and zero otherwise; *THRIFT* is an indicator variable that is equal to one if the target is a savings and loan association and zero otherwise; *SEAST* is an indicator variable that is equal to one if the target and acquirer are located in the Southeast regional compact (AL, AR, FL, GA, KY, LA, MD, MS, NC, SC, TN, VA, and DC) and zero otherwise; *RNEAL* is an indicator variable that is equal to one if the merger announcement date is after 1994 and zero otherwise; and $\varepsilon_{j,t}$ is a random error term. The *MEQUAL* variable is included in the equation to determine whether banks involved in mergers of equals are offered a different price than other banks. The *THRIFT* variable is included in the equation to control for the different charter between banking organizations and thrift institutions. The *SEAST* variable is included to capture whether the southeast regional compact led to differences in bid premium. This indicator variable absorbs the effects of all factors that are common to banking organizations in the Southeast. The *RNEAL* variable is included to capture the impact of the Riegle–Neal Act on bid premiums.

In some specifications, we include indicator variables for the year of the announcement date of the acquisitions that range between 1990 and mid-1998. These variables are introduced to account for the effect of omitted macroeconomic and other variables that may influence the overall level of acquisition activity over time and, thus, the merger premium paid for a given transaction.

total assets, the greater the potential for loan losses and the lower the liquidity of assets, *ceteris paribus*. Everything else held constant, this should lead to a lower bid premium. However, because loans offer the potential for geographical diversification, the loan portfolio could have a positive impact on the bid premium. Thus, we include the ratio of net chargeoffs to loans to capture asset quality more directly.

Another hypothesis we examine is whether state and federal laws on interstate and intrastate branching and banking influence the price offered for target banks. To capture whether the Southeast regional compact led to differences in bid premiums, we include a Southeast indicator variable, which absorbs the effects of all factors that are common to banks in the Southeast.

Our regression equation also includes an indicator variable that captures the impact of the Riegle–Neal Act on the bid premium. A positive coefficient on this variable is consistent with the notion that the universe of actual or potential bidder has increased, resulting in higher acquisition prices. A negative coefficient is consistent with the notion that liberalization of interstate banking laws reduces excess profits, leading to lower merger prices.

Finally, we include indicator variables for the year of the merger announcement date from 1990 to mid-1998 to account for the effects of omitted macroeconomic and other variables that may influence the level of acquisition activity over time and, thus, the merger premium paid for a given transaction.

The results

Table 4 provides a summary of selected financial characteristics of the target, bid premiums, and cumulative abnormal returns (CAR). The standardized cumulative abnormal returns (SCAR) are the cumulative abnormal returns adjusted for the error in forecasting the returns (see box 2). The results show that targets with lower capital-to-asset ratios and higher profitability (larger return on assets) tend to obtain larger stock price gains, as measured by CAR and SCAR, around the merger announcement date than those with high capital-to-asset ratios or lower profitability. In addition, the stock market abnormal returns around the announcement of the merger seem to be larger at

TABLE 4

Selected financial characteristics of target and price information

Financial characteristics	Quartile (1 = lowest)	Range % (of explanatory variable)	Mean values		
			CAR (%)	SCAR (%)	BVPREM
Book value of capital/total assets	1	2.2–6.8	14.0	5.1	2.1
	2	6.8–8.1	11.0	4.2	2.4
	3	8.1–9.4	12.5	4.8	2.1
	4	9.4–25.4	8.3	3.4	2.2
Total loans/total assets	1	22.3–55.5	12.1	4.8	1.9
	2	55.5–63.5	13.3	5.3	2.3
	3	63.5–68.9	9.2	3.9	2.3
	4	68.9–89.2	11.1	3.5	2.2
Total assets (in millions)	1	35.7–272.0	11.4	3.6	2.1
	2	272.0–902.3	13.2	4.4	2.0
	3	902.3–3,276.1	10.2	4.1	2.3
	4	3,276.1–260,159	11.0	5.5	2.2
Return on assets	1	–1.3–0.4	10.8	3.6	1.6
	2	0.4–0.9	12.9	4.8	2.1
	3	0.9–1.2	10.5	3.9	2.2
	4	1.2–2.2	11.5	5.2	2.8
Riegle–Neal (0 before act, 1 after)	0	96 (122)	13.0	4.1	1.7
	1	146 (205)	10.4	4.6	2.4

Notes: CAR is cumulative abnormal returns; SCAR is standardized cumulative abnormal returns; BVPREM is the bid premium offered for the target. CAR and SCAR means are computed using data for a subsample of 242 acquirer institutions. Number of observations in parentheses is out of the 327 observations constituting the entire sample.

TABLE 5

Pricing of bank mergers, regional, and financial characteristics

	Book value premium	Standardized cumulative excess returns
Southeast regional compact		
Yes (AL, AR, FL, GA, KY, LA, MD, MS, NC, SC, TN, VA and DC)	2.4469	5.6950
No	2.0600	3.9400
Difference	0.3869 (2.06)**	1.7550 (1.51)
Before Riegle–Neal Act	2.4687	7.2392
After Riegle–Neal Act	1.7907	2.5294
Difference	0.6780 (3.80)***	4.7098 (2.73)***
Asset characteristics		
Assets ≥$10 billion[a]	2.3612	8.1679
Assets <$10 billion	2.1456	4.0603
Difference	0.2156 (0.89)	4.1076 (3.37)***
Banks versus thrifts		
Banks	2.2500	4.4376
Thrifts	1.4721	3.9604
Difference	0.7779 (7.21)***	0.4772 (0.46)

[a]Excluding banking organizations classified as mergers of equals.
***Indicates significance at the 1 percent level.
**Indicates significance at the 5 percent level.
Note: The t-statistics are in parentheses.

larger target banks. The abnormal returns tend to decline as the target's proportion of loans to assets increases. Similarly, the bid premiums tend to increase with the target's return on assets, and have become larger in the post-Riegle–Neal period. Unlike the abnormal returns, the bid premiums offered for targets seem to be positively correlated with the loan to assets ratio.

Table 5 presents the bid premiums and the standardized cumulative abnormal returns for different target characteristics. The target's abnormal returns around the merger announcement date are significantly larger for targets in the Southeast regional compact. Interestingly, while both the bid premiums and the abnormal returns are generally larger in the post-Riegle–Neal periods (as presented in table 4), they are significantly smaller for target banks in the Southeast.

The results from both tables 4 and 5 suggest that the bid premiums are not statistically different among targets with different asset sizes. However, the stock market reaction (the standardized cumulative abnormal returns) varies depending on the size of the target banks. Overall, the larger the target bank, the larger the standardized cumulative abnormal returns around the merger date (see table 4).[8] Finally, table 5 shows that the standardized cumulative abnormal returns are not statistically different whether the target is a bank or thrift. However, the bid premiums offered for target banks are, on average, significantly larger then those offered for thrifts.

Table 6 separates the megamerger deals from the rest of the sample. The pattern of variation of the bid premiums according to target size is now more evident. Within the large target bank group, there appears to be a U-shaped relationship between total assets and bid premiums: relatively high for the lowest quartile of banking organizations, decreasing to the next quartile, and rising thereafter. There appears to be little if any noticeable pattern in bid premiums for targets with total assets less than $10 billion. While returns are smaller for megamergers of target banks larger than $10 billion than for other mergers, there is no clear pattern of variation in standardized cumulative abnormal returns within each group.

Table 7 provides information on market concentration and merger prices. These numbers suggest that bid premiums increase with concentration, especially for banks with total assets greater than or equal to $10 billion. Thus, it appears that large

TABLE 6

Size and pricing of bank mergers

	Quartile (1 = lowest)	Book value premium	SCAR[b]
Assets ≥$10 billion[a]	1	2.43	8.61
	2	1.83	3.21
	3	2.50	11.65
	4	2.69	9.27
Assets <$10 billion	1	2.16	2.72
	2	2.03	4.84
	3	2.25	3.74
	4	2.15	4.46

[a]Excluding banking organizations classified as mergers of equals.
[b]Standardized cumulative abnormal returns.

TABLE 7

Market concentration and the pricing of bank mergers

	Concentration quartile (1 = lowest)	Book value premium	SCAR[b]
Assets ≥$10 billion[a]	1	1.61	7.68
	2	2.35	11.93
	3	2.43	3.49
	4	2.93	9.59
Assets <$10 billion	1	1.93	4.15
	2	2.48	5.62
	3	2.00	3.05
	4	2.17	3.27

[a]Excluding banking organizations classified as mergers of equals.
[b]Standardized cumulative abnormal returns.

banks pay more for target banks located in less competitive markets.

Table 8 provides detail on megamergers of targets larger than $10 billion, as well as megamergers of equals—the bid premium and the standardized cumulative abnormal return around the merger announcement date are presented for each merger deal. The standardized cumulative abnormal returns, on average, are much larger for megamergers overall than for megamergers of equals. The problem of melding the culture of a large target bank with that of the acquirer is anticipated by the market to be more serious in megamergers of equals deals. Unlike the standardized cumulative abnormal returns, the bid premiums are approximately the same, on average, for both megamergers of equals and other megamergers.

The statistics presented in tables 4, 5, 6, 7, and 8 are averages, and do not control for the other characteristics of the target, the acquiring bank, the deal, and the year of the merger. We control for these characteristics in the regression analysis presented in tables 9–11.

Table 9 presents the regression analysis explaining the bid premiums offered for targets using the financial characteristics of the target and selected factors associated with the transaction, based on equation 1 in box 3. The first three columns present the results using return on book equity as a measure of profitability. The last three columns present the results using return on assets as a measure of profitability. Columns 1 and 4 of table 9 represent the basic model for each measure of profitability, excluding the time indicator variables and the composite of the Southeast region compact indicator and the Riegle–Neal indicator. Columns 2 and 5 expand the basic equation to include the composite term. Finally, columns 3 and 6 add the time indicator variables that control for the year of each merger announcement.

The results in column 1 of table 9 show that more profitable target banks, as measured by higher return on equity, are offered larger bid premiums than less profitable targets. The bid premium increases with the asset size of the target and decreases with the ratio of equity to deposits, although the effects are not statistically significant. Similarly, the loan-to-assets and net-chargeoffs-to-loans ratios are both insignificant in explaining variation in the bid premium across deals. The results also indicate that bid premiums are larger for target banks than thrifts, larger for targets located in the Southeast regional compact, and larger in the post-Riegle–Neal period. The results in table 9 also suggest that the bid premiums tend to be lower for megamergers of equals than for other mergers. However, this effect is not statistically significant at conventional levels.

Table 9, column 2 includes the composite term that interacts the Riegle–Neal indicator variable with the Southeast regional compact indicator variable. The total impact on the bid premium for target banks located in the Southeast regional compact after Riegle–Neal is the sum of the coefficients on the Southeast regional compact indicator variable, 0.4459, and the Riegle–Neal and Southeast regional compact composite indicator variable, –0.2681. Thus, holding everything else constant, in the post-Riegle–Neal period, bid premiums are lower in the Southeast regional compact states than in the pre-Riegle–Neal period.

Column 3 of table 9 reports the results of including time indicator variables (and excluding the Riegle–Neal indicator variable) and the composite term in the basic regression equation. When we add the time indicator variables, the coefficient estimates on return on equity, thrift indicator, and Southeast regional compact indicator are qualitatively similar to those reported in column 1 of table 9. For example, the coefficient on the Southeast regional compact indicator continues to suggest that mergers between banking organizations located in the Southeast states during the post-Riegle–Neal period result in higher bid premiums than those in other states. The effect of 0.4333 in this specification is even greater than the effect of 0.3439 in the basic model in column 1. In the specification in column 3, both the asset size of the target banks and the ratio of equity to deposits are now significantly correlated with bid premiums. The results suggest that larger target banks receive larger bid premiums. This result is consistent with the notion that banks are using their increased freedom to merge in a way intended to increase the value of their deposit insurance, generating higher bid

TABLE 8

Characteristics of large bank mergers during the 1990s

	Acquiring bank	Total assets of target	Book value premium	Standardized excess return	Year
		($ bil.)			
Targets with assets >$10 billion					
Manufacturers Hanover Corporation	Chemical Banking Corporation	61.5	0.70	7.57	1991
C&S/Sovran Corporation	NCNB Corporation	51.4	1.49	4.62	1991
Security Pacific Corporation	BankAmerica Corporation	88.0	1.17	11.50	1991
Ameritrust Corporation	Society Corporation	11.0	1.99	0.65	1991
Manufacturers National Corporation	Comerica Inc.	12.1	1.34	6.46	1991
MNC Financial Inc.	NationsBank Corporation	17.5	1.33	–1.55	1992
Continental Bank Corporation	BankAmerica Corporation	22.5	1.35	10.20	1994
Michigan National Corporation	National Australia Bank Ltd.	10.2	1.69	10.03	1995
Shawmut National Corporation	Fleet Financial Group Inc.	31.3	1.79	9.18	1995
First Fidelity Bancorporation	First Union Corporation	36.2	1.92	13.82	1995
Midlantic Corporation	PNC Bank Corporation	13.3	2.12	6.54	1995
Integra Financial Corporation	National City	13.7	1.97	4.87	1995
Meridian Bancorp Inc.	Corestates Financial Corporation	15.0	2.17	5.71	1995
First Interstate Bancorp	Wells Fargo & Company	55.8	3.35	18.16	1995
BayBanks	Bank of Boston Corporation	10.8	2.22	5.40	1995
Boatmen's Bancshares	NationsBank Corporation	33.7	2.71	14.35	1996
Standard Fed Bancorp	ABN-AMRO Holding NV	13.3	2.05	–2.76	1996
US Bancorp	First Bank System	31.9	3.38	9.76	1997
Central Fidelity Banks Inc.	Wachovia Corporation	10.6	2.81	8.32	1997
Signet Banking Corporation	First Union Corporation	11.7	3.46	18.64	1997
Barnett Banks	NationsBank Corporation	41.4	4.05	10.28	1997
Corestates Financial Corporation	First Union Corporation	45.6	5.39	3.51	1997
First of American Bank	National City	22.1	3.84	12.58	1997
Average		28.7	2.36	8.17	
Mergers of equals					
KeyCorp, Albany, NY	Society Corporation	25.5	1.82	—	1993
BB&T Financial Corporation	Southern National	9.2	2.32	2.29	1994
First Chicago Corporation	NBD Bancorp	65.9	1.30	1.87	1995
Chase Manhattan Corporation	Chemical Banking Corporation	114.0	1.38	3.76	1995
First Chicago NBD Corporation	Banc One Corporation	114.1	3.68	0.52	1998
BankAmerica Corporation	NationsBank Corporation	260.0	3.06	1.52	1998
Wells Fargo & Company	Norwest Corporation	97.5	2.70	–2.13	1998
Average		98.0	2.32	1.30	

premiums with higher asset size. The ratio of equity to deposits (leverage ratio) is significantly negative, indicating that higher leverage targets are offered larger bid premiums than other leveraged institutions. These less-capitalized target banks are viewed by the acquirers as being more efficient in their use of expensive capital funding; thus, the acquirers are willing to pay a larger bid premium. Finally, the time indicators suggest that bid premiums have been increasing over time. For example, merger bid premiums in 1997 were, on average, 0.6692 percentage points below those in 1998, while in 1996 they were 1.2543 percentage points below the 1998 level.

Columns 4 to 6 of table 9 report the results of using return on assets rather than return on equity as a measure of profitability. As in columns 1 to 3, bid premiums increase with profitability as measured by return on assets. Moreover, the asset size of the target banks and the ratio of equity to deposits are statistically significantly related to bid premiums in almost every empirical specification. The coefficients on the thrift and Southeast regional compact indicators are roughly the same as those reported in columns 1–3 of table 9. The model specification fits well, explaining almost 25 percent of the variation in the bid premiums offered for targets across all merger deals.

Table 10 uses the target's standardized excess returns as the dependent variable in the regression equation rather than the bid premium. As shown in

TABLE 9

Relationship between bid premium and target financial characteristics

	Profitability as return on equity			Profitability as return on total assets		
Variables	**Basic controls**	**Southeast compact interacted with Riegle–Neal**	**Time binary variables**	**Basic controls**	**Southeast compact interacted with Riegle–Neal**	**Basic variables**
Return on equity	3.3169 (2.82)***	3.2695 (2.86)***	3.0487 (2.82)***	—	—	—
Return on assets	—	—	—	39.8565 (2.69)***	40.4137 (2.73)***	31.8314 (2.18)**
Natural logarithm of total assets	0.0636 (1.54)	0.0630 (1.52)	0.0758 (1.89)*	0.0681 (1.65)*	0.0676 (1.63)*	0.0812 (2.02)**
Book value of equity to total deposits	–2.9059 (–1.51)	–2.8209 (–1.46)	–3.8521 (–2.11)**	–4.1854 (–2.14)**	–4.1216 (–2.10)**	–4.9113 (–2.64)***
Loans to total assets	0.7767 (1.31)	0.7357 (1.24)	0.6238 (1.12)	0.7742 (1.31)	0.7341 (1.24)	0.5952 (1.06)
Net chargeoffs to loans	–12.9726 (–1.18)	–12.6326 (–1.15)	–10.4068 (–1.00)	–14.7443 (–1.37)	–14.4692 (–1.34)	–14.2260 (–1.38)
Thrift indicator	–0.5178 (–2.21)**	–0.5191 (–2.22)**	–0.4248 (–1.92)*	–0.4846 (–2.01)**	–0.4859 (–2.01)**	–0.4412 (–1.92)*
Megamergers of equals indicator	–0.1573 (–0.32)	–0.1281 (–0.26)	–0.4647 (–0.99)	–0.0554 (–0.11)	–0.0254 (–0.05)	–0.3568 (–0.76)
Southeast regional compact indicator	0.3439 (2.39)**	0.4459 (2.44)**	0.4333 (2.51)**	0.3287 (2.27)**	0.4274 (2.34)**	0.4197 (2.42)**
Riegle–Neal	0.3577 (2.52)**	0.2855 (1.75)*	—	0.3621 (2.54)**	0.2923 (1.79)*	—
(Southeast regional compact) x (Riegle–Neal indicator)	—	–0.2681 (–0.91)	–0.2753 (–0.99)	—	–0.2598 (–0.88)	–0.2554 (–0.92)
1990	—	—	–2.1043 (–4.88)***	—	—	–2.0594 (–4.74)***
1991	—	—	–1.5333 (–4.92)***	—	—	–1.5081 (–4.76)***
1992	—	—	–1.5709 (–4.85)***	—	—	–1.5937 (–4.89)***
1993	—	—	–0.9094 (–3.23)***	—	—	–0.9140 (–3.23)***
1994	—	—	–1.0488 (–3.75)***	—	—	–1.0410 (–3.70)***
1995	—	—	–1.3827 (–5.55)***	—	—	–1.3468 (–5.37)**
1996	—	—	–1.2543 (–4.96)***	—	—	–1.2362 (–4.85)***
1997	—	—	–0.6692 (–2.71)***	—	—	–0.6642 (–2.59)**
Number of observations	327	327	327	327	327	327
Adjusted R^2	0.1337	0.1358	0.2432	0.1319	0.1313	0.2355
F-statistic	6.592	6.693	7.1641	6.503	5.926	6.908

***Indicates significance at the 1 percent level.
**Indicates significance at the 5 percent level.
*Indicates significance at the 10 percent level.
Note: Numbers in parentheses are t-statistics.

TABLE 10

Standardized excess returns of target and target financial characteristics

	Profitability as return on equity			Profitability as return on assets		
Variables	**Basic controls**	**Southeast compact interacted with Riegle–Neal**	**Time indicator**	**Basic controls**	**Southeast compact interacted with Riegle–Neal**	**Time indicator**
Return on equity	0.1101 (1.55)	0.1249 (1.78)*	0.1189 (1.66)*	—	—	—
Return on assets	—	—	—	0.7750 (0.82)	0.9671 (1.03)	1.0524 (1.10)
Natural logarithm of total assets	0.0074 (2.88)***	0.0076 (3.00) ***	0.0073 (2.77)***	0.0077 (3.03)***	0.0079 (3.16)***	0.0076 (2.87) ***
Book value of equity to total deposits	–0.1361 (–1.07) (–2.52)**	–0.1267 (–1.01) (–2.58)**	–0.1127 (–0.89) (–2.24)**	–0.1598 (–1.23) (–2.57)**	–0.1561 (–1.22) (–2.62)***	–0.1427 (–1.01) (–2.25)**
Net chargeoffs to loans	0.1095 (–0.16)	0.0187 (–0.03)	0.0113 (–0.02)	0.4139 (–0.51)	0.2940 (–0.44)	0.2370 (–0.35)
Thrift indicator	0.0158 (1.10)	0.0172 (1.22)	0.0141 (0.97)	0.0124 (0.83)	0.0140 (0.95)	0.0120 (0.79)
Megamergers of equals indicator	–0.0592 (2.05)**	0.0518 (1.82)*	–0.0445 (–1.51)	0.0561 (1.94)*	–0.0484 (1.69)*	–0.0412 (–1.40)
Southeast regional compact indicator	0.0224 (2.47)**	0.0421 (3.77)***	0.0420 (3.65) ***	0.0226 (2.45)**	0.0418 (3.72)***	0.0416 (3.60)***
Riegle–Neal indicator	0.0068 (0.77)	–0.0061 (–0.63)	—	0.0074 (0.84)	–0.0052 (–0.54)	—
(Southeast regional compact) × (Riegle-Neal indicator)	— (–2.95) ***	–0.0548 (–2.89)***	–0.0543	—	–0.0538 (–2.89)***	–0.0536 (–2.84) ***
1990	—	—	0.0504 (1.98)**	—	—	0.0519 (2.03)**
1991	—	—	0.0306 (1.55)	—	—	0.0308 (1.54)*
1992	—	—	0.0267 (1.21)	—	—	0.0253 (1.14)
1993	—	—	0.0217 (1.14)	—	—	0.0206 (1.08)
1994	—	—	0.0258 (1.38)	—	—	0.0264 (1.40)
1995	—	—	0.0335 (1.92)*	—	—	0.0349 (2.00)**
1996	—	—	0.0302 (1.75)*	—	—	0.0303 (1.75)*
1997	—	—	0.0136 (0.80)	—	—	0.0141 (0.83)
Number of observations	242	242	242	242	242	242
Adjusted R^2	0.0694	0.0994	0.0994	0.0625	0.0912	0.0933
F-statistics	2.996	3.659	2.565	2.784	3.4519	2.458

***Indicates significance at the 1 percent level.
**Indicates significance at the 5 percent level.
*Indicates significance at the 10 percent level.
Note: Numbers in parentheses are t-statistics.

TABLE 11

Bid premiums and financial characteristics with market concentration measure

Variables	Profitability as return on equity	Profitability as return on total assets
Return on equity	3.0500 (2.81)***	—
Return on assets	—	30.9000 (2.18)**
Natural logarithm of total assets	0.0760 (1.88)*	0.0815 (2.02)**
Book value of equity to total deposits	–3.8590 (–2.11)**	–4.9265 (–2.64)**
Loans to total assets	0.6239 (1.12)	0.5944 (1.06)
Net chargeoffs to loans	–10.4294 (–1.00)	–14.2521 (–1.38)
Market concentration	–0.0364 (–0.06)	–0.0625 (–0.11)
Thrift indicator	–0.4243 (–1.91)*	–0.4401 (–1.91)*
Southeast regional compact indicator	0.4323 (2.49)**	0.4179 (2.39)**
(Southeast regional compact) x (Riegle–Neal indicator)	–0.2727 (–0.97)	–0.2509 (–0.88)
1990	–2.1054 (–4.87)***	–2.0611 (–4.74)***
1991	–1.5345 (–4.90)***	–1.5101 (–4.76)***
1992	–1.5720 (–4.84)***	–1.5954 (–4.88)***
1993	–0.9100 (–3.23)***	–0.9150 (–3.22)***
1994	–1.0491 (–3.74)***	–1.0416 (–3.69)***
1995	–1.3819 (–5.53)***	–1.3454 (–5.35)***
1996	–1.2540 (–4.95)***	–1.2355 (–4.84)***
1997	–0.6686 (–2.70)***	–0.6431 (–2.58)**
Number of observations	327	327
Adjusted R^2	0.2408	0.2331
F-statistic	6.744	6.504

***Indicates significance at the 1 percent level.
**Indicates significance at the 5 percent level.
*Indicates significance at the 10 percent level.
Note: Numbers in parentheses are t-statistics.

table 10, the standardized excess returns tend to be greater for more profitable banks. However, the effect is only marginally significant. On the other hand, there is a greater stock market reaction for larger target banks, possibly reflecting the fact that bid premiums tend to increase with target bank asset size. The coefficient on the megamergers of equals indicator variable is negative and, in four out of six cases, statistically significant at conventional levels, indicating that the stock market reacts relatively negatively to announcements of such mergers. For example, the coefficient estimate in column 1 of table 10 suggests that standardized excess returns for these megamergers of equals announcements were 5.92 percentage points less than for other merger announcements. This negative response may be due to the higher cost of melding the culture of two large banking organizations. The Southeast regional compact indicator variable has a positive coefficient in all six specifications in table 10, showing that the target's standardized excess returns around the merger announcement date are significantly larger for targets in the Southeast regional compact states. The composite term that interacts the Riegle–Neal and the Southeast regional compact indicators is negative and statistically significant. Thus, relative to the pre-Riegle–Neal period, the stock market reaction is less in the post-Riegle–Neal period to announcements of mergers of banks located in the Southeast regional compact states. This result is consistent with the notion that bank merger prices will be lower when bank acquisition laws are more liberal.

The target's standardized excess returns around the merger announcement date are significantly lower for targets with higher loan-to-asset ratios. The results indicate that the book value of equity to total deposits and net-charge-offs-to-loans ratios are insignificant in explaining variation in standardized excess returns across deals. As in table 5, the results in table 10 show that standardized excess returns are not statistically different for banks and thrifts.

Table 11 presents the regression analysis explaining merger prices using market concentration, in addition to the financial characteristics of the target and control factors used in tables 9 and 10.

The main results remain pretty much unchanged from table 9. For example, the coefficient on the Southeast regional compact indicator continues to be positive, suggesting that mergers between banking organizations in the Southeast during the post-Riegle–Neal period result in higher bid premiums (and standardized excess returns, which are not shown here but are available from the authors) than those in other states. The results in table 11 indicate that market concentration is not significant in explaining either the variation in the bid premiums that the acquiring banks are willing to pay or the standardized excess returns as a result of the merger announcement, when controlling for characteristics of the targets and the transactions.

Conclusion

The wave of bank consolidation in the 1990s has dramatically changed the structure of the U.S. banking industry. The number of banks has significantly declined, with much fewer smaller banks and more large superregional and money center banks. The market shares of large banks have also become much larger as a result of megamergers and mergers of equals. The rapid pace of bank mergers and acquisitions is likely to continue into the future. Moreover, the pace of bank acquisitions of security firms and insurance companies is also likely to rise in the future as a result of the recent enactment of the Gramm–Leach–Bliley Act of 1999.

This article presents evidence on the different motivations affecting merger bid premiums that the acquiring banks are willing to offer for the targets as well as the announcement-period abnormal stock returns. We find that the following target banks are likely to be offered a larger bid premium—more profitable targets with higher returns on assets and/or returns on equity, and less-capitalized target banks with high leverage ratios.

The positive correlation between target size and the standardized abnormal (excess) returns around the merger announcement date implies that the market views the mergers positively when the potential bidders look for large targets that participate in significant markets. However, the abnormal returns are significantly lower for megamergers of equals than for other mergers, probably because the market anticipates problems in melding the cultures of two large banks. Regarding the capital ratio, unlike bank regulators, which favor better-capitalized banks, acquiring banks tend to prefer targets that offer an ample, inexpensive source of funds. In addition, target banks tend to receive larger bid offers than thrifts.

Our results show that banks located in the Southeast regional compact states—the only group that operated as a cohesive unit in our sample period, restricting entry by banks from states outside of the region—receive larger bid premiums than targets in other parts of the country. Overall, the bid premiums are larger in the post-Riegle–Neal period, consistent with the notion that as the universe of actual or potential bidders has expanded, acquisition prices have risen.

If market participants are able to identify in advance the improved performance associated with bank acquisitions, as documented in Cornett and Tehranian (1992), the bid premiums and the announcement-period abnormal stock returns examined here should be positively correlated with the long-term performance of the merged banks. Smaller bid premiums and abnormal returns at targets larger than $10 billion in megamergers of equals suggest that future megamerger applications between banks and other bank or nonbank financial institutions should be monitored more closely. Our results also imply that these megamergers of equals are not perceived by the market to have the benefit of creating a bank that is too big to fail.

NOTES

[1]Prior to the Riegle–Neal Act, banking organizations could conduct interstate banking operations through "nonbank banks"—those that do not meet the definition of bank. Banks are commonly defined as institutions that both accept demand deposits and make commercial loans.

[2]States could individually opt out of this branching authority or choose to adopt an earlier starting date.

[3]Siems (1996) examines bank megamerger deals in 1990–95, and concludes that market powers are not the primary motivation for the mergers.

[4]See Cetorelli (1999) for a discussion of the HHI.

[5]When first introduced the 1982, the DOJ horizontal merger guidelines listed in table 2 applied to all U.S. industries. In 1985, the DOJ modified the 1,800/50 rule for bank mergers to 1,800/200 to recognize the impact of competition from thrifts and nondepository institutions.

As mentioned earlier, the Supreme Court, in the Connecticut National Bank case (1974), concluded that thrifts should not be included in the calculation of concentration measures because they were not offering the cluster of banking services. However, the court did recognize that thrifts could be included if they became significant competitors for a broad range of consumer services. With the passage of the Depository Institutions Deregulation and Monetary Control Act (1980) and the Garn–St. Germain

Act (1982), which effectively deregulated the thrift industry, thrifts were authorized to compete with banks in providing the cluster of products previously unique to commercial banking organizations. By the mid-1980s competition from thrifts had grown to such a point that the Federal Reserve Board changed its rules regarding delegation of authority of the Federal Reserve Banks to give thrifts a weight of 50 percent when calculating concentration numbers, to reflect both actual and potential competition from thrifts. In some cases, it may give 100 percent weight to thrifts when they are significant competitors.

[6]Healy, Palepu, and Ruback (1991) perform a similar study on nonregulated firms.

[7]Due to the sample bias problem stemming from this criterion, the result may not be applicable to small banks.

[8]An exception is evident, however, for megamergers of equals (shown in table 8). That is, the abnormal returns are significantly smaller for money center banks with total assets more than $10 billion compared with the rest of the population. The smaller stock price gain around the merger date for megamergers of equals of targets larger than $10 billion may represent a higher cost of melding the culture of a large target bank with that of the acquirer.

REFERENCES

Adkisson, J. Amanda, and Donald R. Fraser, 1990, "The effect of geographical deregulation on bank acquisition premium," *Journal of Financial Services Research*, Vol. 4, July, pp. 145–155.

Amel, Dean, 1988, "State laws affecting commercial bank branching, multibank holding company expansion, and interstate banking," Board of Governors of the Federal Reserve System, working paper.

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

Baradwaj, Babu, Donald Fraser, and Eugene Furtado, 1990, "Hostile bank takeover offers," *J ournal of Banking and Finance*, Vol. 14, No. 6, pp. 1229–1242.

Beatty, Randolph P., Anthony M. Santomero, and Michael Smirlock, 1987, "Bank merger premiums—Analysis and evidence," New York University, Salomon Brothers Center for the Study of Financial Institutions, monograph, No. 1987a-3.

Benston, G. J., W. C. Hunter, and L. D. Wall, 1995, "Motivations for bank mergers and acquisitions: Enhancing the deposit insurance put option versus earnings diversification," *Journal of Money, Credit, and Banking*, August, pp. 777–788.

Berger, Allen N., Robert DeYoung, Hesna Genay, and Gregory F. Udell, 2000, "Globalization of financial institutions: Evidence from cross-border banking performance," in *Brookings-Wharton Papers on Financial Services: Toward a Global Financial Market*, Washington: Brookings Institution, forthcoming.

Binder, John J., 1988, "The Sherman Antitrust Act and the railroad cartels," *Journal of Law and Economics*, Vol. 31, October, pp. 443–467.

Boyd, John, and Stanley Graham, 1991, "Investigating the bank consolidation trend," *Quarterly Review*, Federal Reserve Bank of Minneapolis, Spring, pp. 3–15.

Bradley, Michael, Anand Desai, and E. Han Kim, 1988, "Synergistic gains from corporate acquisitions and the division between the stockholders of target and acquiring firms," *Journal of Financial Economics*, Vol. 21, No. 1, pp. 3–40.

Cetorelli, Nicola, 1999, "Competitive analysis in banking: Appraisal of the methodologies," Federal Reserve Bank of Chicago, *Economic Perspectives*, Vol. 23, First Quarter, pp. 2–15.

Cheng, David, Benton Gup, and Larry Wall, 1989, "Financial determinants of bank takeovers," *Journal of Money, Credit, and Banking*, Vol. 21, No. 4, November, pp. 524–536.

Cornett, Marcia Millon, and Sankar De, 1991a, "Common stock returns in corporate takeover bids: Evidence from interstate bank mergers," *Journal of Banking and Finance*, Vol. 15, No. 2, pp. 273–295.

_____, 1991b, "Medium of payment in corporate acquisitions: Evidence from interstate bank mergers," *Journal of Money, Credit, and Banking*, Vol. 23, No. 4, pp. 767–776.

Cornett, Marcia Millon, and Hassan Tehranian, 1992, "Changes in corporate performance associated with bank acquisitions," *Journal of Financial Economics,* Vol. 31, April, pp. 211–234.

Desai, Anand, and R. D. Stover, 1985, "Bank holding company acquisitions, stockholder returns, and regulatory uncertainty," *Journal of Financial Research*, Vol. 8, No. 2, pp. 145–156.

DeYoung, Robert, 1999, "Mergers and the changing landscape of commercial banking," Federal Reserve Bank of Chicago, *Chicago Fed Letter*, September, No. 145.

Dubofsky, D., and Donald R. Fraser, 1989, "The differential impact of two significant court decisions concerning banking consolidation," *Journal of Banking and Finance*, Vol. 13, July, pp. 339–354.

Eckbo, Espen, and Herwig Langohr, 1989, "Information disclosure, measure of payment, and takeover premiums: Public and private tender offers in France," *Journal of Financial Economics*, Vol. 24, October, pp. 363–404.

Fraser, Donald R., and James W. Kolari, 1988, "Pricing small bank acquisitions," *Journal of Retail Banking*, Vol. 10, No. 4, Winter, pp. 23–28.

Frieder, Larry A., 1988, "The interstate banking landscape: Legislative policies and rationale," *Contemporary Policy Issues*, Vol. 6, April, pp. 41–66.

Healy, Paul M., Krishna Palepu, and Richard S. Ruback, 1991, "Does corporate performance improve after mergers?," *Journal of Financial Economics*, Vol. 31, April, pp. 135–175.

Holder, Christopher L., 1993, "Competitive considerations in bank mergers and acquisitions: Economic theory, legal foundation, and the Fed," Federal Reserve Bank of Atlanta, *Economic Review*, Vol. 78, January/February, pp. 23–36.

Hunter, William C., and Larry Wall, 1989, "Bank merger motivations: A review of the evidence and examination of key target bank characteristics," Federal Reserve Bank of Atlanta, *Economic Review*, September, pp. 2–19.

Jackson, William E. III, 1992, "Is the market well defined in bank merger and acquisition analysis?," *The Review of Economics and Statistics*, Vol. 74, November, pp. 655–661.

Jensen, Michael, and Richard S. Ruback, 1983, "The market for corporate control: The scientific evidence," *Journal of Financial Economics*, Vol. 11, April, pp. 5–50.

Kane, Edward J., 1998, "Megabank mergers and the too-big-to-fail doctrine: Some preliminary results," Boston College, unpublished paper.

Kwast Myron L., Martha Starr-McCluer, and John D. Wolken, 1997, "Market definition and the analysis of antitrust in banking," *The Antitrust Bulletin,* Vol. 42, Winter, pp. 973–995.

Neely, Walter, 1987, "Banking acquisitions: Acquirer and target shareholder returns," *Financial Management*, Winter, pp. 66–74.

Palia, Darius, 1994, "Recent evidence on bank mergers," *Financial Markets, Institutions, and Instruments*, Vol. 3, No. 5, pp. 36–59.

_____, 1993, "The managerial, regulatory, and financial determinants of bank merger premiums," *Journal of Industrial Economics*, Vol. 41, March, pp. 91–102.

Rhoades, Stephen A., 1987, "Determinants of premiums paid in bank acquisitions," *Atlantic Economic Journal*, Vol. 15, No. 1, March, pp. 20–30.

Rogowski, Robert J., and Donald G. Simonson, 1989, "Bank merger pricing premiums and interstate bidding," in *Bank Mergers: Current Issues and Perspectives,* Benton E. Gup (ed.), Amsterdam: Kluwer Academic Publishers, pp. 87–106.

Rose, Peter S., 1991, "Bidding theory and bank merger premiums: The impact of structural and regulatory factors," *Review of Business and Economic Research*, Vol. 26, No. 2, Spring, pp. 22–40.

Saunders, A., 1997, *Financial Institutions Management: A Modern Perspective*, second edition, Chicago: Richard D. Irwin.

Savage, Donald T., 1993, "Interstate banking: A status report," *Federal Reserve Bulletin*, Vol. 73, December, pp. 1075–1089.

Schipper, Katherine, and Rex Thompson, 1983, "The impact of merger-related regulations on the shareholders of acquiring firms," *Journal of Accounting Research*, Vol. 21, Spring, pp. 184–221.

Scholes, Myron, and Joseph Williams, 1977, "Estimating betas from nonsynchronous data," *Journal of Financial Economics*, Vol. 5, No. 3, pp. 309–327.

Shawky, Hany, Tobias Kilb, and Carsten Staas, 1996, "Determinants of bank merger premiums," *Journal of Economics and Finance*, Vol. 20, No. 1, pp. 117–131.

Siems, Thomas, 1996, "Bank mergers and shareholder wealth: Evidence from 1995's megamerger deals," *Financial Industry Studies*, Federal Reserve Bank of Dallas, August, pp. 1–12.

Stultz, R. M., R. Walking, and M. H. Song, 1990, "The distribution of target ownership and the division of gains in successful takeovers," *Journal of Finance*, July, pp. 817–833.

Sushka, M. E., and Y. Bendeck, 1988, "Bank acquisitions and stock holders' wealth," *Journal of Banking and Finance*, Vol. 12, No. 4, pp. 551–562.

Travlos, Nickolaos, 1987, "Corporate takeover bids, methods of payment, and bidding firms' stock returns," *Journal of Finance*, Vol. 42, September, pp. 943–964.

Trifts, J. W., and K. P. Scanlon, 1987, "Interstate bank mergers: The early evidence," *Journal of Financial Research*, Vol. 10, No. 4, pp. 305–311.

Dollarization in Argentina

François R. Velde and Marcelo Veracierto

Introduction and summary

In January 1999, Argentina announced that it was considering adopting the U.S. dollar as its sole medium of exchange. This policy proposal, which is known as "dollarization," received considerable attention from both policymakers and the media, generating an ongoing debate. This article discusses from a critical perspective some of the issues raised in this debate. Although we do not reach a definite answer on whether Argentina should dollarize, we believe that our work sheds considerable light on the costs and benefits associated with it.

The debate over dollarization is part of a broader, longstanding, and ongoing debate over the relative merits of monetary arrangements. The general question is whether a country's currency should be tied to some anchor, and, if so, to which anchor and how tied. The question involves a variety of issues, depending on the context in which it is raised. In the international context, this question becomes the debate about fixed and flexible exchange rates and optimal currency areas.[1] Dollarization is simply the most extreme form of a fixed exchange rate. When one abstracts from international considerations, as one would for a relatively closed economy like Argentina's, in which international trade matters less, the context is the debate over "rules versus discretion": Should monetary policy be tied to a rigid rule or should central bankers be allowed discretion in their conduct of policy? Dollarization is the ultimate rule, or the total absence of discretion.

While the choice of anchor for monetary systems has been debated for centuries, the question of dollarization has been posed relatively recently. Indeed, Mundell wrote in his classic paper (Mundell, 1961) that "it hardly appears within the realm of political feasibility that national currencies would ever be abandoned in favor of any other arrangement." More recently, Schwartz (1993) wrote in her review of the history of currency boards that "central banks seem to me strongly entrenched and unlikely to be dislodged even if their policies create hyperinflations." Yet currency boards have made a comeback of sorts, with Hong Kong since 1983 and Argentina since 1991 as the most prominent examples.[2] Dollarization has been evoked in Argentina. But the debate has sprung up elsewhere. Just as the European common market led to European monetary union, some have argued that the members of the North American Free Trade Agreement, particularly Mexico, should seriously consider dollarization. Most recently, on January 9, 2000, the president of Ecuador announced plans to immediately dollarize his country's economy, retaining the local currency only for small change.

American officials have repeatedly taken a very balanced position on the matter; while not rejecting the idea out of hand, and while admitting that the U.S. could not prevent a country from adopting the dollar as currency, they have issued strong cautionary notes. At present, following the election of a new president on October 25, Argentina has stated a strong commitment to the current currency board arrangement, and U.S. Secretary of the Treasury Lawrence H. Summers recently concluded that the question of dollarization is not on Argentina's agenda. The topic, however, has now raised interest in academic and business circles.

This article restricts attention to the particular case of Argentina. Argentina has a long history of

François R. Velde and Marcelo Veracierto are economists at the Federal Reserve Bank of Chicago. *The authors thank Fernando Alvarez, Larry Christiano, John Cochrane, and David Marshall for helpful discussions and Loula Merkel for outstanding assistance.*

disastrous monetary policies and repeated hyperinflations, which have led it to peg its currency to the dollar since 1991. Since Argentina is in practice already quite close to being fully dollarized, it presents a good illustration of what is (and is not) required for a successful dollarization and what are the costs and benefits associated with it.

We first present the facts about Argentina's case, in particular the historical background to Argentina's peg to the dollar since 1991. We then describe the possible forms that dollarization could take, present the benefits that have been suggested, consider possible costs and objections, and carry out a rough cost–benefit comparison.

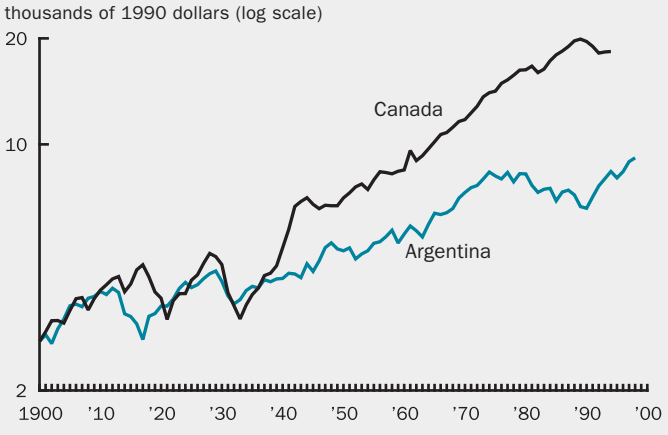
The facts of Argentina's case

At the turn of the twentieth century, Argentina was one of the ten or 15 richest countries, and its gross domestic product (GDP) per capita was only 40 percent lower than that of the world leader (the United Kingdom). In fact, GDP per capita in Argentina stood at the same level as in Canada, a country similar in many respects in terms of physical and human endowments.

Figure 1 shows the subsequent paths taken by Argentina and Canada over the course of the twentieth century. Both were similarly affected by the Great Depression of 1929 and the trade wars that followed in the 1930s. After World War II, however, their paths begin to diverge noticeably. And, while Canada's growth is strong and smooth, Argentina's growth is weaker, and subject to greater fluctuations. The paths take opposite directions in the 1970s, when Argentina's income actually falls. At present, Argentines are half as rich as Canadians. The gap depicted in figure 1 is often viewed as a measure of Argentina's wasted opportunities.

FIGURE 1

GDP per capita, Argentina and Canada



Sources: Maddison (1995) until 1994, Ministerio de Economía (MECON) thereafter.

Perhaps not coincidentally, Argentina has a very long history of unstable monetary policies, stretching back to the nineteenth century. After independence in 1810, it took the country until 1853 to reach a constitutional agreement. Until 1881, the currency consisted mainly of paper money, issued by various local administrations, that was not redeemable in gold or silver. Attempts to set up a monetary system on a gold standard began in 1881 but were not successful until 1899, when the outstanding mass of paper money was made convertible into gold. Convertibility, suspended on August 8, 1914, when World War I broke out, was not resumed until 1927. It was suspended again in December 1929, when the country was hit by a combination of falling commodity prices, mounting government deficit, loss of access to foreign capital markets, and incipient currency speculation (Eichengreen, 1992). A coup in 1930 overthrew the elected government and inaugurated a long period of military regimes interspersed with occasional elections, until full democracy returned in 1983.

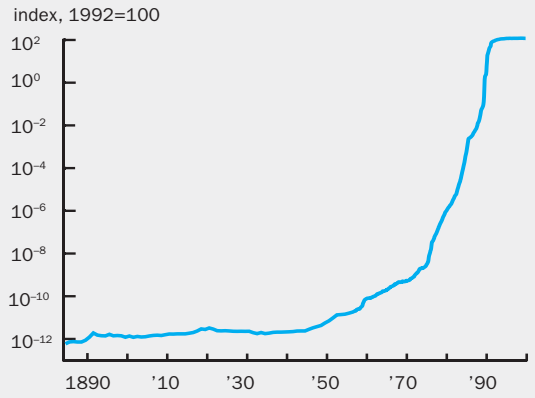
Figure 2 plots the price level over time. It shows how Argentina's familiarity with inflation is long-standing. For example, the price level doubled from 1889 to 1891. Such experiences pale in comparison with what happened after the end of the gold standard in 1929, the establishment of the central bank in 1935, and its role in monetizing deficits (that is, financing deficits with the printing press) from 1943 on. From 1943 to the present, the price level has gone up by a factor of 10 in the U.S.; in the same period, it has gone up by 10^{12} in Argentina. The main recent episodes of high inflation, in 1975, from 1982 to 1985, and from 1987 to 1990, are visible in figure 2 as sharp accelerations of the price level.

Figure 2 also shows that a remarkable change took place in the early 1990s. When Carlos Menem was first elected president of Argentina in May 1989, the inflation rate had reached 78 percent per month. To put an end to inflation, Congress passed the "convertibility law" in March 1991, establishing the convertibility of the *austral* (the Argentine currency since 1985) into the U.S. dollar at a rate of 10,000 australes per dollar. In January 1992 the *peso* replaced the austral, at a rate of 1 peso for 10,000 australes.

The regime instituted by these reforms places strict limits on the Argentine Central Bank's policy. Under the convertibility

FIGURE 2

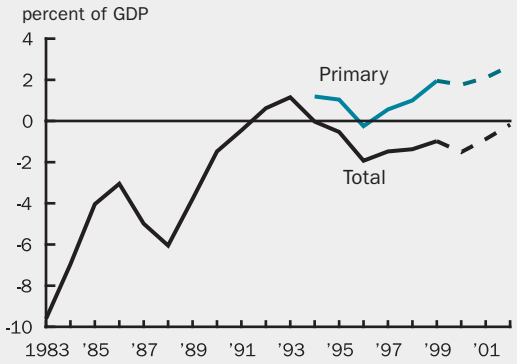
Consumer price index, Argentina



Sources: Rolnick and Weber (1997) and International Monetary Fund.

FIGURE 3

Government surplus as share of GDP



Note: Projections, indicated by dashed lines, for 1999–2002 from budget submitted to Congress in September 1999.
Source: MECON.

law, the central bank must stand ready to sell dollars for pesos at the rate of 1 U.S. dollar per peso. Free reserves, consisting of gold, foreign currency, or deposits and bonds payable in gold and foreign currency, must be maintained at a level no less than 100 percent of the monetary base.

The central bank is forbidden by its charter, passed in 1992, from lending to the government. However, the formula for the classic currency board requires full backing of the currency with foreign reserves only. In Argentina, the bank is allowed to hold Argentine government bonds as part of the backing of the monetary base. But this departure from the classic currency board is minor, for the following reasons. Those holdings must be purchased at market price (so they are not direct loans to the government), they cannot exceed 33 percent of total reserves,[3] and they cannot increase by more than 10 percent in any year.

The peso (or its predecessor the austral) has been convertible with the dollar at a constant rate for over eight years. As figure 2 shows, the Argentine price level was quickly stabilized and has remained stable. Implementation of the currency board arrangement has been accompanied by a number of other reforms. The Argentine government reduced both spending and taxes, and quickly eliminated the deficit, as shown in figure 3.[4] It also privatized many state-owned companies and carried out other major reforms, including trade liberalization, freeing of international capital flows, and deregulation of the banking industry.

These reforms appear to have had beneficial effects in the Argentine economy. Figure 4 shows an index of real GDP. After a long period of stagnation, the growth rate of output went from a –1 percent annual average between 1980 and 1990 to 4.3 percent between 1991 and 1998. The effect on a per capita basis is strikingly displayed in figure 5 (which also plots monthly inflation). Real output per capita fell 23 percent during the 1980s, and this fall was more than reversed up to 1998.

The expansion of the 1990s was interrupted twice, as shown in figure 4: in the aftermath of the Mexican balance of payments crisis in January 1995 (the so-called Tequila effect), and again in the recent international turmoil following the Russian default of August 1998 and the Brazilian devaluation of January 1999 (the "Vodka–Caipirinha effect").

Although the Argentine peso has remained pegged at 1 dollar since 1991, it has not been immune

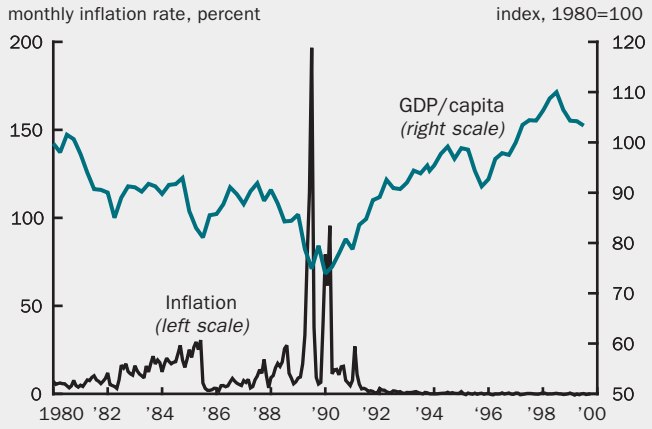
FIGURE 4

Index of Argentine real GDP, 1980:Q1 to 1999:Q2

Note: Based on 1986 prices to 1993; 1993 prices thereafter.
Source: MECON.

FIGURE 5

Argentina's real GDP per capita and inflation

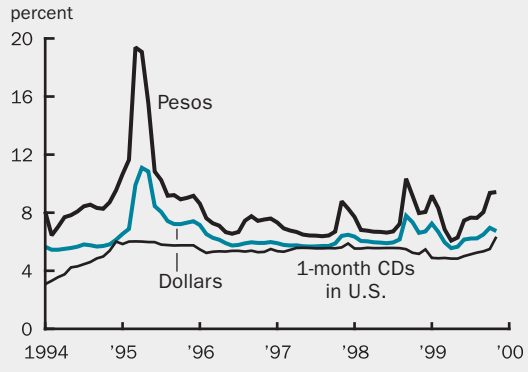


Notes: Real GDP is quarterly from 1980:Q1 to 2000:Q2 and inflation rate is monthly from January 1980 to September 1999.
Sources: MECON and Banco Central de la República Argentina (BCRA).

to speculations about how long this regime will last, and what will replace it. Those questions are raised when currencies elsewhere fall victim to crises. After the Tequila crisis of 1995, and again after the Vodka–Caipirinha crisis of 1998–99, there was intense speculation on a possible devaluation of the peso, in spite of limited links between the affected countries and Argentina. For example, Mexico's share of Argentine exports was only 1.7 percent in 1994, and Argentina's exports overall accounted for just 9 percent of its GDP. Similarly, although Brazil is Argentina's main trading partner, exports to Brazil only represented 3 percent of Argentine GDP in 1998. Interest rates rise sharply with each speculative attack, as shown in figure 6. Furthermore, the premium in interest rates on peso-denominated loans over dollar-denominated loans rises as well, suggesting that the perceived risk of a devaluation is much higher.

FIGURE 6

Interest on 30–59 days time deposits



Notes: Monthly averages from January 1994 to October 1999.
Sources: BCRA and Board of Governors of the Federal Reserve System.

Two days after Brazil devalued the real on January 13, 1999, it became known that Argentina's President Carlos Menem had asked his finance minister to study the feasibility of dollarization. On January 22, the president of Argentina's central bank, Pedro Pou, confirmed that such studies were underway, and that a working group was to be formed by the U.S. Treasury. Over the course of 1999, other events affected the markets' perceptions of Argentina's commitment to the currency board. In mid-May, comments by the creator of the currency board, Domingo Cavallo, were misreported in the *Financial Times* as suggesting that the peg could be modified or abandoned. In August, Ecuador (a Spanish-speaking country south of the Panama Canal, but with no other meaningful relation to Argentina) fell into default on its international debt. The Argentine presidential elections took place on October 25 and ended in the victory of the Radical candidate Fernando de la Rúa, who took over from the Peronist Menem in early December.

Figure 7 shows the response of markets to such news, indicating how sensitive interest rates are to the perception of a possible devaluation. Along with figure 6, it suggests that, most of the time, the level of interest rates in Argentina is not very different from that in the U.S., but that, when doubts are raised about the convertibility of the Argentine peso, interest rates in Argentina can rise quickly to very high levels and be very volatile.

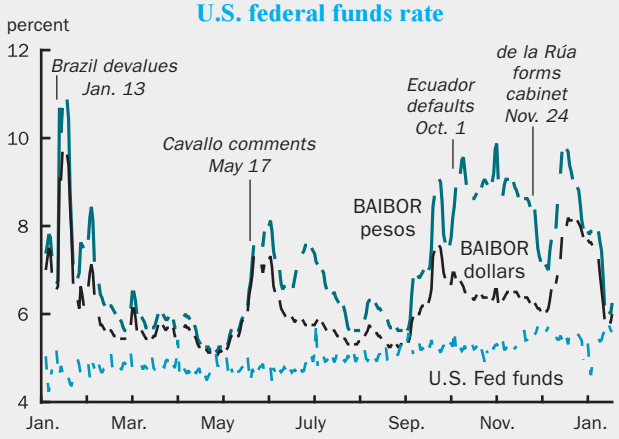
Since markets appear uncertain about Argentina's commitment to its currency board, and since recurrent fears of devaluation have severely affected the economy in the past, it is apparent that Argentina needs to make its currency board fully credible. This is what led the previous Argentine administration to consider the possibility of fully "dollarizing" the economy.

What dollarization is

Dollarization means the total elimination of the Argentine currency, the peso, and its complete replacement with the U.S. dollar. At present, the monetary base in Argentina consists of the peso-denominated currency. If Argentina dollarized, this monetary base would be converted into dollar-denominated currency,

FIGURE 7

Buenos Aires interbank offered rate and U.S. federal funds rate



Notes: Interbank offered rate (BAIBOR) for one-day peso and dollar loans. Rates are graphed daily from January 4, 1999, through January 12, 2000, excluding weekends and holidays.
Sources: BCRA and Board of Governors of the Federal Reserve System.

that is, U.S. Federal Reserve notes. Transactions would be made in dollars, accounts would be kept in dollars, and debts and contracts would be denominated in dollars. The U.S. dollar would be the sole legal tender.

The Argentine economy is already partly dollarized. For example, 61.3 percent of private nonfinancial sector deposits are currently denominated in dollars. The reserve requirements of commercial banks are met with dollar-denominated assets. Argentines are already used to quoting prices and carrying out transactions in dollars. Complete dollarization would not dramatically change their habits and practices.

There are two ways in which dollarization could be implemented. One is for Argentina to proceed on its own; the other is for Argentina to negotiate a formal arrangement with U.S. authorities.

Unilateral dollarization

For Argentina to dollarize, the only requirement is to eliminate the peso-denominated monetary base. Since January 1995, commercial banks have held reserves in dollar-denominated assets instead of peso deposits at the central bank.[5] Thus, the monetary base is just the currency in circulation. To replace the currency, Argentina needs to take the "liquid reserves" that currently back the monetary base, sell them for dollars, and exchange all outstanding peso notes for dollar notes. Once that has been accomplished, the peso has been eliminated, and the only legal tender is the U.S. dollar. Then, all peso-denominated deposits, debts, securities, and contracts are relabeled and become dollar-denominated.

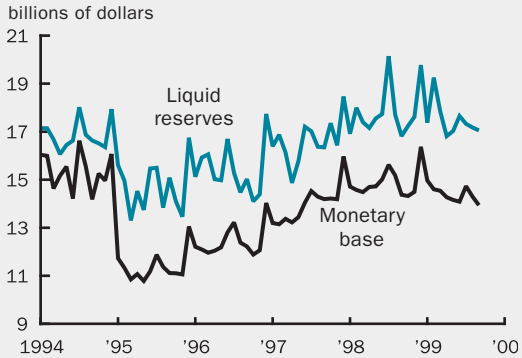
To carry this out, Argentina needs to have enough resources to buy the required amount of dollars. That is already the case under the convertibility law. Figure 8 compares the liquid reserves held by the central bank with M0, the monetary base. As of December 31, 1999, the central bank holds $19.0 billion in reserves (excluding its holdings of Argentine bonds), while the monetary base is $16.5 billion. Thus, Argentina has more than enough to unilaterally liquidate its reserves on the world market and acquire the dollar notes.

In order to dollarize, Argentina has to buy noninterest-bearing dollars with the interest-bearing reserves it has accumulated. These reserves bear interest at present, and therefore are a source of income for Argentina. This income is called seigniorage, and comes from the structure of any central bank's balance sheet: its liabilities (money) bear no interest, while its assets do. But once Argentina's reserves are replaced by dollars, this source of income disappears. Instead, the U.S. will collect the seigniorage. A consequence of dollarization, therefore, is a transfer from Argentina to the U.S.

How large would that transfer be? According to the central bank's income statement, the income on liquid reserves (excluding government bonds) in 1998 amounted to $808 million, an average nominal rate of return of 4.7 percent, or a real rate of return of 3.1 percent (subtracting the U.S. inflation rate) or even 4 percent (subtracting the Argentine inflation rate).

FIGURE 8

Monetary base and liquid reserves



Note: Data from January 1994 through October 1999.
Source: BCRA.

Since liquid reserves averaged $17.2 billion in 1998, in excess of the monetary base which averaged $14.9 billion, only $700 million actually represents seigniorage, that is, income on the reserves that back the monetary base. Since nominal GDP was $298 billion in that year, seigniorage represented 0.2 percent of GDP, or 1.2 percent of government revenues.

This number seems small, seen as a flow. But seigniorage is collected every year. To calculate the value of all future seigniorage that would accrue to the Argentine government from the peso-denominated monetary base if it did not dollarize, we would need to estimate what the monetary base would be in the future and use a discount rate to compute the net present value.

Let M be the current value of the monetary base and R the nominal rate of return on liquid reserves. Suppose that the monetary base grows at a constant rate α, so that, at any future date t, the monetary base is $(1+\alpha)^t M$, and the seigniorage collected on that monetary base is $R(1+\alpha)^t M$. Assume that future sums are discounted at the same nominal rate R. Then the net present value of future seigniorage is given by:

$$\sum_{t=1}^{\infty} \left(\frac{1}{1+R}\right)^t R(1+\alpha)^t M.$$

Using the formula for a geometric sum, the present value of future seigniorage is

$$(1+\alpha)\frac{RM}{R-\alpha}.$$

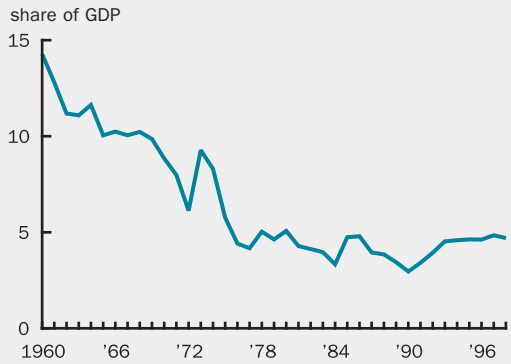
If $\alpha = 0$, that is, the monetary base is not assumed to grow at all, then that net present value is exactly the current monetary base M. If the monetary base grows in the future, then Argentina will continue to bear annual costs, namely, the real assets that it will need to accumulate (through exports) to buy dollars for use at home as currency.

In other countries, the monetary base usually grows at roughly the same rate as nominal output. Figure 9 shows that this has been true in Argentina since the stabilization, although not before.[6] Under the assumption that $M0$ grows at the same rate as nominal output, then α is the sum of the growth rate of real output and the inflation rate. The denominator of our expression becomes the real rate of interest less the real growth rate.

An assumption of 6 percent nominal rate, 4 percent real rate, and 3 percent growth rate yields a present value of six times the current monetary base,

FIGURE 9

Currency in circulation as share of GDP



Source: International Monetary Fund.

or $84 billion. One should remember, however, that this value is very sensitive to the assumptions about rates of interest and growth rates, and that some assumptions will make the denominator of our expression very small. We can, nevertheless, keep in mind a number like 0.2 percent of GDP as the size of the permanent annual transfer from Argentina to the U.S. that would follow dollarization.

Bilateral dollarization

In view of this transfer, it is not surprising that the Argentine government is currently seeking to dollarize the economy within some form of monetary association with the U.S. that would reduce the size of this transfer.

One possible arrangement, which would allow Argentina to avoid the transfer altogether, would be as follows. Instead of letting Argentina sell its reserves on the open market for dollar bills, the Federal Reserve could print an amount of dollar notes equivalent to the total currency in circulation in Argentina ($14 billion) and hand it over to the Argentine government to retire the outstanding peso notes. This would allow Argentina to retain the reserves that are currently backing the peso notes in circulation and to keep the corresponding interest income. From the point of view of the U.S., this operation only involves costlessly printing pieces of paper and shipping them to Argentina. The pieces of paper would remain in Argentina as a medium of exchange, because Argentina needs a medium of exchange. But if Argentina were to introduce a currency again, then Argentine citizens would not need the pieces of paper as a medium of exchange, and would then redeem them for goods and services in the U.S. The $14 billion initially printed by the U.S. to dollarize Argentina would then

result in a transfer of goods from the U.S. to Argentina and a (undesired) monetary expansion in the U.S.

One problem, then, is guaranteeing that pesos are never reintroduced by Argentina. Retiring the peso is a mechanical operation; committing a government to future actions is much harder. This problem could be handled as follows. The monetary association could require Argentina to put the assets that are currently backing the monetary base in an escrow account in a third country (say, Switzerland). As long as Argentina never issues a national currency, it would receive the corresponding interest income, which is the seigniorage. (This seigniorage income could be shared with the U.S., depending on the terms of the treaty). However, if Argentina ever tried to issue a national currency, the U.S. would seize the assets. This arrangement uses Argentina's reserves (excluding Argentine bonds, obviously) as collateral. By making Argentina pay a high price for reintroducing pesos, this arrangement would give enough assurance to the U.S. that Argentina would never renege the monetary association. Argentina would also benefit from such an explicit commitment device. Since investors would be more easily convinced that Argentina would never reintroduce the peso, they would demand lower interest rates on their Argentine investments.

The problem with this arrangement, however, is that it only allows Argentina to retain the seigniorage on the initial stock of currency. But, as Argentina's output grows, demand for media of exchange may grow as well. After dollarization, Argentina would increase its currency stock by acquiring dollar notes with trade surpluses. This would work through arbitrage: If the demand for currency grows but the supply is constant, the price level in Argentina will fall relative to the U.S., prompting an export of goods from Argentina to the U.S. in exchange for dollars. Future increases in the money stock would thus take place in a decentralized way. There is no simple way to extend the escrow arrangement to account for these future increases without estimating each year the growth rate of the monetary base in Argentina. But that monetary base is in dollars, and it would be just as difficult as counting the dollar bills in circulation in, say, the Seventh Federal Reserve District. Yet these future increases could be quite substantial: As we saw in the previous section, the current monetary base may represent only one-sixth of all future seigniorage.

There are, thus, a number of practical difficulties with the idea of a monetary treaty between the U.S. and Argentina. It is also not clear what advantages the U.S. could draw from such an association, which would have to be approved by the U.S. Congress.[7]

The benefits of dollarization

The previous section described how dollarization could be implemented. But what benefits would dollarization have in a country like Argentina?

Credibility

Fixed exchange rates always present credibility problems and are subject to self-fulfilled speculative attacks. The reason is that if investors believe that the central bank will devalue the currency, they will want to exchange their peso assets into dollars, reducing the reserves of the central bank. If, for some reason, everybody believes that a devaluation will take place, the reserves of the central bank could be depleted, forcing it to devalue the peso. The advantage of having a currency board is that investors are guaranteed that the central bank will never run out of reserves (since its reserves exceed the currency in circulation).

Despite the fact that Argentina has been under a currency board since 1991, fears of devaluation are still present, as the Tequila and Vodka–Caipirinha effects have shown. Apparently, what investors fear is that the Argentine government will not be willing to lose all its reserves to maintain the convertibility of the peso, and that it will devalue if the run against the peso is large enough. At first glance, these fears seem unwarranted since the currency board has been, after all, established by law, and it would take another law approved by both houses of Congress to repeal it. However, the Argentine executive does have emergency powers that would allow it to suspend convertibility immediately by decree, subject to ratification by Congress after the fact. Going to the extreme, one can always imagine that a dishonest central bank may disobey the law, or that a coup may take place. (Argentina had coups in 1930, 1943, 1946, 1951, 1966, and 1976.) Certainly a currency board provides a stronger commitment device than having the government promise that it will never devalue the peso, but it is not perfect.

Dollarization would provide a much stronger commitment device, especially if it were done through a bilateral arrangement. If Argentina proceeded unilaterally, one can imagine that the Argentine government could find ways of reintroducing a national currency in the future. But it would be extremely difficult for Argentina to do so if an international treaty explicitly prohibited it. In this sense, a bilateral agreement would make Argentina's commitment more credible than unilateral dollarization.[8] In any case, even if dollarization were done unilaterally, it would be difficult for Argentina to reintroduce a national currency in an unanticipated manner.

It should make foreign investors feel safer about the returns of their investments.

Debt crises

Debt crises can be thought of as situations in which the repayment prospects of a country's sovereign debt (or its private sector debt) are sharply downgraded by international capital markets. In other words, the default risk is reevaluated. A debt crisis can occur because of objective, "fundamental" reasons, such as a radical modification of the components of the government's budget constraint: increased spending (either present, or in the form of future liabilities) or reduced revenues (because of a recession or internal turmoil). Such reasons have been put forward by Burnside, Eichenbaum, and Rebello (1999) to explain the Asian crises.

Alternatively, a debt crisis can be driven purely by expectations on the part of international markets that the government of said country is about to default, resulting in a drying up of lending to that country. The country's government is then faced with the choice of repaying the maturing debt (and thereby maintaining some hope of convincing lenders to return in the future) or simply defaulting and sparing itself the trouble of repaying the existing loans. Cole and Kehoe (1998) have shown that a government may, under normal conditions, have no incentive to default, but would decide to default if faced with foreign lenders who are convinced that it will. Such a default is purely driven by expectations.

At first glance, dollarization per se seems to have little relation to the mechanics of a debt crisis. In fact, the ability to default does not appear to depend on the ability to issue domestic currency. For instance, even in the U.S., where states have not repudiated debt in over a century, bond ratings vary from AAA (Minnesota) to A (New York). When debt crises are due to "fundamental" reasons, dollarization cannot do much to prevent them.

However, dollarization may play an important role in preventing debt crises that are driven purely by expectations. Let us suppose, for example, that the government cares mainly about the revenues it raises over time, say, to spend on its constituents. Defaulting spares the government from having to meet its obligations, increasing the funds available for spending now and in the future. The costs stem from severely diminished access to foreign credit, which can impair the country's growth and the government's tax base. If seigniorage is an option that becomes available after a default (because the government does not care about its international reputation anymore), the cost of default is smaller than if it does not become available. As a consequence, investors will rationally believe that the government will move more easily toward default if seigniorage is an option. This increases the likelihood of a self-fulfilling debt crisis. Since dollarization takes away the government's ability to raise seignorage, it may be a factor in reducing country risk.[9]

Some common objections

Certain issues raised in the debate over dollarization, in our opinion, have obscured the debate. These are issues related to the role of a lender of last resort and to the independence of monetary policy.

The role of a lender of last resort

One of the most frequent objections raised against dollarization is the loss of a domestic lender of last resort. Central banks have long performed the function of providing emergency funds to otherwise sound banks suffering a run, and fulfillment of this function was the main objective of the U.S. Federal Reserve Act of 1913, which established the Federal Reserve System.

The function of a lender of last resort is to be able to instantaneously provide liquidity to a bank. Given that banks obligate themselves to provide funds on demand to depositors, but hold assets that can be difficult to sell quickly, a bank can be in a situation where depositor demands exceed its liquid assets, and the bank is forced to suspend its payments and cease to operate. This can adversely affect the economywide system of payments.

Central banks that are free to create money have a particular ability to provide such liquidity. Since dollarization takes away that ability from central banks, it is feared that it would make bank runs more likely. But there are other ways to marshal liquidity. Furthermore, there are other ways to prevent bank runs. In fact, Argentina has devised such ways, in the wake of the Tequila crisis.

The Tequila crisis in Argentina can be seen in large part as a run on the Argentine banking sector, prompted by speculation and fears about the convertibility of the peso. After Mexico abandoned the peg of its currency in January 1995, total deposits fell 13 percent from January to March 1995, but the composition of deposits remained virtually the same: Dollar-denominated deposits fell from 57.7 percent to 57.2 percent of the total. Withdrawals were affecting dollar deposits as well as peso deposits.

Argentine officials learned the lesson, and implemented several mechanisms to deal with bank runs. One mechanism is the traditional imposition of liquidity requirements on banks. Originally, Argentine banks were required to hold peso reserves at the central bank, just like banks in the U.S. But the banking crisis of 1995 was brought about by doubts over the convertibility of the peso, and, at such a time, peso reserves did not offer strong assurances. This led to a radical change in the reserve requirements. As of August 1995, there are no longer any reserves held at the central bank. Banks now meet the requirements with a variety, broadened over time, of interest-bearing, dollar-denominated financial instruments, either foreign assets or domestic assets held with a put option against an A-rated foreign bank (the put option allows the bank to sell the domestic asset to the foreign bank in exchange for foreign assets). The central bank has total discretion in setting the reserve requirements. Each depositor has a claim on these reserves up to $5,000. In October 1999, these reserves amounted to $17.1 billion, about 21 percent of deposits. By way of comparison, the reserves of the U.S. financial system amount to 1.3 percent of deposits.

A second mechanism is the use of the foreign exchange reserves that the central bank has accumulated in excess of the requirements of the convertibility law. A law passed in April 1995 authorizes the central bank to lend these excess reserves to illiquid banks on a short-term basis against collateral. As of November 23, these reserves stood at about $3.4 billion, or about 4 percent of private sector deposits.

A third mechanism is a deposit insurance fund, created in May 1995, to which banks must contribute on a risk-adjusted basis; it is intended to reach the level of 5 percent of deposits. Deposits are insured up to $30,000 each. Again, by way of comparison, the U.S. Federal Deposit Insurance Corporation's bank insurance fund amounts to 1.3 percent of total insured deposits.

Finally, in December 1996 Argentina arranged a collection of contingent repurchase contracts with a consortium of (currently 14) private foreign banks. Each contract gives the central bank the right to enter at any time t of its choosing, and for a duration T of its choosing up to T_{max} (between two and five years depending on the contract), into a repurchase agreement of Argentine government bonds for U.S. dollars with that foreign bank: The central bank sells the bonds at t and repurchases them at $t + T$. The repurchase price implies a rate of LIBOR (London interbank offered rate) plus 200 basis points. The contracts are renewed every three months by mutual consent. Thus, if a bank cancels its contract at t, it is still obligated to enter into the repurchase agreement up to $t + T_{max}$. In this manner, the central bank can avail itself of liquidity quickly in case of a crisis. In October 1999 the facility amounted to $7.35 billion, about 9 percent of total deposits; the goal is to keep it at about 10 percent of deposits. The cost of the facility is 32 basis points per year ($23 million per year). An interesting clause in the contracts is that the facilities are all void and the foreign banks are freed from all obligations if Argentina defaults on its sovereign external debt.

This last mechanism is of interest because it suggests that, ultimately, the ability to play the role of lender of last resort rests on the government's taxing power. The contingent repurchase facility allows the central bank to translate this future source of funds into immediately available funds without any need for the printing press.

Put together, these mechanisms provide Argentina with protection for about 39 percent of its deposits (that is, M3), or more than 2.4 times the monetary base. How extensive is this protection in comparison with that afforded by a central bank with the discretionary power to act as lender of last resort? The most notable use of the lender-of-last-resort power by the U.S. Federal Reserve in recent times occurred after the stock market crash of October 1987. In the week that followed, the monetary base increased by 1.3 percent, the largest weekly increase of the past 25 years. That action was deemed sufficient to prevent a liquidity crisis in the U.S. financial system. The Argentine central bank's contingent repurchase facility alone provides it with the ability to increase the monetary base by 50 percent. Argentina's protections are thus substantial.

Independence of monetary policy

Another common objection to dollarization is that Argentina would lose its ability to conduct monetary policy: It would be unable to pursue expansionary monetary policy during recessions. On top of this, Argentina may be subject to increases in the U.S. federal funds rate precisely when it most needs the rate to go down, that is, during recessions in Argentina.

The concern that Argentina will not have an independent monetary policy is an important one. However, we need to keep two key points in mind.

First, it is admittedly not clear that dollarization will lead to better outcomes than a *good* independent policy. However, a choice between a good independent policy and dollarization may not be the choice that Argentina faces. Argentina's independent monetary policies of the past are illustrated in the high inflation rates of figure 5. It appears that successive

Argentine governments could not resist the temptation of using the printing press to finance persistently large deficits, with disastrous consequences for the economy.[10] There are theoretical reasons to believe that the country could be better off tying itself to a simple monetary policy rule than resorting to discretionary policy, as box 1 illustrates.

The second key point is that Argentina has already made the decision of surrendering its ability to pursue discretionary monetary policy by introducing a currency board. But Argentina is now in an unpleasant situation. With the currency board, the country has lost the ability to pursue an active monetary policy, but it is still unable to obtain the full benefits of that act of abnegation. Dollarization, for Argentina, is thus not a question of choosing between a rule and discretion, but rather of reaping the benefits of a choice that it has already made.

The worry that dollarization will make Argentina too vulnerable to U.S. interest rate policies is unwarranted. Whether Argentina is vulnerable to U.S. interest rates does not depend on dollarization but on how open the economy is to capital flows. In principle, Argentina could dollarize its economy at the same time as it closes its economy to capital flows, completely isolating itself from U.S. interest rates. Argentina will be vulnerable to U.S. interest rates only as long as it allows for unrestricted capital flows.

In fact, Argentina is extremely open to capital flows at present. Consequently, it is already subject to the effects of variations in international interest rates. Argentina has decided that the benefits of international capital flows more than compensate for the costs of having its interest rates tied to the international interest rate. As figures 6 and 7 show, changes in U.S. interest rates are negligible compared with the sharp increases in interest rates associated with the Tequila and Vodka–Caipirinha effects. Those sharp increases are the source of concern and are what Argentina wants to eliminate by completely dollarizing the economy.

Another concern that has been raised is that once Argentina dollarizes, the U.S. Federal Reserve will be under pressure to take into account economic conditions in Argentina when deciding its interest rate policy. But as we have mentioned, Argentina is already subject to the full consequences of the Fed's decisions, yet exerts no influence on policymaking in the U.S. We would not expect Argentina's influence to become any larger than it is at present.

A cost–benefit analysis

Assuming that dollarization would eliminate Tequila type of crises, would it be in Argentina's best interest to dollarize even if had to do it unilaterally? We estimated earlier that the annual cost in terms of seigniorage is 0.2 percent of GDP. If we think of the loss of seigniorage following dollarization in the same way as the contingent repurchase facility, namely, as an insurance premium against crises of the Tequila type, under what circumstances would 0.2 percent be an actuarially fair price? To answer this question, we have to model the risk that is being insured, however crudely.

As figure 4 shows, the Argentine economy grew at a steady 8 percent annual rate from 1990:Q1 to 1994:Q4 and then again at the same rate from 1995:Q4 to 1998:Q2. The Tequila effect appears as a permanent shock to the output level in 1995, which did not affect growth rates before or after. (Had output continued to grow without interruption, it would now be higher than it is: The loss was never made up).

Let us think of Tequila effects as follows. Every year, a Tequila shock might occur, with some probability, independently of previous occurrences. If the shock occurs, output is lower than it would have been in the absence of a shock. Afterwards, growth resumes at its normal rate, but output is permanently lower than it would have been without the shock. This model embodies what we see in figure 4, namely, that the growth rate was not permanently affected by the Tequila effect, but a sharp reduction in output occurred in 1995. Output is adversely affected through the sharp increases in interest rates shown in figure 7, due to a higher perceived devaluation risk. Dollarization would eliminate this risk, protecting the real economy from these "contagion effects."

For the Tequila effect, the permanent output loss turned out to be about 14 percent. Current forecasts for GDP growth in 2000 suggest that the impact of the Asian crisis will be the same size or greater. We do not know what the annual probability of a Tequila effect is, but we can calculate what it would have to be in order to make Argentina indifferent between dollarizing and not dollarizing. That probability is the annual cost of dollarization (0.2 percent of GDP) divided by the benefit of dollarization (14 percent of GDP), namely 1.4 percent. Given that Argentina has been hit twice in ten years, unilateral dollarization is unambiguously desirable under those assumptions. Put another way, if the annual probability of a Tequila effect is 20 percent (consistent with two occurrences

BOX 1

Rules versus discretion

One of the great lessons of the macroeconomic literature in the past 20 years has been to highlight the temptations inherent in monetary policy, which have come to be known as the time-commitment problem. Aside from raising seigniorage, the other reason for governments to resort to inflation is the Phillips curve. Originally thought of as a firm statistical law that offered a trade-off between inflation and unemployment, it is now mostly seen as a trade-off that depends on the degree to which the private sector fails to correctly anticipate the actual inflation rate (the "expectations-augmented Phillips curve"). The particular temptation that this relation induces was shown by Kydland and Prescott (1977). The following very stark presentation draws on Sargent (1999). The story has two variables, unemployment U and inflation y. It has three components: the government, the private sector's expectations, and the Phillips curve.

The government wants to minimize both unemployment and inflation. Its objective is of the form

$$1) \quad -\frac{1}{2}\left(U^2 + y^2\right).$$

Obviously, the best outcome for the government is $y = 0$ and $U = 0$. The government chooses inflation y from a set of possible values $Y = [0, \bar{y}]$. It does not choose U directly: that is determined by the Phillips curve.

The Phillips curve relates unemployment with its "natural rate" U^* and the degree to which inflation is unanticipated. Let x represent private-sector expectations of inflation:

$$2) \quad U = U^* - \theta(y - x),$$

where $\theta > 0$ is the slope of the Phillips curve; the higher the slope, the more effective unanticipated inflation is in stimulating the economy.

Finally, the private sector sets its expectations of government. We will assume rational expectations in the simplest form, that the private sector is always correct and accurately predicts y:

$$3) \quad x = y.$$

The commitment problem can be thought of as a problem of timing of moves between the government and the private sector. In one configuration, the government moves first and sets inflation before the private sector sets its expectations. The government cannot revisit its choice later on. The predicted outcome is then the solution to the government choosing y to maximize equation 1 subject to equations 2 and 3. Since equation 3 must always hold no matter what the government does, equation 2 becomes $U = U^*$: Unemployment is what it is. All that the government can do is set the inflation rate as low as possible, at $y = 0$.

In another configuration, the government moves last. The problem then becomes the solution to the government choosing y to maximize equation 1 subject to equation 3, *given* x, and, *separately*, equation 2 holding. No matter what x is, the government will want to choose a high value of y to take advantage of the Phillips curve. But the private sector, while moving first, will anticipate this action (equation 2). The result is the same unemployment U^* with high inflation.

This is, of course, a very stylized model, but it conveys the nature of the temptation inherent in the expectations-augmented Phillips curve. One way to resolve it is to somehow arrange for the first timing configuration to prevail rather than the second. But, aside from Athenian democracy and the odd Swiss canton, delegated government is a necessity, which means the government always moves last. The other way to resolve it is to accept the second timing configuration, with the government moving last, but to change the choice set of the government. Dollarization is a way to reduce Y to the single point $\{0\}$. The best outcome is then achieved.

per decade), a permanent loss of output of 1 percent would make the insurance premium actuarially fair.

Conclusion

Argentina's history has made it painfully aware of the risks involved in allowing a central bank, or government, full discretion in the setting of monetary policy. This led Argentina to establish a currency board in 1991, which is one step short of dollarization. In doing so, it has demonstrated that it is feasible for a country to relinquish control over its monetary policy. It has also shown what steps can be taken to address the loss of a lender of last resort.

Nevertheless, Argentina has suffered from several recessions that can in part be linked to speculative attacks on the currency. These attacks, in turn, were prompted by fears that Argentina's commitment to the currency board was less than full. Thus, full backing

for the currency is not enough to instill full confidence in the currency. Investors' fears are understandable, given Argentina's history.

The main argument for Argentina's dollarization above and beyond the currency board is that it would prevent or attenuate the crises that have stunted Argentina's growth in the 1990s. However, before Argentina decides to dollarize it must weigh very carefully the consequences of losing the ability of pursuing an independent monetary policy. The fact that Argentina has followed bad monetary policy in the past does not mean that it could not do much better in the future.

NOTES

[1]An optimal currency area is a geographical area that would benefit from sharing a common currency.

[2]Under a currency board, a country's currency is fully backed by foreign reserves.

[3]As of November 23, 1999, only 3 percent of reserves were actually held in that form.

[4]The projections shown in figure 3 reflect the Fiscal Convertibility Law, passed in August 1999, which requires a balanced budget by 2003.

[5]On January 12, the central bank dollarized the banks' reserves. This explains the sudden drop in the monetary base in figure 8.

[6]Figure 9 shows only the currency in circulation, since the monetary base at present consists of nothing else.

[7]In addition to the political issues involved, it is not clear what would be the best strategy for the U.S. from a revenue maximizing point of view. It is true that the U.S. could obtain some of Argentina's seignorage by joining it in a monetary union, but the U.S. could obtain all of Argentina's seignorage by letting it dollarize unilaterally. The risk of following such a strategy is that Argentina may not be willing to dollarize at all.

[8]No commitment device is absolute. Unless Argentina becomes the fifty-first state of the (North American) Union, it remains a sovereign state, and its Congress has the constitutional authority to establish and regulate a currency, just as the U.S. Congress does. However, if Argentina's reserves were put in an escrow account as collateral in the form discussed in the previous section, Argentina would have no incentives to renege the agreement.

[9]Such an argument hinges on dollarization being difficult to reverse, once accomplished. It is plausible that a government that has just defaulted on its debt would have difficulty generating much of a demand for a new currency it proposed to issue. Without a demand for money, there is no monetary base on which to collect seigniorage. It thus appears that dollarization might reduce the perception of default risk.

[10]Argentina is not the only country in which an independent monetary policy has had bad consequences for the economy. The calculations in Schmitt-Grohé and Uribe (1999) show that dollarization would cost Mexico 2 percent of consumption compared with a variety of reasonable independent policies. But, they find that the actual independent policy that Mexico followed in the past has cost the country 95 percent of its potential consumption.

REFERENCES

Burnside, Craig, Martin Eichenbaum, and Sergio Rebello, 1999, "Prospective deficits and the Asian currency crisis," Northwestern University, mimeo.

Cole, Harold L., and Timothy J. Kehoe, 1998, "Self-fulfilling debt crises," Federal Reserve Bank of Minneapolis, staff report, No. 211.

Eichengreen, Barry, 1992, *Golden Fetters*, London: Oxford University Press.

Kydland, F. E., and C. E. Prescott, 1977, "Rules rather than discretion: The inconsistency of optimal plans," *Journal of Political Economy*, Vol. 85, June, pp. 473–492.

Maddison, Angus, 1995, *Monitoring the World Economy, 1820–1992*, Paris: Organization for Economic Cooperation and Development.

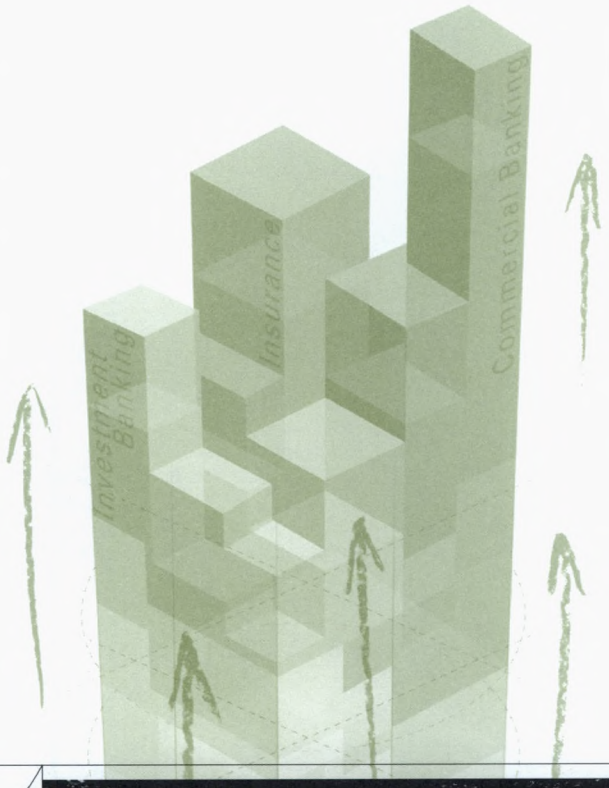
Mundell, Robert A., 1961, "A theory of optimal currency areas," *American Economic Review*, Vol. 51, September, pp. 657–665.

Rolnick, Arthur J., and Warren E. Weber, 1998, "Money, inflation, and output under fiat and commodity standards," *Journal of Political Economy*, Vol. 105, No. 6, pp. 1308–1321.

Sargent, Thomas J., 1999, *The Conquest of American Inflation*, Princeton, NJ: Princeton University Press.

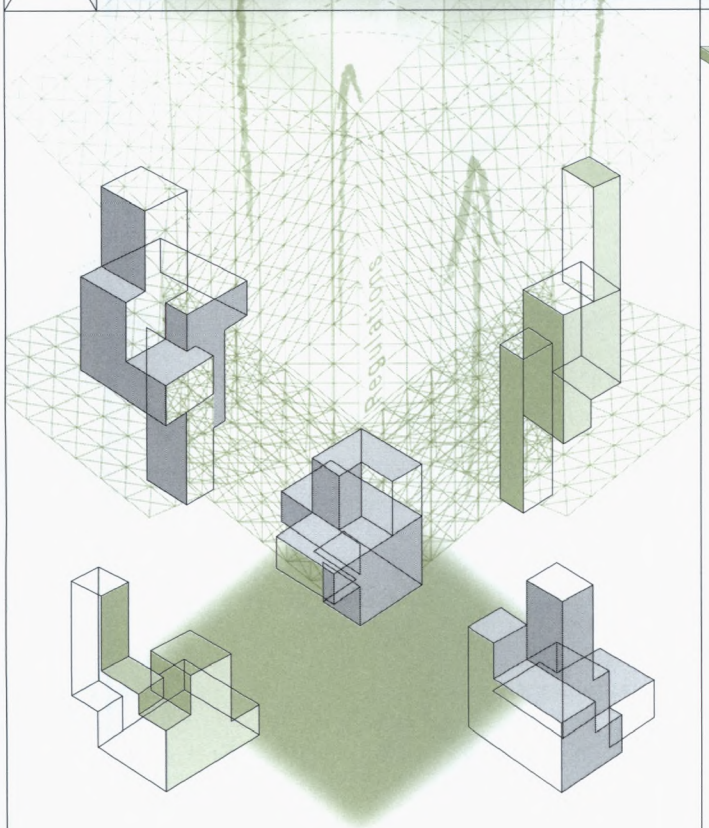
Schmitt-Grohé, Stephanie, and Martín Uribe, 1999, "Stabilization policy and the costs of dollarization," Rutgers University, mimeo.

Schwarz, Anna J., 1993, "Currency boards: Their past, present, and possible future role," *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, pp. 147–187.



On May 3–5, 2000, the Federal Reserve Bank of Chicago will hold its 36th annual Conference on Bank Structure and Competition at the Fairmont Hotel in Chicago. Since its inception, the conference has aimed to encourage an ongoing dialogue and debate on current public policy issues affecting the financial services sector. Each year the conference brings together several hundred financial institution executives, regulators, and academics.

The Changing Financial Industry Structure And Regulation:



The theme for 2000 returns us to the original roots of the conference: financial industry structure and competition. With the recent elimination of geographic barriers to bank expansion, there has been a significant increase in the number and size of bank mergers in the U.S. Similar activity has been occurring in other countries. Perhaps even more importantly, we are also seeing cross-industry mergers and affiliations—a convergence of commercial banks, investment banks, and insurance firms into modern financial service providers. With the passage of recent bank regulatory reform legislation, this trend is expected to continue in the future. While still "separate" in name, in reality, the boundaries between these firms have been significantly decreased or eliminated. What opportunities does this present to providers of financial services? What is driving the merger trend? Is the consolidation occurring for the "right" reasons? What are the implications for industry competition? For current antitrust methodologies? Are there significant efficiency gains

from universal banking or one-stop shopping? What impact might the development of Internet banking have on industry structure? On market definitions and related regulatory issues such as the CRA? With all these changes, has the role of smaller banks been enhanced or diminished?

This will be the primary theme of the 2000 conference. As in past years, much of the program will be devoted to the primary conference theme, but there will be a number of additional sessions on current industry issues. Some highlights of the conference include:

- The keynote address by Federal Reserve Board Chairman Alan Greenspan.

- A discussion of the conference theme from a variety of perspectives by a panel of industry experts. The participants scheduled to attend include James M. McCormick, President, Manhattan Consulting Group; J. Robert Kramer, U.S. Department of Justice; James Chessen, Chief Economist, American Bankers Association; Malcolm Bush, President, Woodstock Institute; and Thomas Brown, President, Second Curve Capital, Inc.

- A luncheon presentation by Donna Tanoue, who in May 1998 became the 17th chairman of the Federal Deposit Insurance Corporation.

- A panel discussion of financial safety net issues in the post-Gramm–Leach–Bliley Act regulatory environment. Participants will include J. Alfred Broaddus, President, Federal Reserve Bank of Richmond; Ernest Patrikis, Senior Vice President and General Counsel, American International Group, Inc.; and Peter Wallison, American Enterprise Institute.

- A discussion of issues arising from alternative financial delivery systems, including the role of electronic payment alternatives, Internet banking, electronic delivery of Treasury payments, and privacy and security issues. Participating panel members will include Roger W. Ferguson, Jr., Vice Chairman, Board of Governors of the Federal Reserve System; Gary Gensler, Under Secretary of the Treasury for Domestic Finance; David Medine, Associate Director for Financial Practices, U.S. Federal Trade Commission; and J. Stephen Baine, Chairman, Wingspan Investment Services, Inc.

- An update and evaluation of bank capital regulation reform as proposed by the Basle Committee on Banking Supervision.

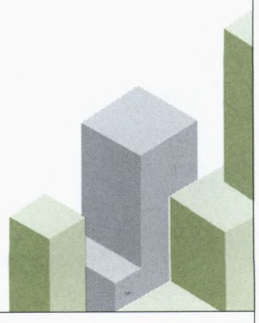
As usual, the Wednesday sessions will showcase more technical research that is of primary interest to research economists in academia and government. The Thursday and Friday sessions are designed to discuss issues that focus on the interests of a broader audience.

Bridging States, Countries, And Industries In The New Millennium

If you are not currently on our mailing list or have changed your address and would like to receive an invitation to the conference please contact:

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

Telephone: 312-322-5775
e-mail: portia.jackson@chi.frb.org



May 3–5, 2000

Black/white differences in wealth

Joseph G. Altonji, Ulrich Doraszelski, and Lewis Segal

Introduction and summary

The gap in wealth holdings between African-Americans and white Americans is enormous—much larger than the gap in earnings. For example, Menchik and Jianakoplos (1997) find that the average wealth of black households is 20 percent of the average wealth of white households in the 1976 National Longitudinal Survey of mature men and 23 percent in the 1989 Survey of Consumer Finances, even though average black income is 60 percent and 50 percent of average white income, respectively, in the two samples. Blau and Graham (1990) use data from the 1976 and 1978 waves of the National Longitudinal Surveys of young men and women and find that, on average, young black families hold only 18 percent of the wealth of young white families, while the corresponding percentage for income is 64.9.

Wealth is important in any society. It influences access to capital for new businesses, is a source of political and social influence, and provides insurance against fluctuations in labor market income. It influences the quality of housing, neighborhoods, and schools a family has access to as well as the ability to finance higher education. The fact that friendships and family ties tend to be within racial groups serves to amplify the effect of the race gap in wealth on the financial, social, and political resources available to blacks relative to whites.[1]

What explains the huge wealth gap? In this article, we summarize some of the results of our ongoing research on this question, drawing heavily on the analysis in Altonji, Doraszelski, and Segal (1999). We focus much of our attention on the most obvious possibility, which is that the wealth gap may arise because whites have higher incomes than blacks and have marriage and fertility patterns that are more favorable to wealth accumulation. Indeed, the existence of a gap in wealth is not surprising in view of the well-established income disparity.[2] Both savings levels and savings rates are positively related to income. Since blacks on average have lower incomes than whites, we would expect blacks to have lower savings. A lower flow of savings translates into less wealth. Similarly, the fact that blacks are less likely to marry, have less stable marriages, and have more children implies that blacks will have less wealth per household than whites. The issue is whether differences in income and demographic patterns can explain the large gap.

Several studies, including those mentioned above, have found large wealth differences even after controlling for differences between blacks and whites in average income and other factors. For example, Blau and Graham (1990) conclude that as little as one-quarter of the wealth gap can be attributed to racial differences in income and demographic variables. There are some limitations to previous studies that lead us to revisit the issue. For example, the wealth of a married couple is likely to depend not only on earnings last year but also on earnings in previous years. Earnings in any one year are influenced by transitory factors, such as whether an individual experiences a layoff or has opportunities to work overtime, and are a very rough indicator of the resources available to a household over the extended time frame in which wealth accumulation takes place. Smith (1995) and Avery and Rendall (1997) base their wealth

Joseph G. Altonji is a professor of economics and associate director of the Institute for Policy Research at Northwestern University, a research associate at the National Bureau of Economic Research, and a consultant to the Federal Reserve Bank of Chicago. Ulrich Doraszelski is a graduate student at Northwestern University. Lewis Segal is vice president of analytics and technology at InfoWorks Incorporated. *The authors' research has been supported by the Institute for Policy Research and the Russell Sage Foundation.*

models on current income alone, and this is not an adequate control for race-related differences in earnings streams. Consider a white family and a black family who have the same income in the previous year. In most cases, the white family will have enjoyed a higher income in other years than the black family, and thus will have higher wealth. Blau and Graham (1990) and Menchik and Jianakoplos (1997) decompose income into current income and the normal or usual flow of income to the household, which we refer to as permanent income. They measure the permanent component of income as the part of income that is predictable given race, sex, age, education, health status, number of children, and geographic location. This approach is a clear improvement over the use of only current income to measure the contribution of differences in income streams to the wealth gap. However, it is inadequate for a number of somewhat technical reasons.[3]

In this article, we take advantage of the fact that our data set, the Panel Study of Income Dynamics (PSID), contains several years of data on the income of each individual in our sample. We use an adjusted average of the income across years as our measure of permanent income in our models of wealth holding.

We also take advantage of the rich data in the PSID to do a better job of controlling for differences in household characteristics that influence wealth holding than has been possible in previous work. The earlier studies control for current demographic variables such as marital status and presence of children. Since wealth at a point in time reflects a flow of savings over many previous years, it is likely to be influenced by demographic histories as well as by current demographic variables. To address this fact, we construct measures of the marriage histories and child bearing and rearing histories of each sample member and add them to our models of wealth.

We use standard *regression techniques* to decompose the race gap in wealth holding into a part that is due to differences between whites and blacks in the income and demographic characteristics that we observe and a part that is not explained by these factors. We can explain most of the difference in wealth holding with income and demographic variables, provided that we use the wealth model that has been estimated on a sample of whites. That is, we find that blacks and whites would have similar wealth levels if 1) the relationship between wealth and income and demographics for blacks was the same as it is for whites, and 2) blacks and whites had the same distributions of income and demographic characteristics. On the other hand, when we ask, "If the relationship between wealth and income and demographics for whites were the same as it is for blacks, how much wealth would whites hold?," we conclude that whites would hold much less wealth than they actually hold. While our results are somewhat sensitive to the particular form of the regression model, they suggest that race-related differences in the sensitivity of wealth to income and demographics are a major factor in determining the race gap in wealth. We draw similar conclusions from separate analyses of home equity, stocks/mutual funds and individual retirement accounts (IRAs), and the value of farms and businesses.

So we have traded one question for another. The question becomes, "Why does the relationship between wealth and income/demographics differ so significantly between blacks and whites?" Racial differences in rates of return, inter vivos transfers and inheritances, and savings behavior could all underlie the race difference in wealth models and contribute to the part of the wealth gap not explained by income and demographics. Blau and Graham (1990) and other researchers have hypothesized that differences in inter vivos transfers and inheritances play a major part in the wealth gap. We provide some indirect evidence on the effects of transfers and gifts on the race differences in wealth models by using data on siblings to estimate the effects of income and demographics on wealth holding. Basically, we estimate the relationship between wealth and income and demographics by regressing differences among siblings in wealth on differences among siblings in income and household characteristics. We do this using a statistical technique called *fixed effects* regression. Using differences among siblings to estimate wealth models largely neutralizes the effects of differences between whites and blacks in inter vivos transfers and inheritances from parents. This is because parental gifts and bequests do not differ greatly among siblings. Consequently, the analysis of siblings provides a way of controlling for the effects of adverse history on the relative position of blacks. Our results for siblings, while somewhat imprecise, confirm our basic finding that wealth holdings are much less strongly related to income and demographic variables for blacks than whites. They tentatively suggest that the race difference in the wealth models is not driven primarily by inter vivos gifts and inheritances.[4]

Data

The data source we use for our study is the PSID, which is collected by the University of Michigan, Institute for Social Research. The PSID is based on a random sample of U.S. households in 1968 and a

separate low-income sample. The households were interviewed annually through 1997, providing many years of income data and long demographic histories for the panel members. Detailed wealth data were collected in 1984, 1989, and 1994, and form the heart of our analysis. We use both the random and the low-income samples without weighting to estimate the wealth models. However, we use survey weights to make our estimates nationally representative when computing decompositions of the wealth gap and descriptive statistics.

The PSID contains a full set of variables only for household heads ("heads") and their spouses ("wives"). Our analysis is based on all persons who were either a head or a wife in at least one of the three years for which wealth data are available. Household heads include the male in a married couple as well as the male or female heads from single adult households. We also created a demographic history for each individual that describes past and present marriages and child bearing and rearing. We use wealth including home equity as the measure of wealth. We also analyze main home equity (house value net of mortgage balance), stocks/mutual funds and IRAs, and wealth in farms/businesses.[5]

Real nonasset family income (deflated by the Consumer Price Index for urban areas) is our measure of current income.[6] We take advantage of the panel nature of the PSID by using all of the available data for an individual when estimating permanent income. Our measure of permanent income is basically a time average of past, current, and future income for each person adjusted for age, marital status, presence of children, and time. The averaging reduces the influence of measurement error and transitory variation in income. See box 1 for details. Our measure has sufficient variation over individuals to permit us to work with nonlinear models of the relationship between income and wealth, in contrast to the linear specifications used in previous studies. In Altonji et al. (1999), we show that the use of linear specifications may lead one to understate the importance of the race gap in income as a source of the wealth gap.

Descriptive statistics

It is useful to begin with a brief discussion of race differences in wealth as well as some of the key income and demographic variables. To save space, we present statistics for the pooled sample of observations for 1984, 1989, and 1994. The weights are normalized so that the means are estimates of the average of the population means across the three years. We provide variable definitions and descriptive statistics for the key variables used in our analysis in tables 1 and 2.

Table 1 provides descriptive statistics for wealth, current income, and permanent income in 1989 dollars. There are separate columns for white couples, black couples, single white males, single black males, single white females, and single black females. A person may be in multiple samples if their marital status changes over time. In the case of couples, the mean of wealth is $54,357 for blacks and $206,386 for whites, a ratio of 0.26. The race gap for income is much smaller, with a mean of $30,236 for blacks and $41,471 for whites, a ratio of 0.73. This is reflected in our permanent income measures, which have a mean of $31,717 for black household heads and

BOX 1

Permanent income

To construct our measure of permanent income, we make use of the panel nature of the PSID. The measure is based on the regression model

$$Y_{it} = X_{it}\gamma + e_{it},$$

where Y_{it} is nonasset family income of person i in year t, and the vector X_{it} consists of a fourth-order polynomial in age (centered at 40), a marital status dummy, an indicator for children, the number of children, and a set of year dummies. In turn, e_{it} is the sum of an individual-specific effect and an idiosyncratic error term, $e_{it} = v_i + u_{it}$. We estimate the parameters of the above equation from race- and gender-specific regressions using all observations in which the person was either a head or wife. Our measure of permanent income is the individual-specific effect v_i, estimated as the person-specific mean of the residuals from the regression. We construct separate measures for the level and the log of permanent income and normalize them to refer to the year of the wealth survey. To ensure the quality of our permanent income measures, we dropped persons with less than four observations in the subsequent analysis. Note that the permanent income variables are normalized to refer to the flow of income at a specific age, 40, for a person in a household with a particular set of characteristics. Consequently, the mean of permanent income may differ substantially from the mean of current income.

TABLE 1

Descriptive statistics for wealth and income variables

	White couples	Black couples	Single white males	Single black males	Single white females	Single black females
Wealth, including main home equity	$206,386.54 (600,271.64)	$54,357.54 (140,186.57)	$63,085.88 (154,851.15)	$15,492.33 (35,991.84)	$71,032.20 (208,611.65)	$15,228.90 (407,11.16)
Main home equity	58,207.66 (74,142.09)	25,694.76 (52,626.45)	18,537.59 (42,916.79)	6,701.92 (16,246.51)	27,739.75 (114,047.48)	9,213.12 (25,739.83)
Farm/business	37,317.57 (253,246.67)	2,695.27 (43,804.92)	8,261.79 (49,746.52)	8,19.13 (22,559.72)	4,406.91 (32,473.38)	104.21 (2,602.34)
Stocks/mutual funds/IRAs	27,745.31 (189,505.85)	4,109.96 (22,228.30)	8,704.90 (63,122.62)	585.52 (3,663.64)	9,105.07 (40,662.64)	516.71 (4,185.33)
Main home equity, excluding 0	69,727.91 (76,037.64)	38,390.03 (60,419.77)	50,076.88 (58,276.33)	30,930.56 (21,651.04)	57,327.17 (158,694.18)	30,740.53 (39,355.64)
Farm/business, excluding 0	182,532.08 (535,903.74)	60,064.02 (198,283.17)	72,540.20 (130,635.90)	38,866.08 (150,564.01)	80,798.82 (114,724.86)	25,752.15 (31,827.34)
Stocks/mutual funds/IRAs, excluding 0	69,013.89 (294,076.23)	26,963.97 (51,238.27)	34,181.45 (121,552.20)	8,724.18 (11,357.29)	38,503.63 (76,551.80)	8,949.50 (15,097.33)
Total taxable nonasset income	41,471.30 (43,812.12)	30,236.65 (19,933.27)	22,446.70 (19,231.00)	14,077.71 (11,878.50)	14,636.77 (26,904.87)	10,622.19 (8,854.27)
Permanent income	45,680.89 (20,976.23)	31,717.42 (12,590.99)	44,098.38 (14,577.26)	27,463.24 (8,161.70)	38,241.75 (12,161.26)	25,103.79 (7,033.50)
Spouse permanent income	42,343.33 (22,991.73)	29,553.00 (13,840.07)				
Number of observations	7,600	2,509	1,395	1,133	2,705	3,179

Notes: Computed from the pooled sample using weights. Standard deviations in parentheses. The weights are normalized so that for each subgroup the means are estimates of the average of the population means for 1984, 1989, and 1994. The definition of permanent income is given in the text.

Source: Authors' calculations based on data from the PSID.

$45,680 for white heads. The permanent income values are $29,553 and $42,343 for black wives and white wives, respectively. The black/white ratios of permanent income are about 0.70. Moreover, the distributions for current and permanent income are much more concentrated than the distributions for wealth. We also report descriptive statistics on several key components of wealth, including home equity, the value of a farm or business, and the value of stocks, mutual funds, and IRAs. For each component we report the mean and standard deviation for the households that have nonzero values, as well as the overall mean and standard deviation including the zero values. It is interesting to note that the race gap in home equity is smaller than the gap in total wealth. With zero values included, the mean of home equity is $25,694 for black couples, which is 44 percent of the value of $58,207 for white couples. In contrast, black couples hold only $4,110 in stocks, mutual funds, and IRAs, which is only 15 percent of the corresponding mean value for whites. Only 15 percent of black households hold wealth in this category, while 40 percent of white households do.

The black self-employment rate is only about one-third of the white self-employment rate, and this ratio has been relatively constant for the past 70 years (see Fairlie and Meyer, 1997). Given this fact, one would expect the value of farms and businesses to be much smaller for blacks than whites. The data confirm this. Only 4 percent of black couples report having assets in a farm or business, while 20 percent of white couples report such assets. Including the zero values, black couples hold an average of $2,695 in farms or businesses, which is only 7 percent of the mean for whites.

The situation for singles mirrors the one for couples. In the case of single women, the mean of wealth is $15,228 for blacks and $71,032 for whites, a ratio of 0.22. The race gap for income is again much smaller, with a mean of $10,622 for blacks and $14,637 for whites, a ratio of 0.73. The means of permanent income of individuals who are single heads of households in 1984, 1989, or 1994 exceed the means of current income dramatically. The numbers are $38,242 for whites and $25,103 for blacks.

In table 2 we present the definitions and descriptive statistics of regional and demographic variables that influence wealth. Many of these show substantial differences across races. Since housing prices vary across regions, and a much higher proportion of blacks live in the South, we control for region and residence in a standard metropolitan statistical area

(SMSA) in our analysis. In the case of couples, the number of children currently living in the family unit is higher for blacks (1.20) than for whites (0.92), although the number of dependents is similar (0.25 for whites and 0.29 for blacks). The difference in the total number of own or adopted children is even bigger, with 2.42 (2.41) for white husbands (wives) and 2.88 (2.85) for black husbands (wives). This points to the potential importance of controlling not only for current demographics but also for demographic histories.

Blacks describe themselves as being in poor or fair health more often than whites.[7] Whites are better educated than blacks, with almost three times as many whites holding a college degree and two times as many whites holding advanced or professional degrees.

TABLE 2

Descriptive statistics for demographic variables

	White couples	Black couples	Single white males	Single black males	Single white females	Single black females
Northeast region	0.23	0.15	0.22	0.15	0.22	0.15
Midwest region	0.29	0.18	0.29	0.22	0.28	0.23
South region	0.30	0.60	0.28	0.55	0.29	0.53
West region	0.18	0.07	0.21	0.08	0.21	0.09
SMSA	0.48	0.64	0.55	0.72	0.56	0.71
Spouse annual hours worked	936.18 (933.12)	979.05 (955.71)				
Age	48.06 (15.18)	47.83 (15.45)	41.08 (17.66)	38.60 (14.69)	53.23 (20.61)	43.90 (17.30)
Spouse age	45.46 (14.71)	44.65 (14.55)				
Number of children in family	0.92 (1.13)	1.20 (1.33)	0.11 (0.46)	0.17 (0.56)	0.36 (0.81)	1.04 (1.28)
Children in family	0.48	0.58	0.07	0.11	0.21	0.53
Number of dependents outside family	0.25 (0.77)	0.29 (0.93)	0.36 (0.81)	0.56 (1.02)	0.16 (0.63)	0.10 (0.42)
Dependents outside family	0.14	0.15	0.21	0.30	0.09	0.07
Health fair or poor	0.14	0.30	0.16	0.25	0.25	0.31
Spouse health fair or poor	0.12	0.29				
Schooling						
0–8	0.08	0.18	0.07	0.12	0.13	0.15
9–11	0.13	0.19	0.14	0.21	0.16	0.28
12–15	0.53	0.52	0.54	0.59	0.55	0.50
16	0.17	0.06	0.17	0.08	0.11	0.05
17+	0.10	0.05	0.08	0.01	0.06	0.02
Spouse schooling						
0–8	0.05	0.09				
9–11	0.12	0.24				
12–15	0.65	0.57				
16	0.13	0.05				
17+	0.06	0.05				
Number of marriages	1.18 (0.49)	1.12 (0.43)	0.64 (0.72)	0.54 (0.67)	0.97 (0.75)	0.71 (0.65)]
Tenure of current marriage	21.83 (15.34)	20.67 (14.85)	1.81 (7.50)	2.43 (6.85)	2.01 (7.76)	3.05 (8.98)
Spouse number of marriages	1.18 (0.47)	1.11 (0.45)				
Number of children born or adopted	2.42 (1.66)	2.88 (2.31)	1.10 (1.65)	1.75 (2.11)	1.99 (1.98)	2.61 (2.37)
Spouse number of children	2.41 (1.65)	2.85 (2.41)				
Number of observations	7,600	2,509	1,395	1,133	2,705	3,179

Notes: Computed from the pooled sample using weights (see table 1). Standard deviations in parentheses.
SMSA refers to standard metropolitan statistical area. Schooling refers to highest level of education.
Source: Authors' calculations based on data from the PSID.

Finally, whites have a slightly higher tendency to marry, as reflected in the number of marriages and the tenure of the current marriage.

Econometric models and methods

Let i index individuals or couples and j index blacks and whites, where j is b for blacks and w for whites. Let W_i^j denote a measure of wealth, Y_i^j a vector of income variables, and X_i^j a vector of demographic variables.

Our basic model specifies wealth to be linear in the income and demographic variables and is given by

$$W_i^w = \alpha_0^w + Y_i^w \alpha^w + X_i^w \beta^w + \varepsilon_i^w$$

$$W_i^b = \alpha_0^b + Y_i^b \alpha^b + X_i^b \beta^b + \varepsilon_i^b,$$

where α_0^w, α^w, and β^w are the regression intercept and slope parameters for whites, ε_i^w is the error term, and α_0^b, α^b, β^b, and ε_i^b are the corresponding parameters and error term for blacks. Separate sets of regressions are specified for single males, single females, and married couples, so the slopes and intercepts depend on sex and marital status as well as on race. The observations are pooled across time with year indicator variables to control for differences over time.

We use the regressions to decompose the difference in wealth between whites and blacks into two parts. The first part is due to the difference between whites and blacks in the average values of income and demographic variables, and the second part is due to a racial difference in the parameters of the wealth model. We refer to the first part as the "explained" gap, meaning "explained by the income and demographics" and to the second as the "unexplained gap." Such decompositions are standard in the literature on group differences, including the studies of the race gap in wealth cited in the introduction. We perform two different decompositions. One uses the parameters of the regression model for whites, α_0^w, α^w, and β^w, to measure the contribution to the wealth gap of the differences between whites and blacks in income Y_i and demographic characteristics X_i. The second decomposition uses the parameters of the regression model for blacks, α_0^b, α^b, and β^b, to compute how much the race differences in income and demographics matter for wealth holding. See box 2 for details.

Basic results

Here, we present decompositions of the race gap into a component explained by differences in income and demographic variables and an unexplained component measuring the portion of the gap that remains after conditioning on income and demographics. First, we discuss our findings for married couples.

Couples

Before turning to the wealth decompositions, we must describe the specification of the regression model that is used to produce them. The dependent variable is the level of wealth. In the case of couples, the controls for income and earnings capacity Y_i are current family income, permanent family income of the husband, and permanent family income of the wife. We also include the squares of current income, head's permanent income, wife's permanent income, and the products of current income with the head's and the wife's permanent income. The vector of geographic and demographic controls, X_i, contains region dummies, a dummy for residence in an SMSA, four education dummies for the husband and four for the wife, the wife's work hours in the previous year,[8] a dummy equal to 1 if the wife's health is fair or poor, and a dummy equal to 1 if the husband's health is fair or poor. It also includes fourth-order polynomials in the age of the husband and the age of the wife (centered at age 40), a dummy equal to 1 if there are children under 18 in the family unit and 0 otherwise, the number of children under 18 in the household, controls for whether the household head has dependents outside of the family unit, the number of dependents outside the family unit, controls for the number of marriages of the head, the respective number for the wife, the tenure of the current marriage, the total number of children of the head, and the total number of children of the wife. Finally, we include year dummies for the 1984 and the 1994 surveys. Estimates of the regression models are reported in Altonji et al. (1999). Our focus here is on the wealth decompositions based on the regression equations listed above rather than on the coefficients of specific variables.

The estimate of the wealth gap is $150,656 with a standard error of $13,872 (table 3, column 5).[9] We emphasize that the group means have substantial standard errors, which is a reflection of the extreme values in the wealth distribution and is not always appreciated in the literature making group comparisons. Using the estimates of α_0^w, α^w, and β^w to assess the importance of the white/black difference in the explanatory variables, we conclude that the race gap in income and demographics explains $101,391, or 67 percent, of the gap for couples (table 3, column 6). We obtain strikingly different results when we use the estimates of α_0^b, α^b, and β^b from the wealth equation for blacks to evaluate the wealth gap. Using these coefficients we find that only 6 percent of the wealth gap is explained by differences in the explanatory variables

BOX 2

Regression decomposition

We evaluate the explanatory power of our wealth models using a regression decomposition. The predicted values of the mean of wealth for whites and blacks are, respectively,

$$\hat{W}^w = \alpha_0^w + \bar{Y}^w\alpha^w + \bar{X}^w\beta^w$$
$$\hat{W}^b = \alpha_0^b + \bar{Y}^b\alpha^b + \bar{X}^b\beta^b,$$

where $\bar{Y}^w$ and $\bar{X}^w$ are weighted means of the income and demographic variables for the sample of whites and $\bar{Y}^b$ and $\bar{X}^b$ are the weighted means for blacks. (Note that we estimate the wealth models without using sampling weights to avoid introducing additional heteroscedasticity into the analysis but weight the observations when performing decompositions so that they will be representative of the U.S. population. As a result, $\hat{W}^w$ and $\hat{W}^b$ differ somewhat from the weighted sample means $\bar{W}^w$ and $\bar{W}^b$ of wealth for whites and blacks, respectively.)

In addition to predicting wealth holdings, we can use our wealth models to ask a counterfactual question, namely "How much wealth would blacks hold if they had the same relationship between income and demographic variables and wealth holdings as whites?" The answer is given by

$$\alpha_0^w + \bar{Y}^b\alpha^w + \bar{X}^b\beta^w.$$

Using the above equations, we can decompose the wealth gap into a part that is explained and a part that remains unexplained. Specifically, it is easy to show that

$$\hat{W}^w - \hat{W}^b = \left\{\left(\bar{Y}^w - \bar{Y}^b\right)\alpha^w + \left(\bar{X}^w - \bar{X}^b\right)\beta^w\right\} + \left\{\alpha_0^w - \alpha_0^b + \bar{Y}^b\left(\alpha^w - \alpha^b\right) + \bar{X}^b\left(\beta^w - \beta^b\right)\right\}.$$

The first term in brackets is the part of the total wealth gap $\hat{W}^w - \hat{W}^b$ between whites and blacks that is explained by racial differences in the mean of income $\bar{Y}^w - \bar{Y}^b$, and the means of the demographic variables $\bar{X}^w - \bar{X}^b$ based on the coefficient estimates from the white sample. In other words, it is an estimate of the contribution of income and demographic differences to the wealth gap, assuming the dependence of wealth on income and demographics for blacks is the same as it is for whites. The second term represents the "unexplained" part of the wealth gap—the difference that arises because the relationship between characteristics and wealth, as summarized by the regression parameters, differs between whites and blacks.

The gap in wealth may also be decomposed using regression coefficients α^b and β^b for blacks to assess the contribution to the wealth gap of the race differences in income and demographics. This wealth decomposition is given by

$$\hat{W}^w - \hat{W}^b = \left\{\left(\bar{Y}^w - \bar{Y}^b\right)\alpha^b + \left(\bar{X}^w - \bar{X}^b\right)\beta^b\right\} + \left\{\alpha_0^w - \alpha_0^b + \bar{Y}^w\left(\alpha^w - \alpha^b\right) + \bar{X}^w\left(\beta^w - \beta^b\right)\right\}.$$

The first term is the portion of the wealth gap explained by income and demographic variables based on the wealth model for blacks. The second term is the unexplained portion.

As we shall see, the coefficients of the wealth model tend to be much larger in absolute value for whites than for blacks. This fact drives a central finding of our study—a much larger portion of the race gap in wealth can be attributed to white advantages in income and demographic characteristics when these differences are evaluated using the wealth model for whites rather than the wealth model for blacks.

(table 3, column 7). This large discrepancy between the white and the black wealth models in the degree to which racial differences in the distributions of the income/human capital and demographic variables can explain the gap in wealth levels is a key theme in our analysis. Underlying this result is the fact that wealth differences among blacks are much less sensitive to differences in income and demographics than wealth differences among whites. Blau and Graham (1990) obtain qualitatively similar results using the National Longitudinal Survey of Young Men and Young Women.

To illustrate the role of the difference in the white and black wealth equations, we compute the index $Y_i\alpha^j$, corresponding to the vector of income variables Y_i for each observation in the combined sample of blacks and whites. We then regress the value of the index $Y_i\alpha^b$ in the combined sample on $Y_i\alpha^w$ and a constant. We would expect a coefficient of 1 on $Y_i\alpha^w$ if the coefficient vectors α^b and α^w are identical. We would expect a coefficient below 1 if the elements of α^b are of the same sign but smaller in absolute value than the corresponding elements of α^w.[10]

The regression of $Y_i\alpha^b$ on $Y_i\alpha^w$ and a constant is 0.1485 (with a standard error of 0.0038). A similar regression involving the $X_i\beta^w$ and $X_i\beta^b$ indexes also shows that differences in demographics have a stronger association with wealth levels for whites than for blacks. The slope coefficient of the regression of $X_i\beta^b$ on $X_i\beta^w$ and a constant is 0.1570 (with a standard error of 0.0024). We obtain qualitatively similar results for single women and single men.

Single women

Table 3 reports wealth decompositions for single women. The specification of the regression model that we use for single women and single men corresponds to the model for couples, with all variables pertaining to a spouse excluded. For single women the estimate of the mean wealth gap is $57,026 (with a standard error of $6,730). Using wealth regression coefficients for the white sample, we find that single black women would have 90 percent of the wealth that white women hold if they had the same income and demographics as whites. This suggests that the large wealth gap is for the most part a reflection of racial differences in income streams, human capital variables, and current and past demographic variables. However, the wealth model for blacks tells a completely different story. Using the wealth coefficients for black single women, we find that only $15,931 or 28 percent of the total gap is attributable to income and demographics.

Single men

Table 3 also reports results for single males. The results parallel those for single females and couples. Using the estimates of the regression model for the white sample, we find that single black men would have 108 percent of the wealth of single white men if they had the same income and demographics as whites. This result, like the result for single females, suggests that the large wealth gap is simply a reflection of racial differences in income streams, human capital variables, and current and past demographic variables. However, we again find that the estimated coefficients of the wealth model for blacks tell a completely different story. Only $13,262 or 27 percent of the total gap of $49,731 ($5,810) is attributable to income and demographics.

Summary

We find that most or all of the race gap in the wealth level for single men and single women and a substantial portion of the gap for married couples would disappear if blacks and whites had the same distribution of income and demographic variables *and* if the slope coefficients of the white wealth equation also held for blacks. However, the wealth models for blacks exhibit much less sensitivity to income and demographics, indicating that both the race gap in the income and demographics *and* race differences in the distribution of wealth *conditional* on income and demographic variables play important roles in the gap in wealth levels.

TABLE 3

Regression decompositions of race gap, level of wealth

Demographic group	White coefficients		Black coefficients		Total gap	Explained gap, white coefficients	Explained gap, black coefficients
	White characteristics	Black characteristics	Black characteristics	White characteristics			
Couples	$203,869	$102,478	$53,213	$62,433	$150,656	$101,391	$9,220
	(7,906)	(20,061)	(11,399)	(13,813)	(13,872)	(67%)	(6%)
Males	64,277	10,326	14,546	27,808	49,731	53,951	13,262
	(5,157)	(11,126)	(2,677)	(3,781)	(5,810)	(108%)	(27%)
Females	70,967	19,864	13,941	29,872	57,026	51,103	15,931
	(6,368)	(12,596)	(2,176)	(5,119)	(6,730)	(90%)	(28%)

Notes: Computed from pooled sample using weights (see table 1). Standard errors in parentheses, columns 1–5. The regression coefficient estimates are estimated without sample weights. The dependent variable in the regressions is the level of wealth in 1989 dollars. Columns 1 and 2 are based on coefficient estimates from the white sample; columns 3 and 4 are based on coefficient estimates from the black sample. The variables included in the model are discussed in the text. Column 1 predicts wealth holdings for whites and column 3 for blacks. Column 2 uses the white coefficient estimates with the black sample to calculate counterfactual wealth holdings for blacks; column 4 uses the black coefficient estimates with the white sample to calculate counterfactual wealth holdings for whites. Column 5 is the difference between columns 1 and 3; column 6 is the difference between columns 1 and 2; and column 7 is the difference between columns 4 and 3. The percentage gap explained is in parentheses in columns 6 and 7. It is 100 times column 6 (and column 7) divided by column 5.

Source: Authors' calculations based on data from the PSID.

Alternative models

In Altonji et al. (1999) we work with a number of alternative econometric models of the wealth gap. While the full analysis is beyond the scope of this article, we provide a summary of what we have learned using the alternatives.

Models with interaction terms

One disadvantage of the linear models of wealth we estimate is that they implicitly restrict the interaction between the effects of income in the effects of demographic variables such as children, age, and marriage history. In Altonji et al. (1999) we experiment with including interactions between the demographic variables in the income terms. One can go only so far in this direction, because the sample sizes are not large enough to allow a rich set of interactions. For the most part, the results are quite consistent with the ones we report here. However, the explained portion of the wealth gap using the black coefficient estimates rises somewhat relative to the results reported in table 3.

Results for the log of wealth

Another standard way to allow for interactions among the explanatory variables is to use the log of wealth as the dependent variable rather than wealth itself. A second reason to use the log of wealth is that the wealth distribution is highly skewed, with a small number of individuals accounting for a very large fraction of total wealth. The use of the log of wealth reduces the impact of outliers. On the other hand, there are a substantial number of people who hold zero or negative wealth. If the value of wealth is less than $1,000, we set the log of wealth to the log of $1,000. (The results are not very sensitive to this threshold.)

In Altonji et al. (1999) we present results for the log of wealth. In the log wealth regressions, we use the log of permanent income and current income as our income measures. We find that the fraction of the gap in the mean of log wealth that is explained by income and demographic variables is large when we use the white regression coefficients to weight the differences in the variables, but smaller than the corresponding estimates when we analyze wealth itself. Second, the fraction of the gap in log wealth explained using the black regression model is substantially larger than the fraction of the gap in wealth itself. For example, for couples, the explained fraction of the wealth gap using the black regression model is 58 percent in the case of log wealth and only 6 percent in the case of the level of wealth. Again, we find that the responsiveness of log wealth to income and demographics is larger for whites than blacks, although the difference is not nearly as large as in the case of wealth itself.

Overall, the log results suggest that income and demographics play a major role in the wealth gap between whites and blacks, but the results based upon the white model are less dramatic than the results for wealth itself, and the portion of the gap explained using the black model is much larger for log wealth than wealth. However, for two reasons we are not sure how much weight to place on these results. First, a large number of observations are affected by the lower bound on wealth. Second, the translation between the log of wealth and wealth itself is not straightforward. In particular, since the log wealth model implies a multiplicative model of wealth, the race difference in the intercepts of the log regression model translates into a smaller response of wealth to income variables and demographic variables for blacks. This parallels our findings using wealth as the dependent variable.

Median regression results

Median regression is a statistical technique to predict the median value of the dependent variable conditional on a set of regressors; in contrast, ordinary least squares regression predicts the conditional mean of the dependent variable. One might prefer to predict median wealth conditional on the income and demographic variables rather than use the more standard mean regression based upon ordinary least squares for two reasons. First, in view of the skewness of the wealth distribution, it may be easier to estimate conditional medians because median regression is more robust to outliers. Second, one may be more interested in the wealth of the "typical," hence median, person with a given set of characteristics than in the mean of wealth for such people.[11] In the case of couples, using the set of explanatory variables that we use for the standard regression models, the total gap in the median is estimated at $85,935, which compares to a gap in the mean of wealth of $150,656. The income and demographic variables account for 62 percent of the wealth gap if we use the median regression function for whites. In contrast, the median wealth regression for blacks implies that demographic characteristics account for only 23 percent of the gap in the conditional median of wealth. For single men, the white median regression implies that 86 percent of the gap is explained, while the black regression implies only 43 percent is explained. The corresponding figures for females are 68 percent and 30 percent, respectively. Overall, income and demographics account for a somewhat smaller percentage of the race

gap in the conditional median of the wealth level than in the conditional mean, particularly in the case of single men and single women. When we use the wealth model for blacks to measure the contribution of observables to the race gap, the results are similar to our findings for the gap in the conditional mean. The black coefficient estimates imply a much smaller role for income and demographics.

Results for subcomponents of wealth

A recurring theme throughout our analysis is the black/white difference in the relationship between wealth and income/demographics. Why are the coefficients so different across the black and white specifications? One hypothesis we can explore is whether the relationship differs for some forms of wealth assets but not for others. Table 4 analyzes main home equity (house value net of mortgage balance), stocks/mutual funds/IRAs, and wealth in farms/businesses for the sample of couples, applying the ordinary least squares regression decomposition technique to the level of each asset. The regressors are the same as those we use for couples in the other models. On average, whites hold $33,079 more home equity than blacks ($57,911 versus $24,832), which is not surprising given the difference in home ownership rates across races. The white regression model explains 78 percent of this gap, compared with the 67 percent explained for the combined wealth assets in table 3. Again, the amount explained using the model for blacks is significantly smaller, only 30 percent.

The unconditional black/white difference in asset holdings is significantly larger for the other two asset categories (stocks and business wealth) and the wealth model for whites explains less of the difference than the overall model (only 61 and 47 percent, respectively). However, the fractions explained by the black regression equation are negligible at only 17 and 3 percent, respectively.

Results for siblings: Indirect evidence on the role of inheritances and parental transfers

Our results for total wealth as well as the individual components of wealth show substantial differences in the sensitivity of wealth holding to income and demographic variables. Because both the income and demographic characteristics of whites are more favorable for wealth holding, we assign higher fractions of the wealth gap to differences in income and demographics when we use the white wealth equations than when we use the black wealth equations. There are at least three possible explanations why wealth holding may be more sensitive to characteristics for whites than for blacks. First, whites may enjoy a higher rate of return on assets, in which case the same level of savings and inter-family transfers would lead to larger wealth levels, magnifying underlying differences that are associated with income and demographics. Second, inheritances and inter vivos transfers are larger among whites than among blacks because the long history of discrimination against blacks has inhibited the accumulation of wealth in the black population.[12] Third, the savings rates of blacks may be less sensitive to the income and demographic variables for reasons that are not clear. A reduced sensitivity of saving rates to income and demographics would lead to a reduced sensitivity of wealth to these variables.

TABLE 4

Regression decompositions of race gap, level of wealth components (couples sample)

Demographic group	White coefficients: White characteristics	White coefficients: Black characteristics	Black coefficients: Black characteristics	Black coefficients: White characteristics	Total gap	Explained gap, white coefficients	Explained gap, black coefficients
Home equity	$57,911	$32,223	$24,832	$34,719	33,079	25,688	9,887
	(978)	(2,482)	(1,336)	(1,620)	(1,656)	(78%)	(30%)
Farm/business	35,844	20,620	3,506	4,385	32,338	15,224	879
	(3,771)	(9,570)	(4,501)	(5,454)	(5,872)	(47%)	(3%)
Stocks/mutual funds/IRAs	27,626	12,495	2,735	6,929	24,891	15,131	4,194
	(2,266)	(5,750)	(583)	(706)	(2,340)	(61%)	(17%)

Notes: Computed from pooled sample of couples using weights (see table 1). Standard errors in parentheses, columns 1–5. The regression coefficient estimates are estimated without sample weights. The dependent variable is the level of the indicated wealth component in 1989 dollars. The definitions of the wealth components are given in table 1. Households with 0 values for a particular component are included in all computations. See table 3 for an explanation of columns.

Source: Authors' calculations based on data from the PSID.

Menchik and Jianakoplos (1997) provide some evidence that blacks experience a lower rate of return on assets. However, the evidence on this point is far from conclusive. In Altonji et al. (1999), we use data on siblings to explore the possibility that differences in intergenerational transfers are the source of differences in wealth holding. The PSID tracks all members of the families sampled in 1968, providing information on siblings after they form their own households. We analyze the data on siblings using fixed effects regression. It amounts to relating differences among siblings in wealth to differences among siblings in the income and demographic variables that determine wealth. These differences should not be affected by parental transfers or expected future transfers that are common to siblings. Consequently, our use of data on siblings largely neutralizes the effects of differences between whites and blacks in inter vivos transfers and inheritances, and provides a way of controlling for the effects of parental resources in our analysis of the link between wealth and income and demographics. Hence, if the analyses based on standard regression and fixed effects regression give similar answers, then we can conclude that race differences in gifts and inheritances that are correlated with income and demographic variables do not explain our finding that wealth levels are more sensitive to income and demographic variables in the case of whites than in the case of blacks. Our approach does not require us to observe the actual transfers; instead it assumes that siblings act under the belief that they will receive similar inheritances. This assumption is consistent with the empirical evidence that inheritances are evenly divided in about 70 percent of the cases and that sibling differences in inter vivos transfers from parents have only a modest relationship to sibling differences in income.[13]

To obtain adequate sample sizes we pool observations on single men, single women, and couples and add control variables for the three demographic groups. The results are presented in detail in Altonji et al. (1999); here, we provide a brief summary. Basically, we find that income and demographic differences over-explain the wealth gap and account for 111 percent of the wealth gap between the samples of black siblings and white siblings when we use the coefficients from the wealth model for whites and only 30 percent of the gap when we use the wealth model for blacks. The decompositions based on the application of fixed effects regression techniques to the sibling samples are similar to what we obtain when we apply standard regression to the sibling samples and to the results reported in table 3. We continue to explain more of the wealth gap using the white coefficients than the black coefficients, particularly when we specify wealth and income in levels. There is little indication that differences in factors such as inheritances or inter vivos transfers that are likely to vary across families provide an explanation for the racial difference in the sensitivity of wealth to income and demographics. However, the standard error of the difference in the percentage of the wealth gap explained by the white and black models is approximately 31.8. Consequently, the results using fixed effects regression are not sufficiently precise to rule out the possibility that inheritances and family transfers partially explain the stronger relationship for whites between income and demographics and wealth. Furthermore, Altonji et al. (1999) point out that there may be an interaction between parental transfers and income and demographics that is missed in our sibling analysis. Nevertheless, our results to date tentatively suggest that differences in savings behavior and/or rates of return may be more important than intergenerational transfers in explaining the very different wealth models that we obtain for whites and blacks.[14]

Conclusion

We use improved income and demographic measures and unique data on siblings to assess the role of differences in income and demographic characteristics, such as marriage patterns and fertility, in the huge disparity in wealth between whites and blacks. When we use the level of wealth as the dependent variable, we can explain a large part of the difference in wealth holdings with income and demographic variables, provided that we estimate the wealth model on a sample of whites. That is, we find that blacks would have wealth levels similar to whites, particularly for single males and females, if 1) the relationship between wealth and income and demographics for blacks was the same as it is for whites, and 2) blacks and whites had the same income and demographic characteristics. On the other hand, we can explain only a small fraction of the race gap when we ask the question: "If the relationship between wealth and income and demographics for whites was the same as it is for blacks, how much wealth would whites hold?" In general, the regression coefficients relating income and demographic characteristics to wealth are much smaller for blacks. The smaller coefficients mean that less of the race gap in wealth is explained by the gap in income and demographics.

Our results are robust to a number of experiments regarding estimation methodology and functional form, which are discussed in more detail in Altonji

et al. (1999). However, they are much less dramatic when we use the log of wealth as the dependent variable, and further research on alternative functional forms is high on our research priorities. But our results suggest that the race gap resulting from the sensitivity of wealth to income and demographics is as important as the race gap in actual income and demographics.

Given the substantial differences in the sensitivity of wealth holding to income and demographics, it becomes important to determine the degree to which race differences in inter vivos transfers and inheritances, savings rates, and rates of return on savings contribute to the unexplained part of the wealth gap. We attempt to isolate the role of differences in transfers and inheritances by analyzing wealth differences among siblings. The fact that we obtain similar results when we relate sibling differences in wealth to sibling differences in income and demographics tentatively suggests that much of the difference between whites and blacks in the effect of income and demographics on wealth is due to differences in savings behavior and/or in rates of return on assets rather than to differences in inter vivos transfers and inheritances. In future research, we intend to investigate the race gap in savings behavior and rates of return by studying differences in the specific assets held and in growth rates.

NOTES

[1]Another important reference on the black/white wealth gap is Oliver and Shapiro (1997).

[2]See Altonji and Blank (1999) for a recent survey of the literature on the black/white gap in earnings.

[3]A cross-sectional decomposition is insufficient to accurately determine the permanent flow of income to an individual. Much of the variation in permanent income is within the categories used by the previous studies to define permanent income. Were wealth a linear function of income, ignoring the within-cell variation would not be much of an issue. Since wealth is a nonlinear function of income, making use of the within-cell variation is necessary to precisely estimate wealth models. Moreover, since high-income individuals tend to have large wealth holdings, failure to accurately measure differences in the distribution of permanent income might lead to an underestimate of the wealth gap that is out of proportion to the difference in the mean of permanent income. The problem is made more severe by the fact that there are substantial differences between whites and blacks in the distribution of income. The limited overlap in the permanent (and current) income distribution makes it difficult to use a wealth model estimated on one group to predict the wealth holding of the other group.

[4]Our results leave open the possibility that the level of inter vivos transfers and inheritances differs between whites and blacks and plays a role in the wealth gap. They rule out such intergenerational transfers as the main reason why the wealth of whites is more sensitive to income and demographic characteristics.

[5]The other components of wealth that are elicited in the wealth surveys are checking/savings, credit card, other real estate, vehicles, and other savings/assets.

[6]Throughout this article, income is short for nonasset income and all income and wealth amounts are expressed in 1989 dollars using the Consumer Price Index for urban consumers.

[7]Smith (1995) finds that "healthier households are wealthier ones" for both blacks and whites. Hence, controlling for health status helps to explain the wealth gap. The question of causality, however, is tricky.

[8]We include this because it has been used in previous studies, although there are some obvious endogeneity issues and these may lead to different biases for whites and blacks. The wealth decompositions in table 3 are not very sensitive to dropping it.

[9]This value is $\hat{W}^w - \hat{W}^b$, the difference in the weighted mean of the predictions of wealth for whites and blacks, respectively, based on the regression model. As we point out in box 2, $\hat{W}^w - \hat{W}^b$ need not be exactly equal to $\bar{W}^w - \bar{W}^b$, the difference in the weighted sample means of wealth for whites and blacks. From columns 1 and 2 of table 1, the latter figure is $152,029.

[10]We work with the indexes rather than the individual coefficients because the individual coefficients are hard to interpret given the nonlinear terms in our model and the strong covariance among some of the regressors. The relationship between the two indexes provides an overall summary of the relative sizes of the elements of α^b and α^w that is weighted by the variability of Y_t.

[11]When we use median regression we measure the wealth gap as the difference between the population-weighted averages of the conditional medians of wealth based on the median regression for the white sample and the distribution of characteristics for the white sample and the median regression and distribution of characteristics for the black sample. We refer to this as the gap in the conditional median of wealth. For more detail, see Altonji et al. (1999).

[12]Indeed, this "sedimentation of racial inequality" is one of three major themes in Oliver and Shapiro's (1997) sociological analysis of the wealth gap. For a simple model that shows how historical barriers to wealth holding among blacks could lead to present day differences in between blacks and whites in the slopes of wealth models, see Altonji et al. (1999).

[13]See Menchik (1980), Wilhelm (1996), and Menchik and David (1983) for evidence on inheritances and Altonji, Hayashi, and Kotkikoff (1997) and McGarry and Schoeni (1995) for evidence on transfers.

[14]The work of Smith (1995), Avery and Rendall (1997), and Menchik and Jianakoplos (1997) suggests that differences in intergenerational transfers contribute to the race gap in wealth. Our study of siblings does not necessarily contradict this work. Our evidence shows that intergenerational transfers are unlikely to explain why the wealth of black families is less sensitive to income and demographic variables than the wealth of white families.

REFERENCES

Altonji, J., and R. Blank, 1999, "Race and gender in the labor market," in *Handbook of Labor Economics*, O. Ashenfelter and D. Card (eds.), Vol. 3C, Amsterdam: North-Holland, pp. 3143–3259.

Altonji, J., U. Doraszelski, and L. Segal, 1999, "The role of permanent income and demographics in black/white differences in wealth," Northwestern University, working paper.

Altonji, J., F. Hayashi, and L. Kotlikoff, 1997, "Parental altruism and inter vivos transfers: Theory and evidence," *Journal of Political Economy*, Vol. 105, No. 6, pp. 1121–1166.

Avery, R., and M. Rendall, 1997, "The contribution of inheritances to black–white wealth disparities in the United States," Board of Governors of the Federal Reserve System, working paper.

Blau, F., and J. Graham, 1990, "Black–white differences in wealth and asset composition," *Quarterly Journal of Economics*, Vol. 105, No. 2, pp. 321–339.

Fairlie, R., and B. Meyer, 1997, "Ethnic and racial self-employment differences and possible explanations," *Journal of Human Resources*, Vol. 31, No. 4, pp. 757–793.

McGarry, K., and R. Schoeni, 1995, "Transfer behavior in the Health and Retirement Study: Measurement and the redistribution of resources within the family," *Journal of Human Resources*, Vol. 30, Supplement, pp. S184–S226.

Menchik, P., 1980, "Primogeniture, equal sharing, and the U.S. distribution of wealth," *Quarterly Journal of Economics*, Vol. 94, No. 2, pp. 299–316.

Menchik, P., and M. David, 1983, "Income distribution, lifetime savings, and bequests," *American Economic Review*, Vol. 73, No. 4, pp. 672–690.

Menchik, P., and N. Jianakoplos, 1997, "Black–white wealth inequality: Is inheritance the reason?," *Economic Inquiry*, Vol. 35, No. 2, pp. 428–442.

Oliver, M., and T. Shapiro, 1997, *Black Wealth/White Wealth*, London: Routledge.

Smith, J., 1995, "Racial and ethnic differences in wealth in the Health and Retirement Study," *Journal of Human Resources*, Vol. 30, Supplement, pp. S158–S183.

Wilhelm, M., 1996, "Bequest behavior and the effect of heirs' earnings: Testing the altruistic model of bequests," *American Economic Review*, Vol. 86, No. 4, pp. 874–892.