Economic Review

FEDERAL RESERVE BANK OF ATLANTA MAY/JUNE 1989

Stock-Market Volatility Forecasting

Interstate Banking Update

BANK PRODUCT DEREGULATION

The Case for Cost Savings

Economic Review

The **Economic Review** seeks to inform the public about Federal Reserve policies and the economic environment and, in particular, to narrow the gap between specialists and concerned laypersons.

Views expressed in the **Economic Review** are not necessarily those of this Bank or of the Federal Reserve System.

Free subscriptions and limited additional copies are available from the Public Information Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., Atlanta, Georgia 30303-2713 (404/521-8788). **Change-of-address notices and subscription cancellations should be sent directly to the Public Information Department. Please include the current mailing label as well as any new information.**

ISSN 0732-1813

VOLUME LXXIV, NO. 3, MAY/JUNE 1989, ECONOMIC REVIEW

FEDERAL RESERVE BANK OF ATLANTA

Does Multiproduct Production in Large Banks Reduce Costs?

William Curt Hunter and Stephen G. Timme

One justification often cited for further deregulation of the products banks offer is the cost savings that can result from joint production of a variety of financial services. In this article, the authors show that, though consumers may save money on transactions costs when a bank provides a number of financial services, large banks do not necessarily incur lower costs by engaging in multiproduct production.

Legal and regulatory changes have greatly influenced the operations of U.S. financial institutions, especially commercial banks and bank holding companies. Two major trends are the liberalization of longstanding restrictions on geographic expansion and the extension of the products legally available from banking organizations. These developments have led to increased interstate banking and to an explosion in the variety of services and products that commercial banks offer. Although individual institutions have long provided a variety of products to consumers, only recently has the ability to offer multiple products under one corporate banner been cited as a major justification for increased deregulation.

The argument in support of allowing banks to provide an unlimited menu of financial products, or global product deregulation, is that such a move will benefit society by making more choices available to consumers while reducing their transactions costs, that is, the monetary and nonmonetary costs such as travel time associated with traveling to different institutions and seeking information. This argument is quite appealing. By having a host of financial services available at one location, consumer welfare may be enhanced even if financial institutions incur higher costs in providing these services. As long as the amount consumers save on transactions costs exceeds whatever increase in bank production costs they would bear as a result of the bank's offering more products, consumers are unambiguously better off.

The call for increased product deregulation in banking has justifications beyond the simple one given above. By offering additional products, some analysts maintain, banks will actually face lower total costs because of certain economies that result from multiproduct production. Thus, consumers will save not only on transactions costs but, to the extent that banks pass on production cost savings, on the total costs of bank products as well. The likelihood that these savings will be shared with consumers depends on the extent to which competition exists among the banks and nonbanks offering these products.

This article addresses the issue of global product deregulation by focusing on bank production efficiency issues. In particular, it takes a look at the available literature and offers some new empirical evidence as to whether, among the largest banks in the economy, cost savings result from multiproduct production.

It is true that widespread innovations in the financial services industry have spawned much of the recent product deregulation in banking. Still, even though the production process is certainly influenced by the status of product deregulation, banks continue to develop new products on their own. Product development personnel, to perform their jobs optimally, must be informed about the extent to which generalized multiproduct production enhances the efficiency and competitiveness of commercial banks. Thus, gauging the efficiency potential in this area is a vital step in policy considerations.

Measuring the Benefits of Multiproduct Production

Economists have tried to measure the cost-reducing benefits of multiproduct production using the notion of economies of scope. Simply put, economies of scope are said to exist when it is cheaper for one company to produce several products than for a number of specialty firms to develop each product separately. Economies of scope defined in this manner are called global or generalized economies of scope, and they typically occur when several different products can be made through a single production process.[1]

Economies of scope can result from the use of fixed inputs in the production process. Since some fixed inputs, including both physical and human capital, are lumpy, that is, not perfectly divisible and thus not easily shifted from one type of production to another, making one subset of products may leave the firm with excess capacity. In this case, the company may benefit from using its surplus resources to create other goods and services. On the other hand, some of the firm's inputs could have a quasi-public characteristic. That is, a valued product might

---

The authors are, respectively, a research officer at the Federal Reserve Bank of Atlanta and an associate professor of finance at Georgia State University.

be obtained simply as a by-product of some other production. A firm's inputs could also be interrelated in such a manner that it is cheaper for the company to create a given product in conjunction with others than to make just that product; this potential saving results from product-specific economies of scope.

A simple example illustrating the concept of economies of scope is that of the producer of wool and mutton, which appears in many elementary economics textbooks. The example demonstrates that, under normal conditions, producing wool and mutton from one herd of sheep is cheaper than producing them from two different herds, one for mutton and the other for wool.

Economies of scope may also derive from demand-side considerations. By making a number of products, a company can avoid the risks associated with single-product production, such as variations in revenues resulting from changes in demand. In this way, the institution actually broadens its product offering in much the same manner and for the same reasons that investors diversify their security portfolios. Thus, the presence of product market uncertainty also provides an impetus for economies of scope.

Based upon the above discussion, one may conclude that the production processes of U.S. commercial banks harbor significant economies of scope. After all, these institutions utilize fixed and indivisible resources in their production processes, employ inputs that have quasi-public features, and face significant uncertainties with regard to the nature and extent of the demand for their products. However, based on some recent empirical studies of economies of scope in the banking industry, the conclusion that banks experience significant economies of scope may not be warranted.

Studying Bank Production

Two basic approaches tend to be used in empirical studies of bank production: the production approach and the intermediation approach. According to the former, a bank's inputs include labor and physical capital such as computers and other types of equipment. Bank output is measured by the number of accounts held at the bank. Costs under the production approach include all of the operating expenses that the institution incurs over the relevant period of analysis. Interest paid on deposits and other borrowed funds is not considered an operating cost and thus is not included in the definition of relevant production costs.

The second approach concentrates on the process of financial intermediation by which banks borrow from savers and subsequently lend to borrowers. Output is measured by the total dollar value of the various products that the bank offers. The relevant definition of production costs for the intermediation approach is total costs, which includes both interest expenses as well as operating costs. The intermediation approach is generally accepted as more relevant for addressing questions related to the long-run competitiveness of banks.[2]

Studies that use the production approach in examining bank costs generally rely on data collected by the Federal Reserve's Functional Cost Analysis (FCA) program, through which certain banks report cost data on a variety of bank functions. However, the results obtained from these studies cannot, in general, be extrapolated to the large bank population because the survey is conducted on a voluntary basis and banks with over $1 billion in total deposits tend to be underrepresented. This weakness associated with use of the FCA data undermines the production approach since the FCA program is the only reliable source for account-based data on large samples of banks. Nonetheless, researchers can offset this shortcoming because account-based data lend themselves more readily to gauging bank output.[3]

Deciding on the proper indicator of bank output is a common problem. Under the production approach the dilemma is less serious since the definition of output is account-based. Thus, one has only to decide which accounts to include in the analysis. On the other hand, under the intermediation approach the researcher must consider carefully which bank assets and liabilities to include in the definition of output. Since the results of cost studies in banking can be sensitive to the definition of output, this consideration is far more important under the intermediation approach. Ideally, each bank product should be considered as a distinct output. However, the analysis quickly becomes intract-

able as the number of outputs is expanded.[4] Thus, under the intermediation approach choices have to be made with respect to the number and types of outputs to include in the analysis.

For most purposes of determining the appropriate definition of bank output under the intermediation approach, the criterion of *value-added* presented by Allen N. Berger, Gerald A. Hanweck, and David B. Humphrey (1987) is sufficient. This standard simply requires that for a product to be considered part of a bank's output, consumers must value the product more highly in the form produced by the bank than if they were to generate the same output themselves. Demand deposit accounts are one example of a bank product that meets this criterion. Consumers value the safety and convenience provided by bank checking accounts more highly than transacting in cash alone and thus are willing to pay for this service.

Measuring Economies of Scope

To determine whether generalized economies of scope are in effect, one usually measures the cost differential that would arise between the joint and independent production of given output levels of all products. This cost differential is then adjusted by dividing by the total costs of joint production. To illustrate this concept, assume that the output to be measured is the dollar volume of loans. In this hypothetical example one bank produces $1 billion in retail loans and $3 billion in wholesale loans each year, while another pair of banks specialize, one issuing $1 billion in retail loans and no wholesale loans and the other granting $3 billion in wholesale loans and no retail loans.

To measure economies of scope in this case one would compute the difference between the total costs associated with the first bank's production and that of the pair. This difference would then be put on a percentage basis by dividing it by the total production costs associated with the multiproduct bank. Clearly, if this index is negative, meaning that the cost of joint production is less than the cost of independent production, global economies of scope are present. An index greater than zero indicates diseconomies of scope.

Several ways to measure product-specific economies of scope are available. The simplest and most widely used procedure involves the examination of pairwise cost complementarities. That is, the analyst determines how the joint creation of two products in the firm's offering affects the marginal cost of producing each product. If the marginal cost of making one product declines with an increase in the output of the second product, and this relationship holds for all products, global cost complementarities exist. In other words, in the two-product case cost complementarities make it possible for the firm to produce the two products more efficiently from a single production process than independently.

The Empirical Evidence

Several studies in the literature examine economies of scope using the production approach.[5] Among these papers, only Thomas Gilligan and Michael Smirlock's (1984) and Gilligan, Smirlock, and William Marshall's (1984) report significant global or generalized economies of scope. On the other hand, many of these studies report significant product-specific economies of scope. In general, significant economies of scope were observed for the following output pairs: investments and loans, deposits and loans, and time deposits and demand deposits.

H. Youn Kim's (1986) and Colin Lawrence and Robert Shay's (1986) studies are representative of research utilizing the intermediation approach in examining economies of scope in banking. Kim found both global and product-specific scope economies in a sample of Israeli banks when output was defined to include four specific products: demand deposits, foreign currency, loans, and securities. However, when output was defined to be only a subset of these four products—for example, loans and securities—neither global nor product-specific economies of scope were found. The study by Lawrence and Shay finds no significant global economies of scope for the sample analyzed, but it does report significant product-specific economies of scope between deposits and investments as well as between deposits and loans. This study

also finds economies of scope between loans and nonbank activities—safe deposit, trust, data services, and other agency-type services as well as nonbank investments and activities like mortgage lending or data processing. However, the study reports strong diseconomies of scope between loans and investments.

These findings must be interpreted with caution for several reasons. First, notwithstanding variations in statistical methodologies, cost and output definitions, and differences in underlying technical assumptions, the results are based almost exclusively on samples consisting of banks with less than $1 billion in total deposits. Thus, these conclusions do not necessarily characterize the nature of joint production in larger banks.[6] Second, since banks generally do not produce only one product, the application of these measures to the banking industry necessitates extrapolating the estimated cost function well beyond the range of the observed data, which perhaps explains the inconclusiveness of the literature's reported results.[7]

The Cost Structure of Large Banks

Despite the body of results available for smaller banks, the structure of multiproduct production among the U.S. economy's largest banks is still fraught with unknowns. Given the importance of the large money center and regional banks in the national and world economies, this ignorance is somewhat surprising. In an effort to answer the question of whether significant cost savings are associated with multiproduct production in the economy's largest banks, the cost structure of a large-bank sample was examined using a methodology immune to the problems noted in the previous section.[8]

This examination utilized the intermediation approach to bank production. Five bank products were included in the analysis: (1) retail loans; (2) wholesale loans; (3) transactions accounts, that is, demand deposits, negotiable orders of withdrawal, and other interest-bearing checking accounts; (4) nontransactions accounts, such as savings accounts and small time deposits; and (5) a proxy for other bank products intended to capture the increasingly important area of off-balance sheet activities such as loan sales, letters of credit, and securitization activities. The proxy for other bank products was defined as annual noninterest income excluding both service charges received on transactions and nontransactions accounts as well as gains and losses on securities and foreign-exchange transactions. Total costs were defined as total noninterest costs plus allocated interest expense as appropriate under the intermediation approach. Since securities are not examined as an output, allocated interest expense is defined as the product of the ratio of total loans to total earning assets and total interest expense.

The Data. Financial data for the 400 largest U.S. banks in terms of total assets were collected from the "Reports of Condition" and "Reports of Income and Dividends" filed with the Federal Reserve during 1986. Forty-seven banks in states with unit banking laws were dropped from the analysis since branch banking structures are regarded as most appropriate for drawing general conclusions concerning the banking industry. Twenty-eight other banks were also excluded from the sample for assorted reasons. For example, data for banks headquartered in Puerto Rico and banks with credit cards as their major source of business were discarded. Also, 13 banks with either incomplete data or significant data discrepancies were eliminated. The final sample included 312 banks. Table 1 presents summary statistics for all banks in the sample and by eight subgroups based on total asset size.

The Methodology. In order to avoid many of the technical problems associated with measuring economies of scope, this study adapts and augments a production-efficiency measure developed by Berger, Hanweck, and Humphrey in examining the cost structure of the banks sampled. These researchers developed a measure called expansion-path subadditivity as an alternative to the traditional measures of economies of scope given earlier.

Berger, Hanweck, and Humphrey contend that as banks expand along a certain path they alter their product mixes, as illustrated by the growth of many institutions through acquisitions or mergers with other banks offering different product mixes. Thus, the expansion-path subadditivity index measures the percentage difference in the total costs of a single firm's pro-

Table 1.
Bank Output for Sample Banks at End of Year 1986*
(billions of dollars)

Asset Size	Number of Banks	Total Assets	Retail Loans	Wholesale Loans	Transactions Accounts	Nontransactions Accounts	Other Output
$ 0.8-1.0	45	$ 0.88 (0.06)	$ 0.17 (0.08)	$ 0.38 (0.10)	$ 0.25 (0.07)	$ 0.51 (0.08)	$ 0.01 (0.01)
$ 1.0-1.5	79	1.23 (0.15)	0.25 (0.14)	0.50 (0.14)	0.36 (0.11)	0.66 (0.13)	0.01 (0.01)
$ 1.5-2.0	36	1.73 (0.21)	0.37 (0.17)	0.69 (0.11)	0.52 (0.15)	0.93 (0.13)	0.01 (0.03)
$ 2.0-3.0	48	2.50 (0.28)	0.42 (0.19)	1.07 (0.25)	0.70 (0.20)	1.29 (0.24)	0.02 (0.01)
$ 3.0-5.0	41	3.85 (0.63)	0.70 (0.35)	1.60 (0.47)	1.10 (0.29)	1.82 (0.44)	0.03 (0.02)
$ 5.0-10.0	40	7.11 (1.43)	1.17 (0.63)	3.08 (1.05)	2.04 (0.75)	3.08 (0.83)	0.06 (0.07)
$ 10.0-25.0	14	13.70 (3.18)	1.98 (1.18)	6.11 (2.57)	3.42 (0.81)	5.52 (1.48)	0.11 (0.07)
$ 25.0 plus	9	46.05 (15.86)	5.39 (4.30)	20.09 (7.55)	12.65 (4.15)	16.88 (8.94)	0.89 (0.60)
Overall	312	4.38 (8.25)	0.68 (1.21)	1.88 (3.69)	1.22 (2.25)	1.92 (3.22)	0.05 (0.18)

* *Standard deviations are given in parentheses below the means.*

Source: Computed at the Federal Reserve Bank of Atlanta from data in "Reports of Condition" and "Reports of Income and Dividends" filed with the Federal Reserve during 1986.

ducing a set of products in given quantities compared with the costs incurred by a pair of companies producing the same set of products whose value, when summed, equals that of the single firm. That is, the expansion-path subadditivity index subtracts from the single firm's total cost of producing a set of financial services the expenses incurred by the two other firms. This amount is then scaled by dividing the sum by the single firm's total cost.

For example, Bank A produces $1 billion in retail loans and $3 billion in wholesale loans, and Bank B produces $0.5 billion in retail loans and $1 billion in wholesale loans. Expansion-path subadditivity measures the difference, on a percentage basis, in the costs of Bank A relative to the costs of Bank B and Bank C, which produces $0.5 billion in retail loans and $2 billion in wholesale loans. An expansion-path subadditivity index of less than zero represents the percentage reduction in total costs that the single bank incurs relative to the total cost incurred by the pair of banks. Thus, the expansion-path subadditivity index is based on a more realistic treatment of bank product mixes and is also more reflective of the manner in which banks are likely to expand in terms of both the number and quantity of their products.

In this article's analysis of the sample banks, the expansion-path subadditivity index is augmented with the grid approach developed by David S. Evans and James J. Heckman (1984) in their analysis of costs in the Bell System. The Evans and Heckman procedure is similar to the Berger, Hanweck, and Humphrey expansion-path subadditivity test in that it compares the percentage difference in costs of a single bank producing a given output mix relative to the same output mix being divided between two "competing" banks. However, unlike the expansion-path subadditivity test, the grid approach requires (1) that each firm produce some percentage of a minimum observed quantity of all products and (2) that each bank's final product mix be consistent with the observed data. For a given range of observed product mixes, a

Table 2.
Expansion-path Subadditivity and Maximum Grid Subadditivity

Asset Size (in billions of dollars, end of year 1986)	Expansion-path Subadditivity (percent)	Maximum Grid Subadditivity (percent)
$ 0.8 - 1.0	NA	NA
$ 1.0 - 1.5	-0.23	-0.71
$ 1.5 - 2.0	-0.50	-0.43
$ 2.0 - 3.0	0.73	-0.12
$ 3.0 - 5.0	0.39	1.51
$ 5.0 - 10.0	1.23	2.78*
$10.0 - 25.0	1.52	4.81*
$25.0 plus	4.24*	16.81*

** Significant at the .01 level.*

grid of admissible combinations of competing firms is constructed by varying the percentage of the minimum quantity produced by each competing firm. At each point within the grid, the total cost of a single firm's producing the combined product mix is compared on a percentage basis to the sum of the total costs of the competing firms. This index at each point is called grid subadditivity.

The expansion-path subadditivity index can be viewed as depicting just one possible division of the mean quantities of the products of a representative bank between two competing banks contained within a grid composed of banks which are most likely to be sources of competition. Clearly, additional insights into the potential cost benefits associated with multiproduct production may be gained by examining the behavior of costs resulting from other divisions of output within the grid, which, by construction, is a representative section of the bank cost function.

A statistical test was conducted to see if each computed grid subadditivity index was equal to zero, that is, if no cost savings or dissavings are associated with multiproduct production. If the maximum value of the grid subadditivity index within a grid is statistically significant and negative, the hypothesis that multibank production leads to no cost savings is rejected. Such a result would lend support to the notion that cost savings do indeed accrue to multiproduct production. A statistically significant positive value of the maximum grid subadditivity index within a grid provides evidence that cost disadvantages (dissavings) are associated with multiproduct production.

To implement both the expansion-path subadditivity and grid subadditivity tests, the overall sample of banks was divided into eight subgroups based on total assets (see Table 1). For each subgroup a measure of expansion-path subadditivity was estimated in addition to a construct of grid subadditivity indices. Within each grid the maximum grid subadditivity index was obtained and reported. This procedure provides a more stringent test of the hypothesis that significant cost savings are associated with multiproduct production by banks.

Results of the Tests. Table 2 reports estimates of the expansion-path subadditivity and maximum grid subadditivity indices for each of the asset subgroups. The expansion-path subadditivity and maximum grid subadditivity levels in this table measure the percent increase (decrease) in total costs associated with a single bank's producing the given products relative to this same output determinant's being divided between two competing banks.

Based strictly on the numerical values of the expansion-path subadditivity index, the results presented in Table 2 indicate that slight cost savings are associated with multiproduct production for the sample banks with total assets up to approximately $2 billion, no cost savings exist for banks with total assets in the $2 billion to $5 billion range, and increasing cost dissavings occur for banks with total assets exceeding $5 billion. With the exception of the largest asset subgroup, however, none of the computed values of expansion-path subadditivity are statistically different from zero. Thus, based on the expansion-path subadditivity indices in Table 2, no statistically significant cost savings or dissavings can be linked with multiproduct production for banks with asset holdings up to $25 billion.

As stated earlier, the grid subadditivity index permits the examination of many more points on the bank cost function and thus generalizes the expansion-path subadditivity measure. The maximum grid subadditivity indices in Table 2 provide slightly different evidence from the

expansion-path subadditivity indices. The numerical values of the maximum grid subadditivity index indicate slight cost savings associated with multiproduct production for banks with assets up to $3 billion and increasing dissavings for institutions with asset sizes above $3 billion. However, an examination of the statistical significance of these numerical values reveals that the maximum grid subadditivity indices for asset sizes up to $5 billion cannot be regarded as different from zero. Thus, the more robust maximum grid subadditivity tests imply that no statistically significant cost savings or dissavings are associated with multiproduct production for banks with up to $5 billion in assets. Beyond the $5 billion level, the index values imply that statistically significant dissavings are involved in multiproduct production. The difference between the asset size at which dissavings set in under the two measures reflects the fact that the maximum grid subadditivity test examines more points on the bank cost function than the expansion-path subadditivity test.

Implications and Conclusions

For commercial banks with up to $5 billion in assets, the results of this study's more reliable index suggest that no appreciable cost savings accrue from multiproduct production, implying that consumers should not expect significant cost savings solely because of efficiencies associated with multiproduct production at banks up to this size. However, multiproduct production in these banks should unambiguously increase consumer welfare if it only increases consumer convenience and lowers transactions costs. Finally, for sample banks with up to $5 billion in assets, multiproduct production does not appear to give such banks a cost advantage relative to other bank and nonbank competitors.

For banks with assets between $5 billion and $25 billion, the results of the grid subadditivity test indicate that multiproduct production actually increases production costs, suggesting that these banks should either alter their product mix or reduce their scale of output. However, the model in this study does not explicitly incorporate convenience to customers, and such banks may be reducing consumers' transactions costs through the array of services offered. Moreover, these banks may engage in multiproduct production also as a means of diversification to reduce risk even though the resulting scale of output or product mix is not optimal when viewed strictly from a production perspective. If adjusted for the output of convenience and the benefits of diversification, the apparent cost disadvantage of multiproduct production, as measured by the grid subadditivity index, should be substantially reduced. Thus, even for the largest banks, multiproduct production does not appear to afford a cost advantage that could be used to the detriment of potential competitors. Instead, the results indicate that product liberalization would likely increase bank customer welfare by reducing transactions costs and helping to diversify banks.

Notes

[1]A related concept, discussed later in this article, is that of product-specific economies of scope, which refer to savings that arise from the joint production of a particular product with other products. That is, producing a given product along with other products is cheaper than producing it independently.

[2]Humphrey (1985) contains a thorough discussion of the production and intermediation approaches to examining bank production.

[3]The question of what banks produce, that is, what the appropriate definition of bank output is, has plagued financial economists for decades. For an interesting discussion of the problems associated with defining bank output, see Sealey and Lindley (1977).

[4]The problems encountered relate to both the nature of the data being analyzed and the form of the cost function that is to be fit statistically. Though a detailed discussion of the econometric problems associated with the measurement of scope economies in banking is beyond the focus of this article, a very readable discussion of the issues is contained in Clark (1988).

[5]The studies by Benston, Hanweck, and Humphrey (1982); Benston et al. (1983); Gilligan and Smirlock (1984); Gilligan, Smirlock, and Marshall (1984); Berger, Hanweck, and Humphrey (1987); and Kolari and Zardkoohi (1987) fall into this category. Some research examines the existence of economies of scope in other financial institutions, particularly savings and loan associations and credit unions. See the studies by Murray and White (1983); Kim (1986); Mester (1987); and Goldstein, McNulty, and Verbrugge (1987).

[6]Recent studies by Hunter and Timme (1986, 1988), Shaffer (1988), and Shaffer and David (1986) examining production economies in samples composed of the largest banks in the economy report evidence that the largest banks operate on a different long-run average cost curve than do smaller banks. Thus, the results obtained from an examination of the smaller banks in the economy should not be extrapolated to the banking industry as a whole.

[7]For a discussion of the technical problems encountered in measuring economies of scope using banking data, the interested reader should see Berger, Hanweck, and Humphrey (1987): 505-06.

[8]The complete details of this study will appear in a forthcoming Federal Reserve Bank of Atlanta working paper.

References

Benston, George J., Gerald A. Hanweck, and David B. Humphrey. "Scale Economies in Banking: A Restructuring and Reassessment." *Journal of Money, Credit and Banking* 14 (November 1982): 435-56.

Benston, George J., Allen N. Berger, Gerald A. Hanweck, and David B. Humphrey. "Economies of Scale and Scope in Banking." Paper presented at a Federal Reserve Bank of Chicago Conference on Bank Structure and Competition, May 1983.

Berger, Allen N., Gerald A. Hanweck, and David B. Humphrey. "Competitive Viability in Banking: Scale, Scope, and Product Mix Economies." *Journal of Monetary Economics* 20 (December 1987): 501-20.

Clark, Jeffery A. "Economies of Scale and Scope at Depository Financial Institutions: A Review of the Literature." Federal Reserve Bank of Kansas City *Economic Review* 73 (September/October 1988): 16-33.

Evans, David S., and James J. Heckman. "A Test for Subadditivity of the Cost Function with an Application to the Bell System." *American Economic Review* 74 (September 1984): 615-23.

Gilligan, Thomas, and Michael Smirlock. "An Empirical Study of the Joint Production and Scale Economies in Commercial Banking." *Journal of Banking and Finance* 8 (March 1984): 67-76.

Gilligan, Thomas, Michael Smirlock, and William Marshall. "Scale and Scope Economies in the Multi-Product Banking Firm." *Journal of Monetary Economics* 13 (March 1984): 393-405.

Goldstein, Stephen J., James E. McNulty, and James Verbrugge. "Scale Economies in the Savings and Loan Industry Before Diversification." *Journal of Economics and Business* 39, no. 3 (1987): 199-207.

Humphrey, David B. "Cost and Scale Economies in Bank Intermediation." In *Handbook for Bank Strategy*, edited by R. Aspinwall and R. Eisenbeis. New York: Wiley & Sons, 1985.

Hunter, William C., and Stephen G. Timme. "Technical Change, Organizational Form, and the Structure of Bank Production." *Journal of Money, Credit and Banking* 18 (May 1986): 152-66.

Hunter, William C., and Stephen G. Timme. "Technological Change and Production Economies in Large U.S. Banks." Federal Reserve Bank of Atlanta Working Paper 88-6 (December 1988).

Kim, H. Youn. "Economies of Scale and Scope in Multiproduct Financial Institutions: Further Evidence from Credit Unions." *Journal of Money, Credit and Banking* 18 (May 1986): 220-26.

Kim, Moshe. "Banking Technology and the Existence of a Consistent Output Aggregate." *Journal of Monetary Economics* 8 (March 1985): 181-95.

Kolari, James, and Asghar Zardkoohi. *Bank Costs, Structure, and Performance.* Lexington, Mass.: D.C. Heath, 1987.

Lawrence, Colin, and Robert Shay. "Technology and Financial Intermediation in the Multiproduct Banking Firm: An Econometric Study of U.S. Banks 1979-1982." In *Technological Innovation, Regulation and the Monetary Econ-*

omy, edited by C. Lawrence and R. Shay. Boston, Mass.: Ballinger, 1986.

Mester, Loretta J. "A Multiproduct Cost Study of Savings and Loans." *Journal of Finance* 42 (June 1987): 423-45.

Murray, John, and Robert White. "Economies of Scale and Economies of Scope in Multiproduct Financial Institutions: A Study of British Columbia Credit Unions." *Journal of Finance* 38 (June 1983): 887-902.

Sealey, C.W., and J.T. Lindley. "Inputs, Outputs, and a Theory of Production and Cost at Depository Financial Institutions." *Journal of Finance* 32 (September 1977): 1251-66.

Shaffer, Sherrill. "A Revenue-Restricted Cost Study of 100 Large Banks." Federal Reserve Bank of New York unpublished research paper, 1988.

Shaffer, Sherrill, and Edmond David. "Economies of Superscale and Interstate Expansion." Federal Reserve Bank of New York unpublished research paper, 1986.

Forecasting Stock-Market Volatility Using Options on Index Futures

Steven P. Feinstein

The ability to anticipate stock-market volatility can be a most helpful tool for investors, market analysts, and policymakers. An Atlanta Fed financial economist examines implied-volatility forecasts derived from options on S&P 500 index futures.

Financial economists, especially those who work for the Federal Reserve, are often asked: Is the stock market going up or down? The most honest answer, and the only one that should be given, is "Yes, most likely." After the requisite resigned chuckle, a more persistent questioner will occasionally follow the first query with "How much is the market likely to fluctuate?" This follow-up is both fair and relevant. Most people who follow the market have a direct investment interest and need to know an asset's degree of risk in order to decide whether the expected return on the investment represents adequate compensation.[1] Volatility forecasts are also important to policymakers, who might use these forecasts to gauge market uncertainty and to determine when to change policy or to make announcements with minimal disruption to the financial markets.

Stock market volatility is constantly changing.[2] Therefore, investors and market watchers must continually update their volatility forecasts, which is no easy task because of the numerous factors involved. Past returns, a firm's leverage, announcements affecting a company or the whole market, option and futures expirations, holidays, expanded hours of trading, and

---

The author is an economist in the financial section of the Atlanta Fed's Research Department. *He thanks Phil Dybvig, Frank King, and Peter Abken for helpful comments, and gratefully acknowledges David Stowe's extensive research assistance. The full data set upon which this article is based is available from the author upon request.*

changes in margin requirements are just some of the influences on stock volatility that have been tested.[3] Though digesting all of these elements and turning them into an integrated forecast is difficult, a short-cut method, pioneered by Henry A. Latane and Richard J. Rendleman, Jr. (1976), is available; this method allows one to infer a volatility forecast from prices of stock options. (An option is a contract that affords the buyer the right, but not the obligation, to buy or sell an asset for a prespecified price on or before some selected date. They differ from futures contracts in that a futures contract requires the purchase or sale of the underlying asset on the given date.) This "implied volatility" should inherently consider all available information on relevant factors.[4]

Stock-index options and stock-index futures began trading in January 1983. The presence of a market for stock-index options makes it possible to compute implied volatilities that pertain to the stock market as a whole rather than to a single security. A sufficient period of time has now elapsed since the introduction of stock-index options and futures that research can determine the viability of using implied volatilities based on these instruments to predict overall stock-market volatility.

This article describes the theory behind the implied-volatility method, searches for the best way to measure it, tests it for efficiency and bias, and illustrates its usefulness for anticipating change in stock-market volatility.[5] Whereas previous studies have addressed these issues using cross-sectional data from several stocks collectively over various forecast periods, this study focuses on the volatility of one vehicle, the Standard and Poor's 500 index, over several years.[6] This time-series approach can—more definitively than earlier studies—establish whether the implied-volatility approach is practical for forecasting volatility on an on-going basis.

Implied-Volatility Theory

A stock option's price depends on the anticipated future volatility of the stock underlying the option. (See the box on page 14 for a simplified explanation of this connection.) This relationship has been formalized by Fischer Black and Myron S. Scholes (1972, 1973) in an explicit option-pricing formula that has held up quite well under extensive empirical testing.[7] The Black-Scholes option-pricing formula relates an option's price to five variables: the stock price; the strike price of the option (that is, the price that is written into the option contract); the interest rate; the time remaining before the option's expiration; and the volatility of the underlying stock price, which in this case is the standard deviation of the stock's returns.[8] The stock price, the strike price, and time to expiration are easily observed, and reasonable proxies exist for the interest rate. The option price itself is determined in the market and is likewise observable. Thus, of the six variables in the Black-Scholes formula (five inputs and one output), only the volatility cannot be directly observed.

In practice, people use forecasts for future stock volatility—the last element of the Black-Scholes formula. Market watchers, such as other investors, policymakers, and regulators, can thus extract the forecast by finding the volatility value that is consistent with the observed option price and the other variables in the option-pricing formula. The two researchers (1972) showed that sizable profits could be reaped from incorporating perfect forecasts of volatility into their formula. One can assume safely, then, that option traders would try to use all pertinent information in order to construct the best possible forecasts of future volatility.

A 1988 paper by this author showed that for at-the-money options—for which the discounted strike price equals the stock price—inserting an unbiased estimate of the standard deviation of future stock returns into the Black-Scholes formula produces an unbiased estimate of the correct Black-Scholes no-arbitrage option price, at which the strategy of trading in both options and stocks yields no unusually high profits.[9] Thus, when pricing at-the-money options, agents' forecasts of volatility should be unbiased, that is, correct on average.

Conceptually, one would like to invert the Black-Scholes formula, such that stock-option volatility could be determined from the five observable variables. Although the formula cannot be inverted explicitly, one can use iterative methods to identify the volatility value that is

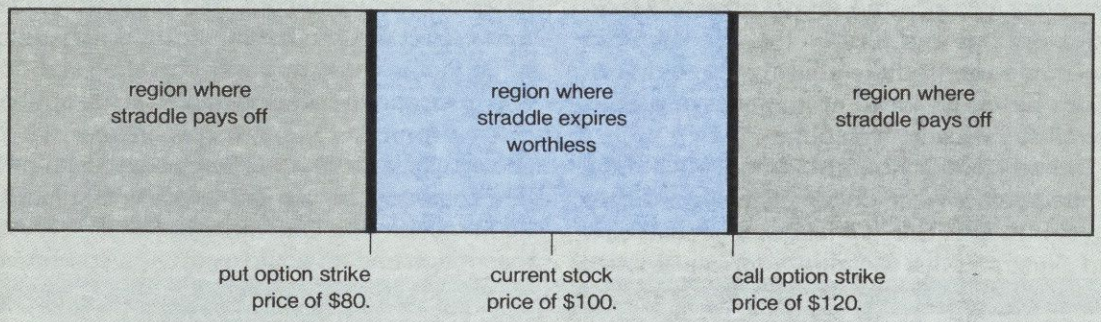
Why Option Prices Reflect Anticipated Stock Volatility

Full appreciation of the relationship between option prices and stock volatility requires close scrutiny of the Black-Scholes formula and an understanding of its derivation. An intuitive grasp, however, can be acquired from consideration of the option strategy known as a "straddle."

A straddle is the purchase of both a call option and a put option on the same stock. The call option will pay off if the stock price rises, whereas the put option will pay off if the stock price falls sufficiently. Suppose, for example, a stock is selling for $100. A straddle may consist of a call option with strike price $120 and a put option with strike price $80. The call and put options will usually sell for considerably less money than these amounts. Should the stock price rise beyond $120, the call will become valuable while the put becomes worthless. Should the stock price fall below $80, the put will be in the money while the call becomes worthless.

The straddle will pay off if the stock price bounces past *either* strike price. The figure below shows the potential stock prices for which the combined straddle portfolio (owning both the put and the call) generates positive payoffs. As long as the stock price moves sufficiently in either direction, whether the stock rises or falls is immaterial.

The more volatile a stock is, the more likely it is to bounce outside the range defined by the strike prices. Consequently, straddles on more volatile stocks have greater likelihoods of producing revenue. Thus, the initial price of a straddle—that is, the sum of the initial put and call prices—must depend on the anticipated stock volatility. Modern option-pricing theory provides methods to value the individual components of the straddle. The dependence on stock volatility is a feature retained by each option individually.



Should the stock prices move past either strike price, represented by the dark vertical lines, the straddle portfolio—consisting of one call and one put option—will become valuable. Should the stock value price remain in the blue-shaded region, near its initial value, the straddle will expire worthless.

Since a more volatile stock is more likely to bounce past the strike prices, options on such stocks must cost more. Consequently, stock option prices reflect anticipated stock volatility.

consistent with the observed option price. This extracted measure is the Black-Scholes implied volatility.

Properties of the Implied Volatility. Implied-volatility forecasts have several advantages over historically based estimates of stock volatility, such as the sample standard deviations from a history of past stock-price variations. Historical projections are necessarily backward-looking; to be used as forecasts, these estimates require an assumption that future performance will resemble the past's. The implied volatility, alternatively, is forward-looking and, by design, a forecast of future vicissitudes.

Unlike historical estimates, the implied volatility is instantaneous—that is, it can be computed from data generated at any given moment. Since option prices quickly incorporate new information, the implied volatility responds immediately, reflecting changes in the volatility forecast. Unlike the historical estimate, which reacts slowly to changing conditions, the implied volatility can detect sudden changes in market sentiment.

Although the computations are more complex for the implied-volatility forecast than for the measurement of historical volatility, much less data are required. Historical volatilities are based on many observations of past stock returns, but the implied volatility requires only current values of the five observable Black-Scholes variables.

Other Consolidated Forms of Multiple Implied Volatilities

On any underlying security, simultaneous trading takes place in several call options, each differing in strike price and maturity date. A *call* option grants the right to buy a fixed amount of an underlying asset at a specified price within a given period of time; a *put* option, on the other hand, is an option contract designed to sell the underlying asset. For the Standard and Poor's (S&P) 500 futures call options, five different strike prices and three different maturity classes are usually available. (A *maturity class* refers to all options with the same expiration date.)

Each of these contracts provides its own implied volatility. Furthermore, the put options actively traded on these contracts provide implied volatilities too. On any given trading day, an investor can purchase a put or call on an S&P 500 index future with one of three expiration dates and five prices corresponding to each of the dates. Thus, the market actually offers approximately 30 different S&P 500 index futures options at a given time, each of which can yield an implied-volatility forecast. Since the many implied volatilities that can be derived from these options rarely agree, even within the same maturity class, researchers face a dilemma: when forecasting volatility over the life span of one maturity class, which implied volatility should be trusted?

Various schemes have been suggested to incorporate the multiple implied volatilities into one optimal forecast. Each plan tries to produce a consolidated forecast in which bias is minimized and efficiency, maximized. The approaches range from choosing one particular contract's implied volatility and discarding all others to constructing various weighted or simple averages from the multiple implied volatilities. A later section of this article searches for the best implied-volatility form by examining the performance of the established consolidation methods and by considering one new technique.[10]

An Innovative Implied-Volatility Form. Averaging several implied volatilities from options with different strike prices has the desirable effect of reducing the impact of any data error or anomalous price, but this method also has an undesirable consequence. J.S. Butler and Barry Schachter (1986) showed that implied volatilities from most options should necessarily be biased, particularly for options that are well in or out of the money. (With in-the-money options, the futures price relates to the exercise price in such a manner that the purchaser profits from exercising the option; for a call option, the index price exceeds the exercise price, and for a put option, the index price is lower than the exercise price. At expiration, out-of-the-money options result in a complete loss of the investor's money.) Although this bias should not be present in at-the-money options (Steven P. Feinstein [1988]), averaging implied volatilities derived from out-of-the-money options together with those from at-the-money options contaminates the theoretically unbiased at-the-money implied volatilities with systematic errors.

With this problem in mind, another strategy for reducing noise while still focusing on at-the-money options was investigated; a measure was constructed from intertemporal averaging, that is, averaging the implied volatilities from at-the-money call options observed on a small number of successive days. Some error is introduced because not completely up-to-date volatility forecasts are included; also, the implied-volatility forecasts encompass the lagged dates as well as the remaining future horizon. Yet, as long as the forecast horizon is long relative to the number of lags averaged, the informational deficiency and slight mismatch in forecast horizons should cause little damage.

The exact construction of this new form of implied-volatility forecast is given in Appendix I, and the established forms are discussed in more detail in the data section that follows. The new forecast method and the established implied-volatility forms were compared on the basis of efficiency and bias. As will be shown, the new technique proved valuable.

Empirical Tests

Implied-volatility forecasts were constructed and compared with the stock-market volatility that actually transpired. A naive forecast, consisting of the sample standard deviation of stock-market returns from the 20 trading days (roughly one month) up to and including the date of the implied-volatility forecast, provided a basis for comparison. Mean squared errors and mean absolute errors were computed as measures of the forecasts' efficiency, and the various forms of implied-volatility forecasts were ranked according to these indicators. The forecasts were tested for bias by examining whether mean errors were significantly different from zero.

Time Series Design. Care was taken to ensure that all observations of forecast errors were statistically independent. Satisfying this need for independence required constructing a specially designed time series for the following reason: Since the forecast spans for all implied volatilities from the same maturity class overlap, their forecast errors are correlated. For instance, note that on any given day, the forecast span of an implied volatility is the time interval from that date out to the expiration date of the option. An implied volatility derived from an option of the same maturity class on a subsequent day will forecast over the shorter span from that later day out to the same expiration date. Thus, implied volatilities from the same maturity class taken on different days project across overlapping time intervals. Since the spans include common days, the forecasting errors in implied volatilities from the same maturity class are correlated, even when the forecasts are made on different dates. An example will further explain this problem: The forecast error in the implied volatility derived from the June 1983 contract on April 1, 1983, is correlated with the forecast error in the implied volatility derived from the June 1983 contract on April 30, 1983; both forecast spans include May 1983, and any surprises in actual volatility during May 1983 will appear in both forecast errors. The overlap of forecast spans is illustrated in Chart 1.

To avoid this problem, one date was chosen from the life of each maturity class. Observations on these dates alone made up the time series of forecasts. The exact dates were selected so that no forecast horizons would overlap and all forecast spans would be equal in length.[11] Specifically, from each maturity class, the implied volatility from the 57th trading day prior to option expiration was included. For example, the implied volatility from the June 1984 class was sampled on March 26, 1984, when the June contracts were 57 trading days from maturity. This procedure resulted in a sample of implied volatilities on the S&P 500 futures such that each observation represented a 57-day forecast with no overlap of forecast spans.[12]

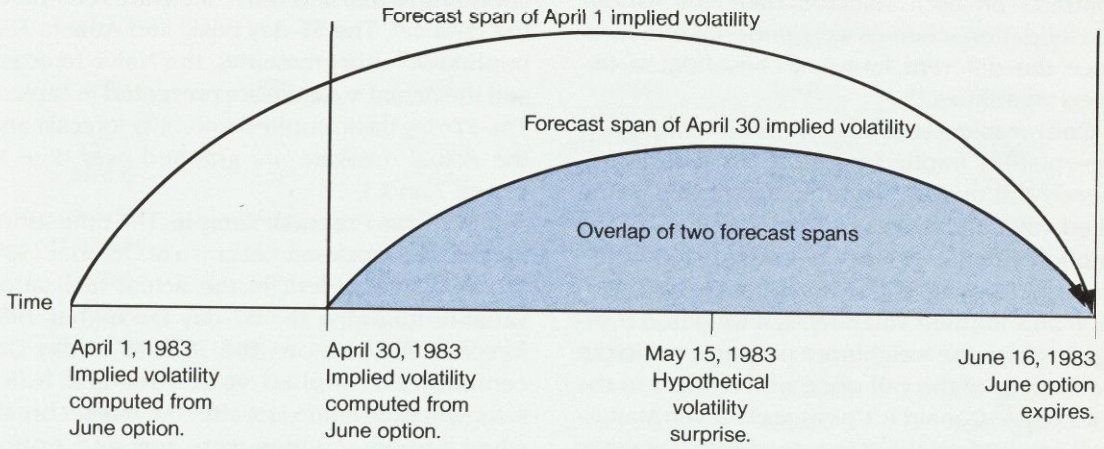
Standard interval lengths shorter than 57 days can also produce time series free of overlapping forecasts. To test how implied volatilities performed over shorter horizons, this study constructed two such series of briefer forecasts. These additional time series forecast over 38- and 19-day horizons. Thus, the necessary data were available to test how well implied volatilities based on stock-index options forecast stock-market volatility over 57, 38, and 19 trading days.

The Data

Options on S&P 500 Index Futures. The empirical tests were conducted on implied volatilities from options on S&P 500 index futures. The S&P 500 is a broad-based stock-market index that generally represents the stock market as a whole. Another advantage of this particular instrument is that the S&P 500 index future and its respective option are popular and actively traded. Consequently, the end-of-day settlement prices, which are easily obtained from daily newspapers and which are used in this study, almost always represent actual trades that took place simultaneously at the market's close. This simultaneity is necessary to compute implied volatilities precisely.

Furthermore, since the index futures option is traded as a single instrument, it avoids the problem of nonsynchronization that occurs when using stock index options. (*Nonsynchronization* refers to the fact that a stock index incorporates the last traded prices of the stocks in the index, which are not necessarily the result of recent trades.) Investors have access to

Chart 1.
An Example of the Correlation of Forecast Errors



Suppose the implied volatilities forecast on the two dates in April are perfect forecasts with the exception that neither predicts a surprise on May 15, 1983. This surprise will appear in both measures' forecast errors. These errors would thus be correlated.

all information about the underlying stocks, even news affecting infrequently traded equities, when pricing the index future. Thus, index futures are apt to reflect all available information, whereas the quoted index might not.[13] Additionally, although anticipated dividends must be subtracted from stock prices in order to price options on stocks, this correction is not needed when pricing options on stock futures. Futures pay no dividends, and so the market's pricing of the future has already, in effect, made the adjustment.

To derive implied volatilities from options on futures, one must use a slightly different option-pricing formula. The pricing formula for options on futures is quite similar to the formula for pricing options on stocks. The only difference is that the discounted future price takes the place of the stock price (Black [1976]).[14]

By construction, then, the implied volatilities examined in this study were forecasts of the standard deviations of the percent changes in the discounted S&P 500 index futures. These standard deviations are closely related to the volatility of the S&P 500 index itself since the discounted index futures price must equal the dividend-corrected, full-information index price in order to preclude arbitrage possibilities.

The Raw Data. The Chicago Mercantile Exchange provided settlement prices of the futures and calls, as well as each option's strike price and expiration date. For the risk-free interest rate this study incorporated the annualized yield converted from the average of the bid and asked discounts of the particular Treasury bill that would be the first to mature after each respective option. Thus, the term of the risk-free rates coincides closely with that of the options.

For each of the 3 forecast series, 23 separate implied volatility observations were constructed, one from each of the maturity classes between June 1983 and December 1988. This study excluded the March 1983 contract, the first S&P 500 index futures option marketed, because trading on it began with less than 57 trading days to maturity.

Construction of the Actual Realized Standard Deviations. In order to assess the forecast accuracy of each implied volatility, the actual fluctuation in the S&P 500 futures subsequent to each forecast was measured. Actual volatilities were determined by computing the sample standard deviations of percent changes in discounted S&P 500 future prices.[15]

Construction of the Implied Volatilities. Each implied volatility was computed via an iterative

search procedure that located the volatility value in the pricing formula that is consistent with the values of the observable variables. After implied volatilities were computed for all contracts on each selected date, the various consolidation schemes were undertaken to produce the different forms of consolidated implied volatilities.[16]

One measure—the basic implied volatility—is the volatility implied from the just-out-of-the-money call option.[17] Another, the average implied volatility, is the arithmetic average of the implied volatilities from all call options of the same maturity class.[18] A third, Latane and Rendleman's implied volatility, is a weighted average in which the weights are proportional to the derivatives of the call price with respect to the volatility.[19] Donald P. Chiras and Steven Manaster's implied volatility is a weighted average in which the weights are proportional to the elasticities instead of the derivatives.[20] The put implied volatility is derived from the just-out-of-the-money put option. The put-and-call implied volatility is the average of the implied volatilities from the just-out-of-the-money call option and the just-out-of-the-money put option.[21] Finally, the Atlanta Fed implied volatility is a new measure: the weighted intertemporal average of implied volatilities from just-out-of-the-money call options. (This last measure is described in greater detail in Appendix 1.)

All implied volatilities were computed to be forecasts of per-trading-day standard deviations; this method differs from previous studies that constructed implied volatilities based on calendar days. A sample standard deviation of stock returns is a measure of standard deviation per trading day, and so to be comparable the resulting implied volatility should be per trading day as well.

The Naive Forecast—A Standard for Comparison. A conventional history-based forecast of volatility, referred to in this article as the Naive forecast, was also constructed for purposes of comparison. This measure is the sample standard deviation of discounted futures returns over the 20 days up to and including the date of each implied-volatility observation. *Actual* is the realization variable—the actual sample standard deviation of discounted stock futures returns over the respective forecast intervals.

The full data set contained a series of 23 observations of each of the implied-volatility forms and for each of the three forecast horizons described above: 57, 38, and 19 days. Corresponding Actual and Naive measures complete the data set. The 57-day basic and Atlanta Fed implied-volatility forecasts, the Naive forecast, and the Actual measure are presented in Table 1. The 57-day basic implied-volatility forecast and the Actual measure are graphed over time in Charts 2 and 3.

Full versus Precrash Sample. The time series include the stock-market crash of October 1987. The crash is evident in the actual realization variable following the 57-day December 1987 forecast. Furthermore, the 38- and 19-day December 1987 implied volatilities and Naive forecasts were made just after the market break, when option volumes were low and option prices were much higher than they had ever been. Table 1 indicates that implied volatilities for the 57-day forecast were indeed much higher following the crash than before, but they gradually moderated from those extreme levels. So that the crash's effects on the measures could be analyzed, all of the tests were performed on the entire data sample and then again on just the precrash sample, that is, the sample period preceding but not including the December 1987 contract. Means and standard deviations of all measures are presented in Tables 2 and 3. The statistics are given for both the entire sample period and for the precrash sample.

Results of the Tests

How the Various Implied Volatilities Performed. Tables 2 and 3 reveal that the various consolidation methods produce similar forecasts. Yet the mean squared forecast errors presented in Table 4, which show the efficiency of the different projections, reveal a definite pattern. The Atlanta Fed intertemporal average exhibited smaller mean squared errors—indicating greater accuracy—among the 57- and 38-day forecasts, and this result held for both the full sample and for the precrash sample.

The results were different among the 19-day forecasts. The basic as well as the put-and-call implied volatilities outperformed that of the

Table 1.
Selected 57-Day Implied-Volatility Forecasts and Realizations
(standard deviation of daily percent returns)

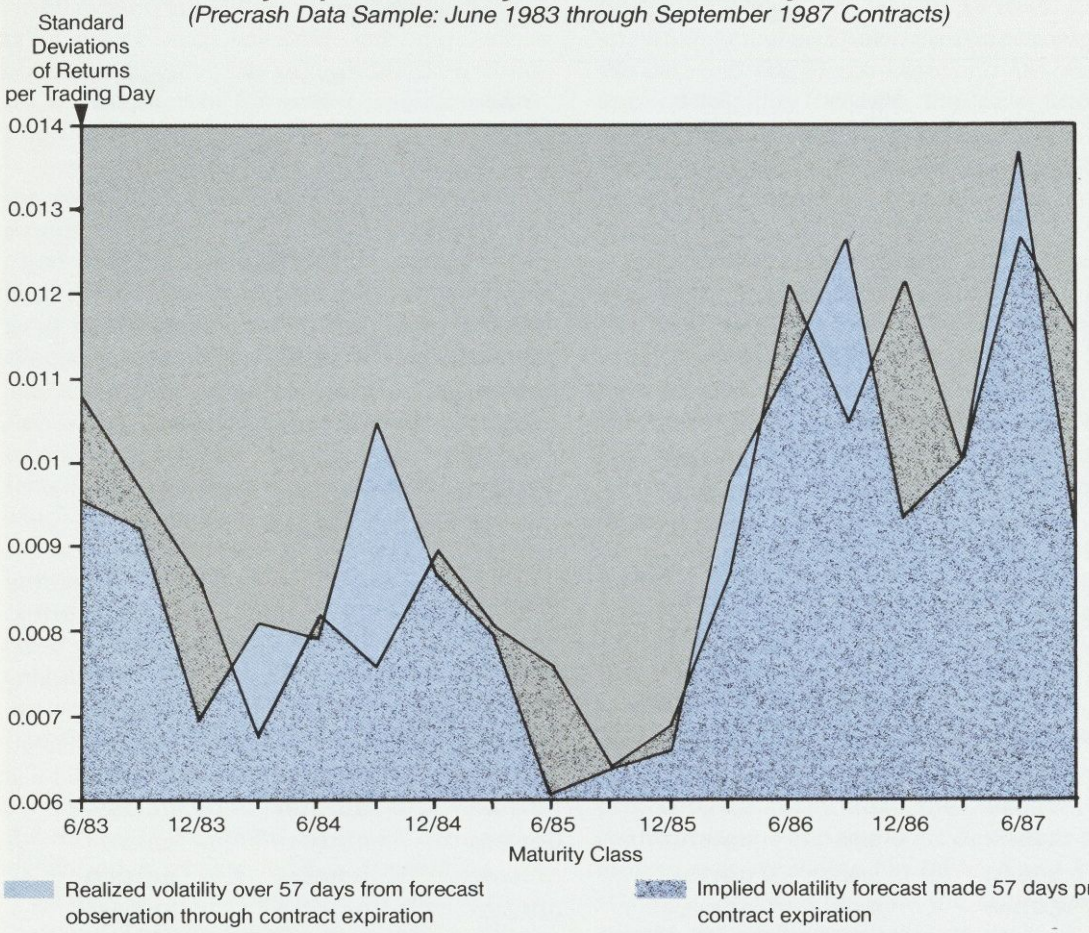
Date	Maturity Class	Actual Volatility	Naive Forecast	Basic Implied Volatility	Atlanta Fed Implied Volatility
3/25/83	6/83	0.009499	0.011625	0.010791	0.010670
6/24/83	9/83	0.009158	0.010762	0.009684	0.009325
9/23/83	12/83	0.006917	0.006858	0.008598	0.008647
12/22/83	3/84	0.008069	0.007162	0.006768	0.006646
3/26/84	6/84	0.007894	0.007256	0.008194	0.008242
7/2/84	9/84	0.010431	0.007442	0.007562	0.007672
10/2/84	12/84	0.008636	0.009235	0.008943	0.008920
12/21/84	3/85	0.007875	0.009512	0.008035	0.008239
4/1/85	6/85	0.006050	0.006548	0.007570	0.007557
7/1/85	9/85	0.006344	0.005609	0.006394	0.006371
10/1/85	12/85	0.006583	0.008745	0.006862	0.006852
12/30/85	3/86	0.009730	0.007854	0.008712	0.009400
4/1/86	6/86	0.011083	0.009217	0.012115	0.012186
6/30/86	9/86	0.012601	0.010009	0.010401	0.010393
9/30/86	12/86	0.009273	0.015980	0.012199	0.011917
12/29/86	3/87	0.009985	0.007493	0.009912	0.009863
3/30/87	6/87	0.013617	0.010685	0.012707	0.012358
6/29/87	9/87	0.009013	0.006973	0.011430	0.011508
9/29/87	12/87	0.057259	0.012423	0.012459	0.012589
12/28/87	3/88	0.016797	0.018950	0.019941	0.019174
3/28/88	6/88	0.012347	0.010582	0.016645	0.015559
6/27/88	9/88	0.010308	0.012757	0.013649	0.012955
9/27/88	12/88	0.007728	0.007137	0.010946	0.010901

Source: Computed at the Federal Reserve Bank of Atlanta from data provided by the Chicago Mercantile Exchange.

Atlanta Fed over the full sample, and the latter displayed the highest mean squared error over the precrash sample. Among the 19-day implied-volatility forecasts, the basic form showed minimum mean squared errors over both samples, and understandably so. Relative to the shorter forecast horizon, the error introduced in the other samples by averaging implied volatilities that were not up-to-date is greater than any errors in the basic format stemming from errors in variables. The implication from the 19-day forecast results is clear: for longer-term forecasts intertemporal averaging of implied volatilities is helpful, but for shorter forecasts the lone current implied volatility from the just-out-of-the-money option dominates. Averaging among the multiple contracts with different strike prices appears to provide no benefit over simply selecting the just-out-of-the-money option.

Comparisons between the Implied Volatilities and the Naive Forecast. Table 4 also compares the various implied volatilities and the historically based forecast. The 57- and 38-day implied volatilities outperformed the Naive forecast over both the full and precrash samples, and the 19-day implied volatilities did better than the Naive forecast over the full sam-

Chart 2.
57-Day Implied-Volatility Forecast vs. 57-Day Realization
(Precrash Data Sample: June 1983 through September 1987 Contracts)



In most periods prior to the crash, the basic implied-volatility forecast generally appears to have anticipated stock-market volatility shifts.

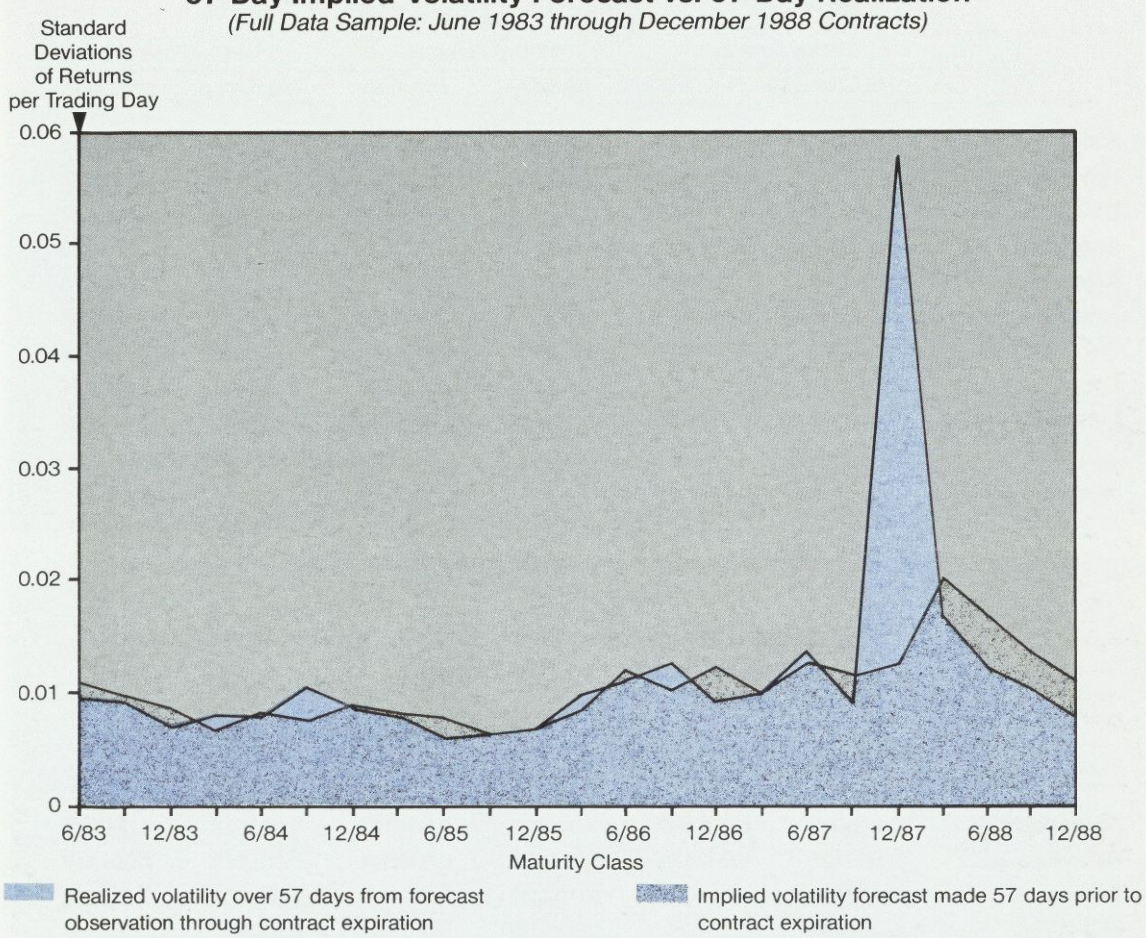
Source: See Table 1.

ple. Only in the precrash period, among the 19-day forecasts, did the Naive forecast outperform the group of implied volatilities.

There are several possible explanations for the failure of the implied-volatility forecasts among the 19-day projections in the precrash period. Perhaps the possibility of a large market shift has always been accounted for in the determination of implied volatilities. When such a break did occur in October 1987, this consideration was validated and the implied volatilities outperformed the Naive forecast. Markets seem to have been anticipating a crash, in the form of a "crash premium" in options prices, which produced large positive forecast errors. This scenario, however, does not adequately explain why the Naive forecast outperformed the implied volatilities only over the short horizon.

Another explanation might be that the Naive predictor functions well over short horizons in normal circumstances, causing it to outperform the implied volatilities among the 19-day forecasts in the precrash sample. After the crash, however, the Naive forecast could not reliably

Chart 3.
57-Day Implied-Volatility Forecast vs. 57-Day Realization
(Full Data Sample: June 1983 through December 1988 Contracts)



The volatility in the period during which the crash occurred was of a much higher magnitude than previous volatility or forecasts. Following the crash, the implied volatility correctly forecast falling volatility.

Source: See Table 1.

predict a return to normality. Other possible explanations exist, and future research will be directed at addressing this phenomenon. The matter is primarily of academic interest, though. The crash occurred, and over *all* available data the implied-volatility forecasts of long, medium, and short lengths performed better on the basis of mean squared error than did the Naive alternative.

Mean Absolute Errors. The mean squared error just discussed assigns greater weight to larger errors than to smaller ones. Consequently, over a period such as 1983-88 in which one enormous market shift transpired, the rankings could be based primarily on how well each forecast performed near the crash. To test that possibility, mean absolute errors were computed, and the forecasts were ranked accordingly. The mean absolute error statistic, which is the mean of the absolute values of the errors, weighs all errors equally. Table 5 presents these statistics and the rankings.

The rankings for the long- and medium-range forecasts are virtually the same as those pre-

Table 2.
Means of Implied-Volatility Forecasts and Realizations

	57-Day Forecasts		38-Day Forecasts		19-Day Forecasts	
	All Data	Precrash	All Data	Precrash	All Data	Precrash
Actual Volatility	0.011617	0.009042	0.009431	0.008727	0.008965	0.008451
Naive Forecast	0.009601	0.008831	0.013853	0.009326	0.009779	0.008664
Basic	0.010457	0.009270	0.012349	0.009388	0.010405	0.009197
Atlanta Fed	0.010345	0.009265	0.012036	0.009342	0.010601	0.009363
Average	0.010596	0.009384	0.012482	0.009554	0.010682	0.009415
Latane-Rendleman	0.010615	0.009410	0.012536	0.009628	0.010736	0.009505
Chiras-Manaster	0.010608	0.009346	0.012504	0.009488	0.010845	0.009424
Put	0.010621	0.009364	0.012555	0.009503	0.010684	0.009328
Put-and-Call	0.010539	0.009317	0.012452	0.009445	0.010545	0.009262

Source: See Table 1.

Table 3.
Standard Deviations of Implied-Volatility Forecasts and Realizations

	57-Day Forecasts		38-Day Forecasts		19-Day Forecasts	
	All Data	Precrash	All Data	Precrash	All Data	Precrash
Actual Volatility	0.010040	0.001998	0.003422	0.001827	0.003073	0.002271
Naive Forecast	0.003121	0.002370	0.016701	0.003009	0.005253	0.001747
Basic	0.003196	0.001929	0.010236	0.002323	0.003874	0.001889
Atlanta Fed	0.002957	0.001869	0.008700	0.002168	0.003970	0.002133
Average	0.003184	0.001904	0.010115	0.002325	0.003973	0.001854
Latane-Rendleman	0.003172	0.001897	0.010100	0.002285	0.003947	0.001853
Chiras-Manaster	0.003266	0.001924	0.010216	0.002407	0.004160	0.001917
Put	0.003343	0.001982	0.010246	0.002385	0.004198	0.001905
Put-and-Call	0.003267	0.001952	0.010239	0.002349	0.003983	0.001891

Source: See Table 1.

viously stated, with the set of implied volatilities besting the Naive forecast in both the full and precrash samples. Among the short-range forecasts, however, the rankings showed an important difference: The Naive predictor showed the lowest mean absolute error in both the full sample and the precrash sample. Since, relative to mean squared error, the mean absolute error statistic deemphasizes performance of the predictors around the time of the October 1987 crash, the difference between the two rankings would seem to pinpoint the crash as the time when the historically based forecast suffered relative to the implied volatilities, thus adding support to the hypothesis that the Naive forecast is effective over short horizons in calm conditions but will fail in a more turbulent environment.

Which consolidated implied-volatility form worked best on the basis of mean absolute

Table 4.
Mean Squared Errors and Rankings for Implied-Volatility Forecasts

57-Day Forecasts			
All Data		Precrash	
Atlanta Fed	0.0000898366	Atlanta Fed	0.0000020959
Basic	0.0000911502	Basic	0.0000021930
Average	0.0000913429	Chiras-Manaster	0.0000022796
Latane-Rendleman	0.0000913626	Average	0.0000023230
Put-and-Call	0.0000915179	Latane-Rendleman	0.0000023640
Chiras-Manaster	0.0000916063	Put-and-Call	0.0000024107
Put	0.0000919392	Put	0.0000026653
Naive Forecast	0.0000925160	Naive Forecast	0.0000057465
38-Day Forecasts			
All Data		Precrash	
Atlanta Fed	0.0000411145	Atlanta Fed	0.0000037958
Average	0.0000613298	Basic	0.0000045031
Latane-Rendleman	0.0000613702	Put-and-Call	0.0000045916
Basic	0.0000622101	Latane-Rendleman	0.0000046010
Put-and-Call	0.0000628733	Average	0.0000046580
Chiras-Manaster	0.0000629718	Put	0.0000047388
Put	0.0000636279	Chiras-Manaster	0.0000048418
Naive Forecast	0.0002089594	Naive Forecast	0.0000074591
19-Day Forecasts			
All Data		Precrash	
Basic	0.0000091484	Naive Forecast	0.0000054929
Put-and-Call	0.0000100337	Basic	0.0000062446
Atlanta Fed	0.0000102795	Put-and-Call	0.0000063854
Average	0.0000104832	Average	0.0000065595
Latane-Rendleman	0.0000105135	Put	0.0000065824
Put	0.0000110382	Chiras-Manaster	0.0000066718
Chiras-Manaster	0.0000120698	Latane-Rendleman	0.0000067284
Naive Forecast	0.0000133162	Atlanta Fed	0.0000070297

Source: See Table 1.

error? Again, the rankings among the implied volatilities were topped by measures constructed exclusively from at-the-money call options. The Atlanta Fed implied volatility was the most accurate forecast among the 57- and 38-day forecasts. The basic implied volatility gave the best short-term implied-volatility forecast.

Tests for Bias. Tests were performed to determine if the mean errors were significantly different from zero. A significantly positive mean error would indicate that the forecast was biased upward, predicting too great a volatility. Significantly negative mean error would indicate a downward bias and an underestimation

Table 5.
Mean Absolute Errors and Rankings for Implied-Volatility Forecasts

57-Day Forecasts			
All Data		Precrash	
Atlanta Fed	0.003317	Atlanta Fed	0.001122
Basic	0.003463	Basic	0.001158
Put-and-Call	0.003556	Chiras-Manaster	0.001210
Average	0.003576	Put-and-Call	0.001223
Chiras-Manaster	0.003590	Average	0.001246
Latane-Rendleman	0.003592	Latane-Rendleman	0.001269
Put	0.003658	Put	0.001299
Naive Forecast	0.003750	Naive Forecast	0.001914
38-Day Forecasts			
All Data		Precrash	
Atlanta Fed	0.003275	Atlanta Fed	0.001471
Basic	0.003617	Basic	0.001554
Put-and-Call	0.003712	Put-and-Call	0.001601
Average	0.003715	Put	0.001648
Latane-Rendleman	0.003719	Chiras-Manaster	0.001673
Chiras-Manaster	0.003787	Average	0.001675
Put	0.003806	Latane-Rendleman	0.001686
Naive Forecast	0.005504	Naive Forecast	0.001982
19-Day Forecasts			
All Data		Precrash	
Naive Forecast	0.002274	Naive Forecast	0.001820
Basic	0.002305	Basic	0.001851
Put-and-Call	0.002426	Put-and-Call	0.001892
Average	0.002469	Average	0.001924
Atlanta Fed	0.002490	Put	0.001934
Latane-Rendleman	0.002511	Chiras-Manaster	0.001949
Put	0.002547	Latane-Rendleman	0.002000
Chiras-Manaster	0.002644	Atlanta Fed	0.002004

Source: See Table 1.

of future volatility. Table 6 presents the results of the tests.

The 57-day forecasts showed no significant bias, and most of the 38- and 19-day implied volatilities did not exhibit bias prior to the crash. Over the entire sample period, however, the implied volatilities did indeed show significant or nearly significant bias, indicating that the implied-volatility forecasts moved up too much when they rose after the crash. The full-sample mean forecast errors of the 38-day implied volatilities were positive and approximately four times greater than the mean errors before the crash. The mean error in the full sample for

Table 6.
Results of t-tests on Implied-Volatility Forecasts

	Mean Errors		t-statistics of Hypothesis that Mean Error = 0	
	57-Day Forecasts		57-Day Forecasts	
	All Data	Precrash	All Data	Precrash
Naive Forecast	-0.00201	-0.000210	-1.00573	-0.404260
Basic	-0.00116	0.000228	-0.57436	0.715403
Average	-0.00102	0.000342	-0.50399	1.056126
Latane-Rendleman	-0.00100	0.000367	-0.49476	1.128814
Chiras-Manaster	-0.00100	0.000303	-0.49714	0.941579
Put	-0.00099	0.000321	-0.49008	0.920722
Put-and-Call	-0.00107	0.000275	-0.53214	0.824701
Atlanta Fed	-0.00127	0.000222	-0.63513	0.660540
	38-Day Forecasts		38-Day Forecasts	
	All Data	Precrash	All Data	Precrash
Naive Forecast	0.004421	0.000598	1.506796	1.028655
Basic	0.002917	0.000660	1.867351	1.499926
Average	0.003050	0.000826	1.984053	1.900026
Latane-Rendleman	0.003104	0.000900	2.024598	2.120899
Chiras-Manaster	0.003072	0.000760	1.970059	1.687388
Put	0.003123	0.000775	1.995856	1.746285
Put-and-Call	0.003020	0.000717	1.932285	1.628309
Atlanta Fed	0.002605	0.000614	2.085444	1.410572
	19-Day Forecasts		19-Day Forecasts	
	All Data	Precrash	All Data	Precrash
Naive Forecast	0.000814	0.000212	1.073707	0.417751
Basic	0.001440	0.000746	2.539720	1.433457
Average	0.001717	0.000964	2.934312	1.861818
Latane-Rendleman	0.001771	0.001054	3.058654	2.038533
Chiras-Manaster	0.001880	0.000973	3.019127	1.863861
Put	0.001719	0.000876	2.837983	1.666446
Put-and-Call	0.001580	0.000811	2.699463	1.553787
Atlanta Fed	0.001636	0.000912	2.783363	1.554338

Source: See Table 1.

the 19-day implied volatilities approximately doubled that in the precrash sample. Perhaps thinner trading in options following the crash led to mispricing, or maybe option prices rose to account for the possibility of a repeat market break that never materialized. Only time will tell whether this systematic positive bias is a persistent phenomenon or one confined to the aftermath of the crash.

Note that the Naive predictor never exhibited significant bias, but this outcome could be more a result of greater dispersion among the Naive

predictor's forecast errors than a result of on-target forecasting. In both cases where the Naive forecast appeared unbiased and the implied volatilities appeared biased, errors in the historically based forecast exhibited far greater dispersion.

Using Implied Volatilities to Predict Volatility Changes

Although 38- and 19-day implied-volatility forecasts appear to be biased predictors of future swings in the market, these measures served quite well as harbingers of the *direction* of change in stock-market volatility. Whereas the Naive forecast assumes that the future will mirror the past, an implied-volatility forecast can be applied to predict whether future market performance will be more or less turbulent than in the past. A test was devised to see how useful the implied volatilities were for this purpose.

For this test, a drop in volatility is defined as occurring when a forecast span exhibits lower volatility than did the 20 days prior to the forecast date. A rise takes place when the span displays greater volatility than did the previous 20 days. The ratio of the realized or actual standard deviation to the Naive forecast will then indicate whether the period experienced a rise or a fall. If the former divided by the latter is greater than one, volatility has climbed; if the ratio is less than one, volatility has declined. In addition, the ratio of the implied volatility to the Naive forecast indicates whether a fall or a rise was predicted. If the implied volatility divided by the Naive forecast is greater than one, a forecast of greater volatility is issued. If that ratio is less than one, a fall in volatility is predicted.[22]

The 23 57-day spans in the full data sample contained 14 upswings and 9 downturns in volatility. The Atlanta Fed implied volatility correctly predicted the direction of these changes 19 times.[23] Six of the 9 falls and 13 of the 14 rises were correctly predicted, a success rate unlikely to have happened completely by chance. The probability of this outcome would be only 0.1 percent if the predictor were equally likely to succeed or fail. Around the 38-day forecasts, 16 falls and 7 rises took place. The Atlanta Fed implied volatility correctly predicted 17 of these changes—11 falls and 6 rises, a result that had only a 2.9 percent probability of occurring strictly by chance. For the 19-day forecasts, 11 falls and 12 rises took place; the Atlanta Fed implied volatility correctly predicted 17 of the 23 changes, 6 falls and 11 rises.

The indicator succeeded in forecasting a greater percentage of the rises than the falls, but the implied volatilities were more likely to be correct when predicting a fall than when predicting a rise, as can be seen by studying the results from a different perspective. The 57-day implied volatilities predicted rises 16 times, and 13 of these forecasts turned out to be correct. On the other hand, six of the seven "fall" predictions turned out to be accurate. The 38-day implied volatilities predicted falls 12 times, and on 11 occasions falls actually occurred. Among the 11 predictions of rises, only 6 were correct. The 19-day implied volatilities predicted rises 16 times, 11 times correctly. Six of the seven predictions of falls were accurate.

Thus, the implied volatility "fall" forecasts were correct more often than the "rise" forecasts, an effect that is plausibly related to the positive bias detected in the 38-day and 19-day forecasts. If forecasts are biased upward, they will predict upturns too often and fail to predict some declines. When a fall is projected, however, the forecast reliability is great. Overall, though, the implied volatilities were useful in predicting changes in stock-market volatility, succeeding more than two-thirds of the time. This result cannot be wholly attributed to positive bias coupled with a rising trend in volatility, since nearly as many falls as rises took place over the period.

Conclusions and Suggestions for Future Research

This study demonstrates the usefulness of implied volatilities in forecasting stock-market vicissitudes. When forecasting 57 and 38 days into the future, implied volatilities were more efficient than the history-based sample of standard deviation of past returns (the Naive forecast). Among 19-day forecasts, however, the Naive forecast appeared to do better than the implied volatilities before the stock market

crash of 1987, although the implied volatilities outperformed the Naive forecast after the crash.

Apparently, the Naive predictor functions best over short horizons when the market is calm. This result is quite reasonable in light of the Naive predictor's design, which is based on the past. One has no way of knowing ahead of time, however, whether calmness will persist. Thus, the implied volatility is a necessary adjunct even when forecasting over very short horizons. Moreover, the Naive predictor is least reliable when volatility is changing, and this is just when a reliable predictor is needed most.

The Atlanta Fed implied volatility was the most efficient form of implied-volatility forecast when projecting over 57- and 38-day horizons, while the basic implied volatility was the most efficient among the 19-day implied-volatility forecasts. Remember that both the Atlanta Fed and the basic implied volatilities are constructed exclusively from just-out-of-the-money options, providing empirical support for the proposition that such options offer the most reliable implied-volatility forecasts. The forms that used away-from-the-money options fared worst, as predicted.

No statistically significant bias could be detected in the 57-day implied-volatility forecasts. The same can be said for the 38-day and 19-day forecasts prior to the 1987 crash with the one exception of the Latane and Rendleman implied-volatility form. In the data sample that included the crash, however, the implied-volatility forecasts displayed significant positive bias. Apparently, since the crash, implied-volatility forecasts have been too high relative to subsequently realized actual stock-market volatility.

This bias may be due to a flaw in the option-pricing model, or it may reflect mistakes made by options market investors after the crash. Even in view of this result, implied volatilities might still have been the best forecasts attainable with all then-available information, and, furthermore, they might have been statistically unbiased in an a priori sense. Those postcrash implied volatilities may have correctly incorporated the possibility of another stock-market crash, yet the implied volatilities now appear to have been biased because that scenario did not happen.[24] There is no way to know. Whether the short- and medium-term implied volatilities will continue to appear biased can only be determined with continued monitoring and testing of future data.

The final test in this study demonstrated how an implied-volatility measure can be used as a leading indicator of stock-market volatility changes. The 57-day Atlanta Fed implied-volatility indicator was quite successful in predicting whether stock-market volatility would rise or fall.

Aside from the direct uses discussed in this study, implied volatility has a variety of other applications for both investors and policymakers. Since active trading now takes place in futures and options on fixed-income securities and on foreign currencies, in addition to stock indexes, implied volatilities from those options can be used to forecast the volatility of bond prices and exchange rates. Future testing should include these applications. Implied volatilities can also provide investors with up-to-date measures of the riskiness in various investment markets.

Also, policymakers can use these indicators to assess market uncertainty and to plan policy changes or announcements accordingly. Authorities can determine whether new policy has calmed or unsettled the markets. This information cannot be inferred from a price response alone since a large price swing might simply reflect adjustment to a new, stable level. Similarly, a small price response might disguise considerable uncertainty and a lack of consensus about equilibrium prices. The implied volatility, on the other hand, will indicate the new level of uncertainty in the form of traders' forecasts of future price stability.

The basic and Atlanta Fed implied-volatility measures will be maintained and continually updated by the Atlanta Fed and are available upon request from the author.

Appendix 1.
The Design of the Atlanta Fed Stock-Market Volatility Indicator

The Atlanta Fed implied-volatility indicator is a weighted average of current and lagged implied volatilities from just-out-of-the-money call options. The implied volatility from the one just-out-of-the-money call option on the current day is averaged together with the one implied volatility from the just-out-of-the-money option from each of the previous four days.

Averaging several closely related implied volatilities makes the resultant measure less sensitive to errors in measuring variables or to any temporary price anomalies. Including away-from-the-money options, however, should theoretically contaminate the resultant measure with systematic bias. (See Butler and Schachter [1986] and Feinstein [1988] for more on this topic.) Intertemporal averaging offers the benefit of averaging without having to rely on away-from-the-money options.

Intertemporal averaging has a drawback, nonetheless. Lagged implied volatilities are informationally deficient, that is, they contain a few days less information than current implied volatilities. Moreover, lagged implied volatilities forecast over slightly different spans. A lagged implied volatility's forecast span includes its immediate future, a few days that are not in the forecast span of the most current implied volatility. If the days following the lagged implied volatilities are anticipated to be unlike those in the remainder of the forecast span, this mismatch in spans could introduce noise into the averaged measure.

The problems with intertemporal averaging can be minimized by proper design of the averaging weights. Since lagged implied volatilities are as likely as the current implied volatility to suffer from errors in the raw variables, they should be given—at most—equal weight in the averaging. But, because lagged implied volatilities are deficient with respect to the information they contain, they reasonably should be given less weight and those weights should decline with the length of the lag.

A five-day information deficiency is more serious relative to a short forecast than to a long forecast. Similarly, a five-day mismatch of forecast spans is more serious when the forecast horizon is near than when it is distant. Thus, the pattern of weights should be related to the length of the forecast span. The weights should decline gradually with lag length when the horizon is distant, but for a nearby horizon the decline should be steep. For example, in constructing a 19-day implied volatility forecast, the average should place the greatest weight on the most current implied volatility, but in constructing a 57-day forecast the average should be more equally weighted.

The following weighting scheme, which has these properties, was developed and used to construct the Atlanta Fed implied volatility:

$$\omega_L = \frac{\left(\frac{\Upsilon}{60}\right)^L}{\sum_{i=0}^{4} \left(\frac{\Upsilon}{60}\right)^i},$$

where ω_L is the weight assigned to the lagged implied volatility; L is the lag length, ranging from zero to four; and Υ is the length of the forecast span, that is, the number of trading days between the forecast date and option expiration. The weights ω_L decline geometrically with lag length L, and the base factor varies with forecast span Υ. The weights are normalized to sum to one.

The formula for the Atlanta Fed implied volatility is thus given by:

$$\sum_{L=0}^{4} \omega_L \cdot IV_L,$$

where IV_L is the implied volatility from the just-out-of-the-money option observed on the Lth day prior to the current day.

Notes

[1]See Feinstein (1987) for definitions and descriptions of stock-market volatility and a discussion of its costs.

[2]See, for example, Officer (1973) or Feinstein (1987).

[3]See Black (1976b) regarding the effects of recent stock returns and firm leverage; Pattel and Wolfson (1979, 1981) regarding the effects of announcements; Stoll and Whaley (1986) and Feinstein and Goetzmann (1988) regarding option and future expiration effects; Fama (1965) and French (1980) regarding holidays and trading hours; and Hardouvelis (1988) regarding margin requirements.

[4]The reader unfamiliar with option fundamentals can find the necessary introduction in Feinstein and Goetzmann (1988).

[5]A forecast is unbiased in the statistical sense if it is expected to be correct on average, which is a rather nice property to expect from a forecast. If, over time, a forecast method consistently produced forecasts that erred by some nonzero amount, that method's projections would be deemed biased. The "bias" would be the quantity, either negative or positive, by which the forecasts usually missed. No forecast method, however, can be counted on to produce perfect forecasts every time. Efficiency, in the statistical sense, measures the size of a forecasts' errors. A more efficient forecast usually has smaller errors.

Unbiasedness and efficiency are desirable qualities in a forecast, but sometimes there is a trade-off between the two. A biased forecast that is very efficient is often better than an unbiased forecast which is inefficient. That is, a method that produces forecasts which are systematically too high or too low (biased) but quite close to the correct value (efficient) may be preferable to a method that produces forecasts that are on average correct but usually off by a large magnitude. This article describes the bias of the implied volatilities but ranks the measures by efficiency. Further discussion on optimal choice of a forecast is beyond the scope of this paper.

[6]Cross-sectional studies were conducted by Latane and Rendleman (1976), Chiras and Manaster (1978), Schmalensee and Trippi (1978), and Beckers (1981).

[7]Tests of the Black-Scholes model include those by Macbeth and Merville (1979).

[8]In the original formulation of their model Black and Scholes required that the stock volatility be constant and known. Merton (1973), however, showed that the volatility need not be constant and that the pricing formula is easily generalized to account for volatility that changes with time. Yet in Merton's formulation, too, the stock volatility is assumed to be nonstochastic.

[9]That is, the strike price discounted at the risk-free rate over the term of the option. The *risk-free rate* is the interest rate on risk-free bonds, such as those issued by the U.S. Treasury.

[10]Also, different implied volatilities can be derived by using option pricing models other than the Black-Scholes model. These models either specify different assumptions regarding the process that drives stock prices, or they attempt to account for some factor ignored by the Black-Scholes model. Most of these models include additional unobservable variables, and so they require ad hoc assumptions in order to produce implied-volatility forecasts. Rubinstein (1985) showed, however, that under reasonable assumptions about the additional unobservable variables, the alternative models based on different stock processes produce implied volatilities that coincide with the Black-Scholes implied volatility for at-the-money moderate-term options. Consequently, little is lost by restricting the analysis in this article to Black-Scholes implied volatilities.

The early exercise futures option-pricing model developed by Whaley (1986) has no more variables than the Black-Scholes model and thus provides implied volatilities without ad hoc assumptions. When implied volatilities constructed from this model were tested, the measure did not deviate significantly from the at-the-money Black-Scholes implied volatility and performed virtually the same in each test. Thus, Whaley's numbers are not presented here since these results can be inferred from the Black-Scholes form.

[11]This construction was necessary to avoid the heteroskedasticity attributable to the fact that implied volatilities from the same maturity class on different dates forecast over spans of varying length. The variances of the forecast errors will differ when the length of the forecast intervals differ.

[12]This length, 57 days, was the shortest interval between contract expirations and thus the longest interval length for which forecast spans did not overlap. If the interval were longer, 60 trading days for instance, the 60-day span forecast by the December 1984 implied volatility would overlap the forecast span of the March 1985 implied volatility. The March 1985 forecast would have been made three days before the December contract expired. Using 57 days, no forecast spans overlap, and so all forecast errors should be uncorrelated.

[13]This point is central in Kawaller, Koch, and Koch (1988).

[14]That is, the future price discounted at the risk-free rate over the term of the option.

[15]Note that the sample standard deviation is actually only an estimator for the "true" standard deviation parameter of the stock return process, and a biased one at that. The true volatility parameter of a stock process cannot be directly observed and so must be estimated; however, when investors speak of stock volatility, they are usually referring to the observable sample standard deviation. For the purposes of this article, the ability of the implied volatility to forecast the variable of common interest is assessed, that is, the subsequently realized sample standard deviation of stock market returns, corrected for bias. The bias correction involves premultiplying the sample standard deviation by a factor that was near one for the cases presented in this article. The formula for the correction factor is given in Cox and Rubinstein (1985): 256. Whenever *sample standard deviations* are mentioned in this article, the reader should understand that the bias-corrected version is intended.

[16]Options with volume below 50 trades or prices below 50 cents were dropped prior to consolidation. When volume is low, the future and option prices are less likely to be synchronous. When price is low, the bid-ask spread and transaction costs tend to distort the pricing.

[17]See Beckers (1981).

[18]See Schmalensee and Trippi (1978).

[19]See Latane and Rendleman (1976). This derivative is the sensitivity of the call price level to changes in the volatility.

[20]See Chiras and Manaster (1978). This elasticity is the sensitivity of percent call price changes to percent changes in the volatility.

[21]This method is close to that used by the Chicago Mercantile Exchange in the "Volatility Corner" feature in their periodical *Market Perspectives*.

[22]Frank King, associate director of research at the Atlanta Fed, originally suggested this application of the implied-volatility forecasts.

[23]To promote computational and expositional parsimony, the test was performed only on the Atlanta Fed implied-volatility form, which generally had the best results in the efficiency tests.

[24]This is an example of the well-known "peso problem" in testing for unbiasedness of forecasts. See Frankel and Meese (1987) for a good discussion of the peso problem.

References

Beckers, Stan. "Standard Deviations Implied in Option Prices as Predictors of Future Stock Price Variability." *Journal of Banking and Finance* 5 (September 1981): 363-81.

Black, Fischer. "The Pricing of Commodity Contracts." *Journal of Financial Economics* 3 (January/March 1976a): 167-79.

_____. "Studies of Stock Price Volatility Changes." *Proceeding of the American Statistical Association* (1976b): 177-81.

Black, Fischer, and Myron S. Scholes. "The Valuation of Options Contracts and a Test of Market Efficiency." *Journal of Finance* 27 (May 1972): 399-418.

Black, Fischer, and Myron S. Scholes. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81 (May/June 1973): 637-59.

Butler, J.S., and Barry Schachter. "Unbiased Estimation of the Black/Scholes Formula." *Journal of Financial Economics* 15 (March 1986): 341-57.

Chicago Mercantile Exchange. *Rules of the Chicago Mercantile Exchange.* Chicago Mercantile Exchange, 1983.

_____. *Market Perspectives.* Various issues.

Chiras, Donald P., and Steven Manaster. "The Information Content of Option Prices and a Test of Market Efficiency." *Journal of Financial Economics* 6 (June/September 1978): 213-34.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross. "The Relation Between Forward Prices and Futures Prices." *Journal of Financial Economics* 9 (December 1981): 321-46.

Cox, John C., and Mark Rubinstein. *Option Markets.* Englewood Cliffs, N.J.: Prentice-Hall, 1985.

Fama, Eugene F. "The Behavior of Stock Market Prices." *Journal of Business* 38 (January 1965): 34-109.

Feinstein, Steven P. "Stock Market Volatility." Federal Reserve Bank of Atlanta *Economic Review* 72 (November/December 1987): 42-47.

_____. "A Source of Unbiased Implied Volatility Forecasts." Federal Reserve Bank of Atlanta Working Paper 88-9 (December 1988).

Feinstein, Steven P., and William N. Goetzmann. "The Effect of the 'Triple Witching Hour' on Stock Market Volatility." Federal Reserve Bank of Atlanta *Economic Review* 73 (September/October 1988): 2-18.

Frankel, Jeffrey A., and Richard Meese. "Are Exchange Rates Excessively Variable?" National Bureau of Economic Research Working Paper No. 2249 (May 1987).

French, Kenneth R. "Stock Returns and the Weekend Effect." *Journal of Financial Economics* 8 (August 1980): 55-69.

Hardouvelis, Gikas. "Margin Requirements and Stock Market Volatility." Federal Reserve Bank of New York *Quarterly Review* 13 (Summer 1988): 80-89.

Kawaller, Ira G., Paul D. Koch, and Timothy W. Koch. "The Relationship between the S&P 500 Index and S&P 500 Index Futures Prices." Federal Reserve Bank of Atlanta *Economic Review* 73 (May/June 1988): 2-10.

Latane, Henry A., and Richard J. Rendleman, Jr. "Standard Deviations of Stock Price Ratios Implied in Option Prices." *Journal of Finance* 31 (May 1976): 369-81.

Macbeth, James D., and Larry J. Merville. "An Empirical Examination of the Black-Scholes Call Option Pricing Model." *Journal of Finance* 34 (December 1979): 1173-86.

Merton, Robert C. "Theory of Rational Option Pricing." *Bell Journal of Economics and Management Science* (Spring 1973): 141-83.

Officer, Robert R. "The Variability of the Market Factor of the New York Stock Exchange." *Journal of Business* 46, no. 3 (1973): 434-53.

Patell, James M., and Mark A. Wolfson. "Anticipated Information Releases Reflected in Call Option Prices." *Journal of Accounting and Economics* 1 (August 1979): 117-40.

Patell, James M., and Mark A. Wolfson. "The *Ex Ante* and *Ex Post* Price Effects of Quarterly Earnings Announcements Reflected in Option and Stock Prices." *Journal of Accounting Research* 19, no. 2 (1981): 434-58.

Rubinstein, Mark. "Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes from August 23, 1976 through August 31, 1978." *Journal of Finance* 40 (June 1985): 455-80.

Schmalensee, Richard, and Robert R. Trippi. "Common-Stock Volatility Expectations Implied by Option Premia." *Journal of Finance* 33 (March 1978): 129-47.

Stoll, Hans R., and Robert E. Whaley. *Expiration Day Effects of Index Options and Futures.* Monograph Series in Finance and Economics, Monograph 1986-3. Salomon Brothers Center for the Study of Financial Institutions, Graduate School of Business Administration, New York University, 1986.

Whaley, Robert E. "Valuation of American Futures Options: Theory and Empirical Tests." *Journal of Finance* 41 (March 1986): 127-50.

Working Paper Series Available

The Research Department of the Federal Reserve Bank of Atlanta publishes a working paper series to stimulate professional discussion and exploration of economic subjects. We welcome readers of the *Economic Review* to complete and return the form below in order to receive our recently released working papers. If you would like a copy of any of the current papers, simply check the box next to the paper's number and return the form to the Public Information Department, Federal Reserve Bank of Atlanta.

89-1 Peter A. Abken
A Survey and Analysis of Index-Linked Certificates of Deposit

89-2 Aruna Srinivasan
Costs of Financial Intermediation under Regulation: Development Banks and Commercial Banks

Previous working papers are also available. For more information, contact the Public Information Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, NW, Atlanta, Georgia 30303-2713 (404/521-8788).

☐ Please start my subscription to the Working Paper Series.

☐ Please send me a copy of the following working papers:

☐ 89-1 ☐ 89-2

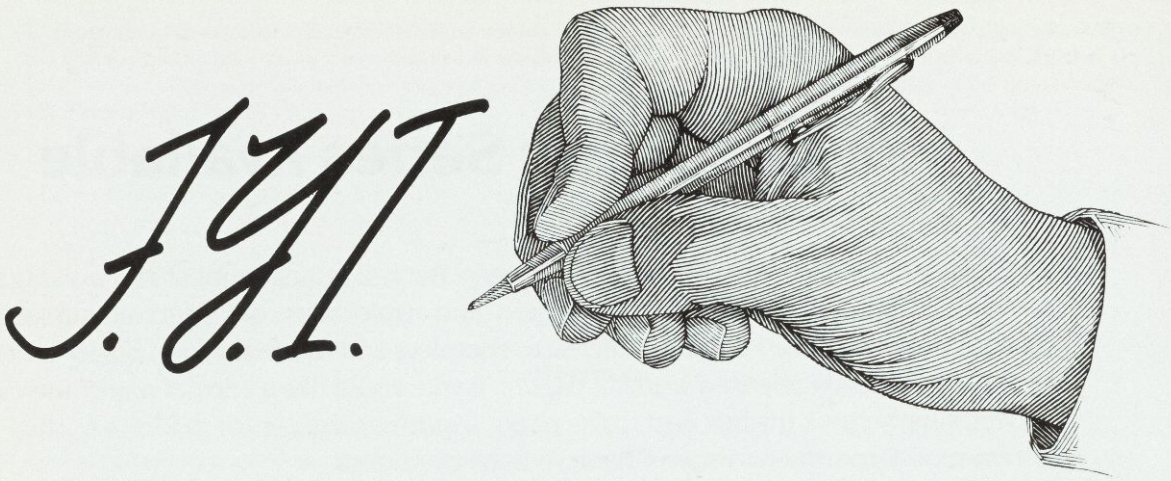
Name _____

Address _____

_____

City _____ State _____ ZIP _____

Return to the Public Information Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, NW, Atlanta, GA 30303-2713



Interstate Banking Developments in the 1980s

B. Frank King, Sheila L. Tschinkel, and David D. Whitehead

A 1983 *Economic Review* article published by this Bank depicted an extensive interstate presence by commercial banking organizations, even though state and national statutes had sought to constrain the establishment of full-service offices across state lines.[1] Banks and bank holding companies had successfully used a variety of methods to cross these boundaries, to the point where domestic and foreign banking organizations had almost 7,500 offices outside their home states. Two years later an update of the interstate situation noted that interstate banking presence was increasing and several states were enacting legislation to allow full-service interstate banking, though at the time no mergers or acquisitions had actually been consummated under these new laws.[2]

Developments since that time have been quite dramatic, and the country has moved significantly, albeit haltingly, toward nationwide banking. By the end of 1988, all but three states allowed interstate acquisitions of at least some of their banks. A new count reveals no less than 14,600 offices of banking organizations existed outside the organizations' home states. Of these, 7,500 offices were permitted to offer all banking services. Thrift institutions operated more than 1,600 interstate offices.

The passage of state laws that allow full-service interstate branches has not entirely eliminated incentives for using other means of establishing interstate offices. Some states continue to be excluded from regional compacts that allow interstate banking with other reciprocating states. Also, under some circumstances, banking organizations still seem to prefer operating more limited nonbank offices in order to maintain an out-of-home-state presence. In addition, economic weakness in some areas and the problems faced by a number of thrift institutions have also spurred the growth of

---

The authors are, respectively, vice president and associate director of research, senior vice president and director of research, and research officer in charge of the financial section of the research department at the Atlanta Fed.

They acknowledge Pamela S. Frisbee, an economic analyst in the financial section of the Atlanta Fed's research department, who gathered and organized the data for this article.

interstate banking offices. After reviewing the legislative evolution of interstate banking, this article describes its recent developments, including changes in the top 50 banking organizations, outlines some explanations for its progress, and assesses some general policy implications of the nation's present course of action regarding banking expansion across state lines.

The Legislative Background of Interstate Banking

Full-service Offices. As the unsteady course of interstate banking indicates, laws governing interstate expansion by banking organizations have developed in a complex way. Throughout the nation's history each state has determined branching rules for banks under its charter. The McFadden Act, originally passed in 1927 and amended in 1933, clarified such restrictions for nationally chartered banks. In its 1933 version, the act allowed national banks the same geographic branching rights as those allowed to state-chartered institutions in their home states. That is, a national bank based in Tennessee would have the same branching rights as a bank chartered in that state. The effect of this legislation was to limit branching activity of national banks to a single state at most.

The McFadden Act left open the possibility of crossing state lines by using a bank holding company, which could own separate bank charters. Bank holding companies had not been widely used for interstate expansion before passage of the 1956 Bank Holding Company Act, but concerns about their possible proliferation were sufficient to prompt Congress to add the Douglas Amendment as part of the act. This law prohibited bank holding companies from acquiring banks outside their home state unless the other state explicitly allowed such purchases. Banking organizations that already used the holding company form to establish an interstate presence were permitted to maintain their existing interstate ties. This provision of the Douglas Amendment accounts for a few large regional organizations, such as First Interstate Bancorp in the West as well as First Bank System and Norwest Corporation in the upper Midwest.

The McFadden Act and the Douglas Amendment seemed to have closed the door to any additional interstate banking from 1956 to 1975. Then, however, after a special study of its financial laws and regulations, Maine passed legislation allowing out-of-state bank holding companies to acquire Maine-chartered institutions if bank holding companies headquartered in Maine were permitted reciprocal rights. Thus, Maine was the first state to take advantage of the Douglas Amendment's provision authorizing the entry of out-of-state holding companies. (Maine's reciprocity requirement was later dropped.)

Other states initially showed little interest in permitting this cross-state activity until 1982, when New York passed nationwide reciprocal banking legislation and Massachusetts led the New England states into a regional banking compact. Policymakers in New England sought to develop the region's relatively small banks into larger regional institutions that were deemed more effective in attracting capital to the region. The 1982 Massachusetts regional reciprocal law included Maine, Connecticut, Rhode Island, New Hampshire, and Vermont. Rhode Island and Connecticut enacted similar legislation in 1983, naming the same group of partner states.

With capital attraction motives similar to the New England states' and with the desire to build banking organizations large enough to resist takeover by banks from outside their region, other states began to consider and enter into regional compacts. Only the southeastern states were able to create a fairly uniform region, but even these states' laws differed in their lists of partner states. Several other states have adopted reciprocal laws lacking in uniformity and often including states that permit widely varying degrees of entry into other markets. Indiana's reciprocal law, for instance, currently includes 11 partners. Four of these have national reciprocal laws; three have regional reciprocal laws that exclude Indiana; and one allows interstate expansion by only one company.

Some states, mainly the large money center banks' headquarters, were excluded from each "region." New York, for example, is not mentioned in any regional legislation. This use of regional reciprocal banking laws to exclude organizations from outside a region evoked a constitutional issue of discrimination, which

was resolved in 1985 when the Supreme Court ruled that states could, in fact, define their own regional partners.

Even though many states initially took the regional-compact approach to interstate banking, broader access has dominated recent developments. Several states have opened their doors to banks from any state, either on a reciprocal or nonreciprocal basis. Some of these states, like Rhode Island, did so after initially allowing only regional entry into their markets. Others, like Arizona, have gone directly to national entry, thinking that less restrictive rules would encourage greater interest in acquisition of their banks, including perhaps weak or failing ones. At this writing, all but three states have enacted some form of interstate banking legislation. Twenty-one states and the District of Columbia allow national entry of one sort or another. Table 1 shows each state's type of legislation and its effective date.

As noted, southeastern states moved fairly early to adopt interstate banking by setting up a regional reciprocal compact that in 1985 became effective for several states. Except for Louisiana—where a weak economic and banking environment led to the permission of nationwide entry starting January 1, 1989—and Kentucky, the Southeast's approach to interstate banking has continued to be through regional compacts. As a group, the states of New England and the Southeast had the most restrictive interstate banking laws. Recently, however, Rhode Island has joined Maine with a national interstate law, and Vermont is scheduled to do so next year.

One effect of these restrictions is that large southeastern banks which have been formed during the last four years may have difficulty maintaining their relative size as banks from other states with newer regional and national interstate banking laws expand. For example, Michigan, New Jersey, and Ohio already have nationwide reciprocal banking, and other states are liberalizing their laws.

Though not used as widely as the state interstate banking laws, the emergency provisions of the 1982 Garn-St Germain Act provided another way for banking organizations to acquire and operate full-service offices in more than one state. This law permitted out-of-state banking organizations to acquire certain large, troubled commercial banks and insured mutual savings banks. Its provisions were modified and extended by the Competitive Equality in Banking Act of 1987, which also authorized the Federal Deposit Insurance Corporation (FDIC) to arrange interstate takeovers of institutions with assets of more than $500 million as long as the FDIC granted the necessary financial assistance. In addition, some states enacted laws that allow out-of-state banks and thrifts to acquire failing in-state institutions.

Foreign banks may own American-chartered banks and bank holding companies. Foreign banks that owned U.S. banks were not limited to one state until 1978. The passage of the International Banking Act in that year placed banks domiciled outside the United States on essentially the same footing as purely domestic institutions; foreign banks were required to choose one state in which they would own a bank or holding company and operate according to the laws of that state. Grandfather provisions permitted these firms to maintain their existing interstate systems.

Offices Limited to International Transactions. Foreign banks themselves were also allowed to have offices that perform only particular transactions in much the same way that U.S. banks could establish an interstate presence on a limited basis. Interstate expansion of foreign banks is controlled by state laws that vary in effect from prohibiting expansion outright to allowing only offices that provide financial services related to international transactions. Foreign institutions consequently can and do operate offices in several states.

Congress has long allowed U.S. banks to compete with foreign firms in the financing of international trade. Both domestic and foreign banks and bank holding companies were permitted by the Edge Act of 1919 to establish banking corporations, provided they serviced only those firms engaged in international trade. As U.S. trade links with the rest of the world expanded, so did the growth of Edge Act corporations. This structure allowed many banks to establish a corporate presence in another state.

Nonbank Banks. As Congress was seeking to limit geographic expansion through commercial banks' full-service offices, innovation in other areas was steadily eroding the legal barriers to interstate banking. A noteworthy but short-lived innovation that demonstrated banks' desires

Table 1.
Interstate Banking Legislation by State
(as of February 1, 1989)

State	Effective Date	Area	Number of Partner States
Alabama	Currently	Reciprocal, 12 states and D.C. (AR, FL, GA, KY, LA, MD, MS, NC, SC, TN, VA, WV).	13
Alaska	Currently	National, no reciprocity.	50
Arizona	Currently	National, no reciprocity.	50
Arkansas	Currently	Reciprocal, 16 states and D.C. (AL, FL, GA, KS, LA, MD, MS, MO, NE, NC, OK, SC, TN, TX, VA, WV). Reciprocity hinges on commitments to community reinvestment.	17
California	Currently	Reciprocal, 11 states (AK, AZ, CO, HI, ID, NV, NM, OR, TX, UT, WA).	11
	January 1, 1991	National, reciprocal.	
Colorado	Currently	Reciprocal, 7 states (AZ, KS, NE, NM, OK, UT, WY).	7
	January 1, 1991	National, reciprocal.	
Connecticut	Currently	Reciprocal, 5 states (MA, ME, NH, RI, VT).	5
Delaware	Currently	Reciprocal, 5 states and D.C. (MD, NJ, OH, PA, VA). Special-purpose banks permitted.	6
	June 30, 1990	National, reciprocal.	
District of Columbia	Currently	Nationwide, no reciprocity if community development commitments are made.	50
Florida	Currently	Reciprocal, 11 states and D.C. (AL, AR, GA, LA, MD, MS, NC, SC, TN, VA, WV). Under a 1972 law, NCNB and Northern Trust Corporation are grandfathered and can make further acquisitions.	12
Georgia	Currently	Reciprocal, 10 states and D.C. (AL, FL, KY, LA, MD, MS, NC, SC, TN, VA).	11
Hawaii	None		0
Idaho	Currently	National, no reciprocity.	50
Illinois	Currently	Reciprocal, 6 states (IA, IN, KY, MI, MO, WI). Nationwide, organizations may acquire failed institutions if the failed institution is larger than $1 billion in assets. Under a 1981 law, General Bancshares Corporation is grandfathered and can make further acquisitions in the state.	6
	December 1, 1990	National, reciprocal.	
Indiana	Currently	Reciprocal, 11 states (IA, IL, KY, MI, MO, OH, PA, TN, VA, WI, WV).	11
	July 1, 1992	National, reciprocal.	
Iowa	1972	Under a 1972 law, Norwest Corporation is grandfathered and is permitted to acquire banks in Iowa.	0
Kansas	None		0
Kentucky	Currently	National, reciprocal.	31*
Louisiana	Currently	National, reciprocal.	29*

continued on next page

Table 1 continued

State	Effective Date	Area	Number of Partner States
Maine	Currently	National, no reciprocity.	50
Maryland	Currently	Reciprocal, 14 states and D.C. (AL, AR, DE, FL, GA, KY, LA, MS, NC, PA, SC, TN, VA, WV) and special-purpose banks.	15
Massachusetts	Currently	Reciprocal, 5 states (CT, ME, NH, RI, VT).	5
Michigan	Currently	National, reciprocal.	20*
Minnesota	Currently	Reciprocal, 11 states (CO, IA, ID, IL, KS, MO, MT, ND, SD, WA, WY).	11
Mississippi	Currently July 1, 1990	Reciprocal, 4 states (AL, AR, LA, TN). Reciprocal, 13 states (AL, AR, FL, GA, KY, LA, MO, NC, SC, TN, TX, VA, WV).	4
Missouri	Currently	Reciprocal, 8 states (AR, IA, IL, KS, KY, NE, OK, TN).	8
Montana	None		0
Nebraska	Currently January 1, 1990 January 1, 1991	Special-purpose banks. Reciprocal, 10 states (CO, IA, KS, MN, MO, MT, ND, SD, WI, WY). National, reciprocal.	0
Nevada	Currently	National, no reciprocity.	50
New Hampshire	Currently	Reciprocal, 5 states (CT, MA, ME, RI, VT).	5
New Jersey	Currently	National, reciprocal.	21*
New Mexico	Currently January 1, 1990	Nationwide acquisition of failing banks. National, no reciprocity.	50
New York	Currently	National, reciprocal.	19*
North Carolina	Currently	Reciprocal, 12 states and D.C. (AL, AR, FL, GA, KY, LA, MD, MS, SC, TN, VA, WV).	13
North Dakota	Currently	A grandfathered interstate banking organization is permitted to sell its North Dakota banks to out-of-state bank holding companies.	0
Ohio	Currently	National, reciprocal.	23*
Oklahoma	Currently	National, no reciprocity.	50
Oregon	Currently July 1, 1989	8 states, no reciprocity (AK, AZ, CA, HI, ID, NV, UT, WA). National, no reciprocity.	8
Pennsylvania	Currently March 4, 1990	Reciprocal, 7 states and D.C. (DE, KY, MD, NJ, OH, VA, WV). National, reciprocal.	8
Rhode Island	Currently	National, reciprocal.	23*
South Carolina	Currently	Reciprocal, 12 states and D.C. (AL, AR, FL, GA, KY, LA, MD, MS, NC, TN, VA, WV).	13
South Dakota	Currently	National, reciprocal and special-purpose banks.	21*
Tennessee	Currently	Reciprocal, 13 states (AL, AR, FL, GA, IN, KY, LA, MO, MS, NC, SC, VA, WV).	13
Texas	Currently	National, no reciprocity.	50
Utah	Currently	National, no reciprocity.	50

continued on next page

Table 1 continued

State	Effective Date	Area	Number of Partner States
Vermont	Currently February 1, 1990	Reciprocal, 5 states (CT, MA, ME, NH, RI). National, reciprocal.	5
Virginia	Currently	Reciprocal, 12 states and D.C. (AL, AR, FL, GA, KY, LA, MD, MS, NC, SC, TN, WV).	13
Washington	Currently	National, reciprocal. Failing institutions may be acquired by organizations from any state.	21*
West Virginia	Currently	National, reciprocal.	29*
Wisconsin	Currently	Reciprocal, 8 states (IA, IL, IN, KY, MI, MN, MO, OH).	8
Wyoming	Currently	National, no reciprocity.	50

* *Does not count the two states where nationwide entry by acquisition of failing banks is possible.*

Source: Compiled by the Federal Reserve Bank of Atlanta.

to expand across state lines with deposit-gathering offices was the nonbank bank.

The wording of the Bank Holding Company Act defines a bank as any institution that both accepts demand deposits and makes commercial loans. By engaging in only one of these activities, several financial and nonfinancial companies were able to obtain bank charters and to qualify for FDIC deposit insurance in any state they chose to enter. Since these firms had not met its dual criteria, they were not banks for the purposes of the Bank Holding Company Act. The term *nonbank bank* derived from the fact that these institutions could perform some of the functions of a full-service bank but not all.

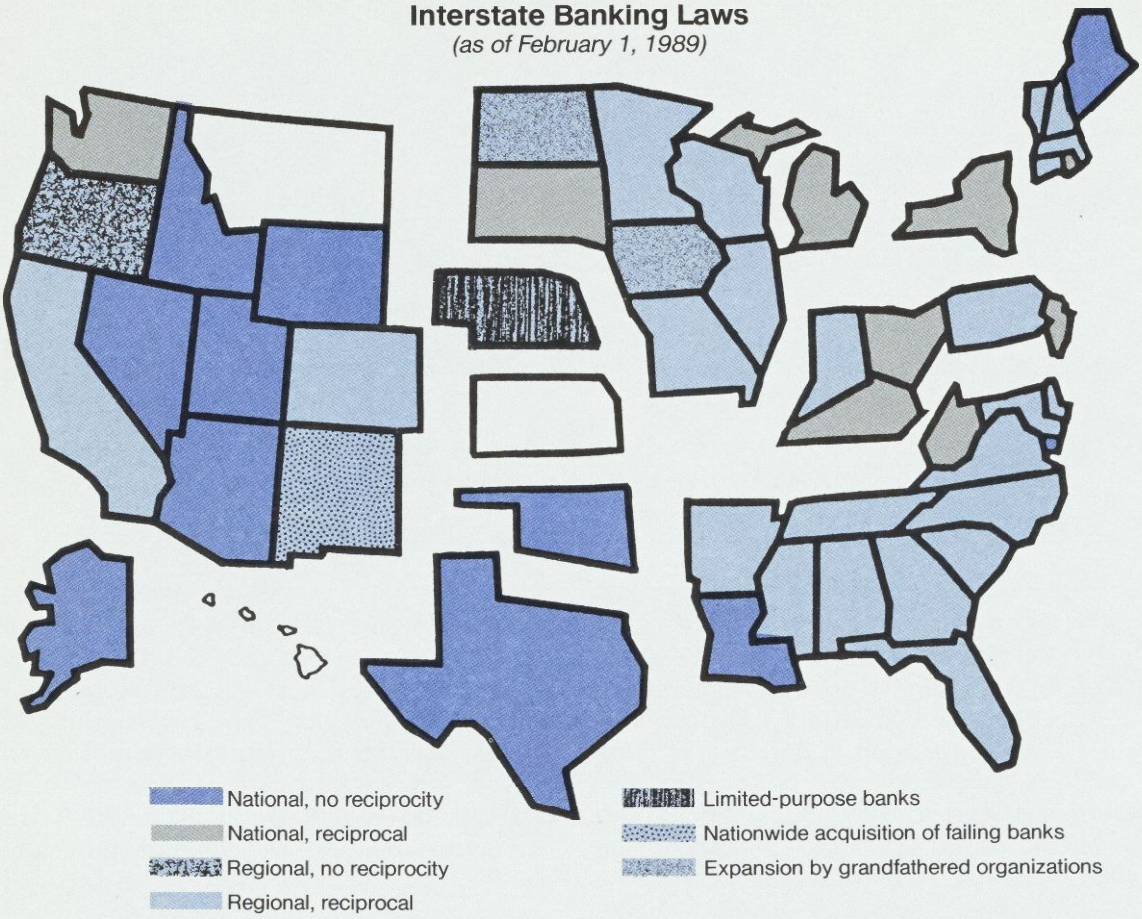
By 1983 congressional and regulatory concerns over a rash of nonbank bank charter applications led the Comptroller of the Currency to declare a moratorium on processing such requests. Though the Federal Reserve Board also sought to halt the establishment of these organizations, the central bank began processing applications after the Supreme Court in 1986 upheld the legality of nonbank banks. The situation remained unresolved until the passage of the Competitive Equality in Banking Act of 1987, which effectively prohibited the establishment of new nonbank banks. The existing ones were grandfathered and, as of this writing, 166 of them exist.

Nonbank Offices. Another way banking organizations innovated around limits to geographic expansion was to establish loan production offices on an interstate basis. Though doing little more than maintaining a staff of calling officers, these divisions of a banking company generate business for the head office and help establish a corporate identity in other states.

Until regional interstate banking took hold, the major avenue used by bank holding companies to move across state lines was the establishment of offices of nonbank subsidiaries under section 4(c)(8) of the Bank Holding Company Act as amended in 1970. This section allows bank holding companies to engage in certain activities other than taking deposits through subsidiaries established for this purpose. The laws prohibiting banking organizations from crossing state lines with full-service offices do not apply to 4(c)(8) subsidiaries because they do not meet the dual criteria for qualifying as a bank. A 4(c)(8) subsidiary may branch without McFadden Act or Douglas Amendment restrictions, giving banking organizations an opportunity to offer many services on a nationwide basis.

Section 4(c)(8) gave the Federal Reserve Board the authority to determine the activities in which subsidiaries formed under 4(c)(8) could engage. Various types are permitted provided that they are "so closely related to banking or managing or controlling banks as to be a proper incident thereto." Since 1970 the Federal Reserve has authorized many such activities by

Chart 1.
Interstate Banking Laws
(as of February 1, 1989)



regulation (that is, those generally approved for all holding companies) and by order (through case-by-case approvals resulting from special circumstances). Certain types of business have also been denied, however. (Appendix 2 to this article shows the activities permitted and denied as of January 31, 1989.)

Thrift Institution Offices. Thrift institutions also operate interstate offices that take deposits and make loans. Although the savings and loan industry never fell under the federal prohibitions relating to interstate banking, for many years the Federal Home Loan Bank System had precluded such activity by regulation and general policy. Starting in 1981, though, the Federal Home Loan Bank Board began allowing interstate mergers when an institution was in danger of failing. The Garn-St Germain Act of 1982 established provisions for these types of mergers. In 1986 the Federal Home Loan Bank Board issued a regulation similar to the Douglas Amendment on interstate activities of savings and loan institutions and mutual savings banks. In essence, this regulation permits interstate acquisitions for thrifts parallel to those for commercial banks.

Developments in Interstate Banking

The study that appeared in this Bank's *Economic Review* in 1983 reported estimates based on an extensive inventory of interstate offices operated by banking organizations. This inventory, which has remained unique to this day,

Table 2.
Changes in Interstate Banking Presence
(1983-88)

Type of Office	Number Reported in 1988	Number Reported in 1983	Change in Number Reported	Percent Change
Bank Offices Controlled by Domestic Bank Holding Company	7,364	1,258	6,106	485
Bank Offices Controlled by Foreign Bank Holding Company	128	148	-20	-14
Total Bank Offices	**7,492**	**1,406**	**6,086**	**433**
Offices of Foreign Banks	302	241	61	25
Domestic Edge Act Corporations	79	143	-64	-45
Total Offices for Foreign Transactions	**381**	**384**	**-3**	**-1**
Section 4(c)(8) Offices	6,446	5,500	946	17
Loan Production Offices	332	202	130	64
Total Nonbank Offices	**6,778**	**5,702**	**1,076**	**19**
Total Offices of Banks	**14,651**	**7,492**	**7,159**	**95**
Thrift Institutions	1,616	N.A.		
Total Interstate Offices	**16,267**	**N.A.**		

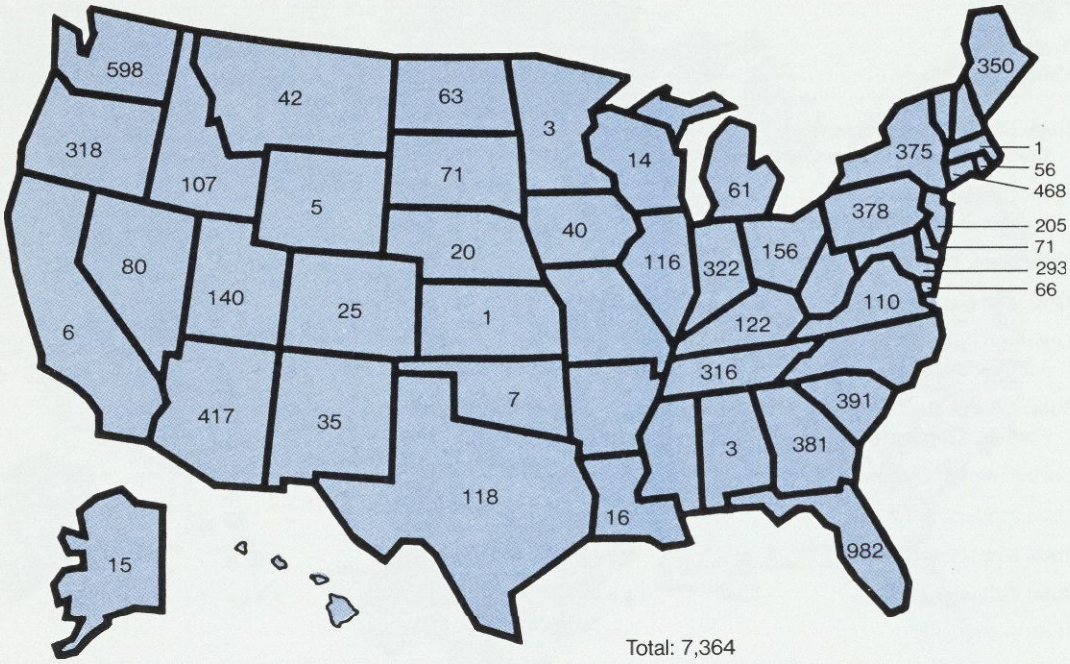
Sources: See Appendix 3 and David D. Whitehead, *A Guide to Interstate Banking,* Federal Reserve Bank of Atlanta, 1983.

found a rather large network of interstate offices, relatively few of which were permitted to offer a full range of banking services. All these full-service offices were in grandfathered banks. The vast majority of interstate offices were nonbank offices such as 4(c)(8) subsidiaries; loan production offices; foreign banks' branches, agencies, and representative offices; and Edge Act offices. The current data reported below come from new counts of interstate offices. (See Appendix 1 for a summary of interstate activity and Appendix 3 for sources.) Particularly in dealing with 4(c)(8) and loan production offices, the totals may not be 100 percent inclusive.

Full-service Offices. Between the time of the 1983 report and the end of 1988, as Table 2 clearly shows, the most dramatic development has been the rapid growth in domestically owned full-service interstate banking offices, which have quadrupled to more than 7,300 during this period. This spread of interstate full-service offices has not been uniform throughout the nation, however. Instead, several patterns of expansion are emerging, some related to each state's history of interstate banking and intrastate branching legislation and others related to features of the state's economy.

In some states that are part of regional compacts, a history of more liberal intrastate expansion laws led to the development of larger banks and their greater penetration throughout the state. The advent of interstate legislation helped these larger banks move quickly into surrounding states. North Carolina and Virginia are examples of states in which local banks rapidly acquired a substantial presence in neighboring states; at the same time, however, few if any banks from other states included in the regional compacts sought entry into those two states' markets. Similarly, the bigger Massachusetts economy and a fairly liberal intrastate expansion environment produced larger banks

Chart 2.
Number of Banking Offices in a Given State Controlled by Out-of-State Domestic Bank Holding Companies



relative to those in other states in the New England compact. Only one bank—with just one office—has entered Massachusetts, while Massachusetts banks have spread.

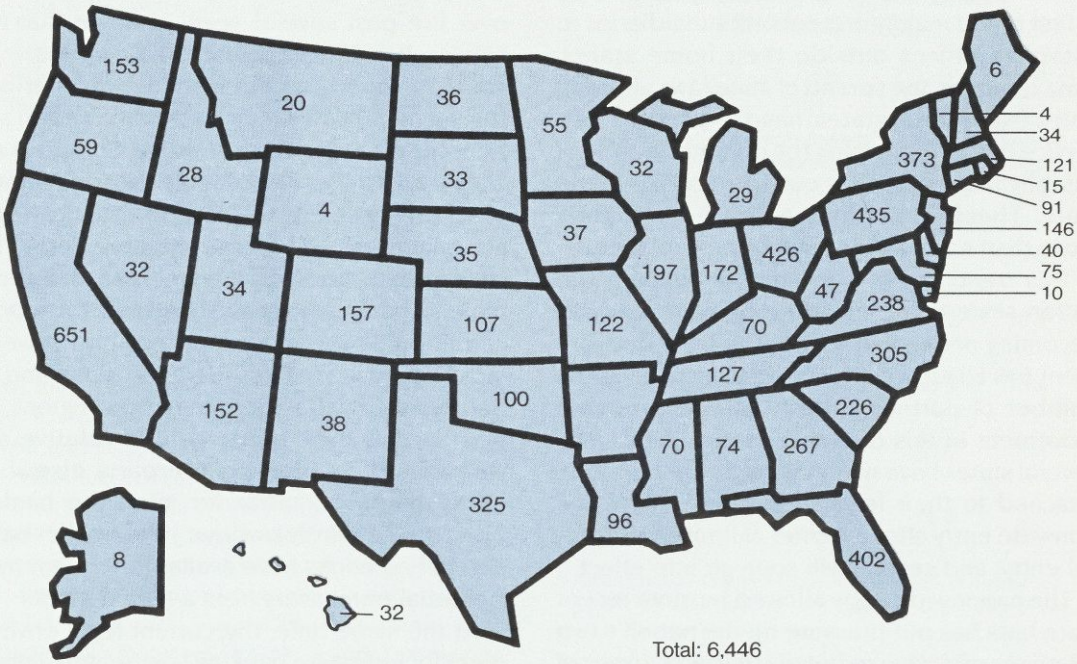
Elsewhere, large economies or rapid economic growth, and thus a favorable banking market, seem to be important factors in attracting both full- and limited-service offices from other states. Eight of the top 15 states in the number of interstate nonbank offices in 1981 and 9 of the top 15 in 1988 also ranked among the top 15 states in interstate full-service bank offices. Florida, where the growth of full-service banks owned by out-of-state organizations has been especially rapid, along with Georgia, New York, Pennsylvania, and Arizona, exemplifies this linkage.

Clearly, the length of time since enabling legislation was enacted to permit interstate banking has also affected the pattern. Developments in the New England compact's states illustrate this effect. Out-of-state banks have made substantial penetration into Connecticut and Rhode Island, whose regional laws became effective in 1983. New Hampshire and Vermont, which first allowed interstate banking in late 1987, have not yet seen banks enter. More liberal legislation in terms of partner states also has resulted in considerable out-of-state entry. Maine and Arizona, which allow nationwide nonreciprocal entry, have experienced substantial inroads by out-of-state banking organizations.

At the same time, the number of full-service offices of U.S. banks directly owned by foreign banks has declined. A significant decrease in California, where a major foreign-owned bank merged with a domestic bank, explains much of the drop.

Offices Limited to International Transactions. The number of offices limited to international transactions has declined slightly. The total of foreign banks' branch, Edge Act, and agency offices, which concentrate on investment-banking types of services, increased modestly during the period. Both the concentration and growth of these offices were greatest in California, Illinois, New York, Florida, and Georgia. The use of domestic Edge Act corporations, on the

Chart 3.
Number of Interstate 4(c)(8) Offices in a Given State Established by Nonbank Subsidiaries



other hand, has actually waned as other types of offices that can make international loans and take international deposits have expanded. A reduction in U.S. banks' interest in international lending may also have played a part in the decline. At the end of 1988, 79 domestic Edge Act offices were operating in 16 states, down from 143 offices in 18 states in the 1983 survey. Most domestic Edge Act offices are still in New York, California, and Florida, owing to the active international banking environments in those states.

Nonbank Offices. The use of this earlier alternative means of gaining presence across state lines has continued, but its spread has not been as dramatic as that of full-service interstate banking offices. Out-of-state loan production offices, for example, now exist in most states.

For domestic banks, the count of section 4(c)(8) subsidiaries increased from 5,500 to almost 6,500, and loan production offices rose from 202 to more than 325 over the same period. In contrast to the data on interstate banking offices, the total of 4(c)(8) offices masks developments in the spread of this type of nonbank operation. When bank holding companies merge across state lines, nonbank subsidiaries of the acquired company typically are transferred to the acquiring company. Former nonbank subsidiaries of the acquired company in the acquirer's state are thus no longer counted as interstate. Thus, the tally of interstate 4(c)(8) subsidiaries underestimates cross-boundary expansion since declines in the count of 4(c)(8) subsidiaries from this source partially offset new openings of interstate 4(c)(8) offices. The data in Table 2 thus indicate that interstate 4(c)(8) offices continued to increase in number even though alternatives for interstate expansion have broadened.

Thrift Institution Offices. In the thrift industry, financial difficulties faced by certain institutions have spurred interstate activity. The number of thrifts whose offices cross state lines has increased from 29 to 57 over the six-year period covered by the 1983 and 1988 surveys. As mentioned earlier, interstate thrift offices currently number 1,616.

Summary and Implications

Interstate banking has expanded significantly since the early 1980s. Bank holding companies at first relied mainly on nonbank subsidiaries to establish offices outside their home states. Now, however, the spread of state laws allowing entry from other states has resulted in full-service offices' becoming the dominant mode of establishing a presence outside a bank's home state. These operations account for slightly more than half of all interstate bank offices.

This trend is likely to grow throughout the nation since more interstate banking laws are becoming national in scope. The legal environment has already shifted toward increasing the number of partner states. Moreover, further movement in this direction seems likely since several states have national trigger mechanisms attached to their legislation. These allow nationwide entry after a period of limited reciprocal entry, and several will soon go into effect.

The nationwide entry allowed by more recent state laws has put pressure on the nation's two relatively exclusive regional compacts, those of New England and the Southeast. Attractive merger partners in both regions are growing scarcer, thus limiting the expansion capabilities of large banks within the narrowly defined regions. Banks in newly formed compacts and in states with nationwide expansion possibilities can potentially grow larger than banks in New England and the Southeast. In the more limited regions, the number of bidders for banks that would consider selling out is also limited. In New England, Rhode Island and Vermont have more recently recognized both problems and adopted national reciprocal laws. In the Southeast, Kentucky and Louisiana have made similar moves.

Expansion of state laws permitting nationwide entry, along with opportunities for banks in states with such laws, is pressuring state legislatures to enact, once and for all, nationwide banking with full-service offices. Firms' choices over the past several years indicate that full-service branches are preferable in many instances to offices limited to international transactions or to the activities allowed in section 4(c)(8) of the Bank Holding Company Act. There is little evidence that expanded interstate presence has resulted in the most egregious kinds of public harm often predicted. The country has witnessed the emergence of superregional banks, whose size relative to the money-center banks has grown appreciably. However, superregionals in states in New England and the Southeast, which adhere to narrow regions, are now seeing their newly gained relative size threatened by their counterparts elsewhere under newer and broader interstate banking laws. On the purely positive side, smaller banks and failing banks have available to them more potential purchasers from a national pool.

At the same time, the current legal environment for interstate banking has created a patchwork of laws which is sustaining the use of limited-service banking offices in areas with narrow banking compacts. Aside from states that restrict interstate banking to others subscribing to a regional compact, some states still allow little or no out-of-state entry. These arrangements are probably inefficient since banks need to use alternative means to evade geographic restraint. In addition, it renders large banks whose home-state charters impose geographic limitations less able to grow relative to banks elsewhere in the nation. Nationwide interstate banking could level this dimension of an "uneven playing field" and achieve a more equitable arrangement for all banks.

Changes in the Largest Banking Organizations during the Time of Interstate Banking

An important development of this decade has been a major restructuring of the list of the nation's largest banks.[1] Not only has a significant group of banks moved up into the ranks of the nation's largest banks, but relative rankings have also shifted dramatically and banks in the lower part of the rankings have increased in size relative to banks ranking in the top 10. Most of the lower-ranked banks that have grown in this manner can trace their size increase at least partly to expan-

sion allowed by interstate banking laws. Banks throughout the top 50 list now operate full-service offices in multiple states.

Three phenomena, all involving consolidation of relatively large banks, account for these shifts:

- a set of intrastate and interstate mergers of large troubled banks, such as Seafirst, into other large banks, such as BankAmerica Corp.;
- another group of intrastate consolidations of large healthy banks, such as that of Bank of New York Co. and Irving Bank Corp.; and
- a group of interstate mergers, such as that of Sun Banks of Florida and Trust Company of Georgia, that were allowed by new state laws.

An analysis of changes in the largest 50 banks in the nation, as measured by asset size, between the end of 1982 and the end of 1988 indicates the type and magnitude of the changes during the development of state-sponsored interstate banking (see next page). More than one-quarter of the banking organizations ranked in the top 50 in asset size at the end of 1982 have been replaced. Thirteen organizations have moved onto the top-50 list since 1982. Several of these banks have made large increases in rank; 3 were not even in the largest 100 in 1982. Of the banks displaced from the top-50 list, 11 were merged out of existence, and 2 have shifted out of the rankings.

Not only has the makeup of this list changed, but shifts in rank have also been significant. Among 1988's 50 largest banks, 18 institutions had moved up in rank by 10 or more places since 1982; 8 ascended 40 or more places. Mergers in this group of fast-climbing banks have resulted in the formation of a group of banks that financial analysts and the press have dubbed "superregionals." Such companies are neither new nor unique. In fact, the 1983 Atlanta Fed study documented the existence of a few large regional bank holding companies that operated under grandfather provisions of the Bank Holding Company Act long before the recent wave of state laws allowing interstate entry. In addition, interstate acquisitions have not been limited to new members of the top 50. Several of the country's 10 largest banks have made interstate bank and thrift acquisitions during this decade.

Nevertheless, new large banks formed mainly through interstate mergers account for most of the major increases in rank and in relative size recorded during this decade among more sizable institutions. For instance, 9 banks entered the top 25 between 1982 and 1988. Of these, eight owe a major part of their external growth to interstate mergers allowed under new state laws. (The ninth, Barnett Banks of Florida, grew mainly by acquisitions in Florida, but that company also made interstate acquisitions in Georgia.) Another institution, Bank of Boston, moved up within the top 25 mainly through interstate mergers. Thus, the states' goal of building larger banks headquartered in their region was achieved under the regional interstate pacts: 3 of the 9 new entrants to the top 25 were in the Northeast and 4 were in the Southeast.

Of the second 25 largest banks, upward moves from outside the top 50 accounted for the presence of 8 institutions. All rose in the ranking through at least some interstate acquisitions. Two of the banks moving up into the second 25 are headquartered in the Southeast region.

The largest banks in the country, most often located in money centers and often precluded from interstate acquisitions of all but failing institutions, lost size relative to the rest of the top 50 between 1982 and 1988. Outright shrinkage by some large troubled banks makes comparisons imperfect, but some indication of relative size changes can be found by comparing the average assets for each successive group of 10 banks in the ranking with average assets for the top 10 (see

Table 1.
Comparison of Average Assets by Ranking Group for Top 50 U.S. Banking Organizations 1982 and 1988

Ranking Group	Average Assets (billion $)		Average Assets in Ranking Group to Average Assets of Top 10 (percent)	
	1988	1982	1988	1982
Largest 10	84.8	66.5	100.0	100.0
11th - 20th Largest	35.2	22.2	41.0	33.4
21st - 30th Largest	26.2	12.5	30.5	18.8
31st - 40th Largest	20.0	7.3	23.6	11.0
41st - 50th Largest	13.1	6.3	15.4	9.5

Source: Compiled by the Federal Reserve Bank of Atlanta from data obtained from the Board of Governors of the Federal Reserve System.

Table 2.
Top 50 Banking Organizations, 1982-88
(ranked by consolidated assets)

Banking Organization	1988	1982	Rank Change 1982 to 1988
Citicorp, New York	1	1	0
Chase Manhattan Corp, New York	2	3	+1
BankAmerica Corp., California	3	2	-1
J.P. Morgan & Co., New York	4	4	0
Security Pacific Corp., California	5	10	+5
Chemical Banking Corp., New York	6	6	0
Manufacturers Hanover Corp., New York	7	4	-3
First Interstate Bancorp., California	8	8	0
Bankers Trust New York Corp., New York	9	9	0
Bank of New York Co., New York	10	24	+14
Wells Fargo & Co., California	11	13	+2
First Chicago Corp., Illinois	12	11	-1
PNC Financial Corp., Pennsylvania	13	32	+19
Bank of Boston Corp., Massachusetts	14	18	+4
Bank of New England Corp., Massachusetts	15	60	+45
Mellon Bank Corp., Pennsylvania	16	14	-2
Continental Bank Corp., Illinois	17	7	-10
NCNB Corp., North Carolina	18	26	+8
First Fidelity Bancorp., New Jersey	19	89	+70
Suntrust Banks, Georgia	20	71	+51
Fleet/Norstar Financial Group, Rhode Island	21	70	+49
First Union Corp., North Carolina	22	46	+24
Shawmut National Corp., Massachusetts	23	55	+32
Marine Midland Banks, New York	24	16	-8
Barnett Banks, Florida	25	38	+13
Bank One Corp., Ohio	26	62	+36
Natwest Holdings, New York	27	37	+10
Republic New York Corp., New York	28	29	+1
First Bank System, Minnesota	29	22	-7
NBD Bancorp., Michigan	30	25	-5
Sovran Financial Corp., Virginia	31	80	+49
Norwest Corp., Minnesota	32	20	-12
First Wachovia Corp., North Carolina	33	40	+7
National City Corp., Ohio	34	41	+7
Citizens and Southern Corp., Georgia	35	46	+11
Midlantic Corp., New Jersey	36	73	+37
Bank of Tokyo*	37	31	-6
MNC Financial, Maryland	38	61	+23
MCorp, Texas	39	27	-12

continued on next page

Table 2 continued

Banking Organization	1988	1982	Rank Change 1982 to 1988
Corestates Financial Corp., Pennsylvania	40	44	+4
Southeast Banking Corp., Florida	41	34	-7
Boatmen's Bancshares, Missouri	42	112	+70
Keycorp, New York	43	109	+66
U.S. Bancorp, Oregon	44	51	+7
Bankmont Financial Corp., Illinois	45	36	-9
Barclays USA, Inc., Delaware	46	136	+90
First City Bancorp. of Texas	47	23	-24
Valley National Corp., Arizona	48	35	-13
Michigan National Corp., Michigan	49	43	-6
Comerica, Michigan	50	33	-17

* *Assets of U.S. banks only.*

Source: See Table 1.

Table 1). For each group of 10 below the top group, mean assets were a considerably larger proportion of mean assets of the 10 largest banks in 1988 than they were in 1982. (Removing BankAmerica Corp. from the top 10 in each year has little impact on relative mean assets.)

Clearly, banks with substantial operations in several states have become more important during this decade, and the state-granted opportunity to consolidate across state lines has been a major factor in the growth. Banks in the Southeast, in particular, have expanded. With their two- to three-year head start, banks in the southeastern compact accounted for 9 of the top 50 institutions in 1988 as compared to 6 in 1982. Though 4 of the 25 largest banks in 1988 were southeastern banks, no institutions from this region were present in the top 25 in 1982. More recently strong gains in size and rank have come from states in the East and the Midwest, which have only recently enacted regional laws or expanded their list of partner states. It is possible that banks in states new to interstate banking and in states with national entry will catch up with southeastern banks since state legislatures have in the past two or three years made the process easier.

Note

[1]The term *bank* here includes all organizations operating a full-service banking business in the United States. All but one of the organizations listed in the top 50 in 1988 were domestically chartered bank holding companies. (Three of these were owned by foreign organizations.) The other is a foreign institution that directly owns three U.S.-chartered banks.

Appendix 1.
Summary Table of Interstate Activity

State	Bank Offices				Offices Limited to International Transactions				Nonbank Offices				
	Domestic		Foreign										
	Out-of-State Banking Organizations Represented	Banking Offices Controlled by Out-of-State Bank Holding Companies	Out-of-State Banking Organizatons Represented	Banking Offices Controlled by Out-of-State Bank Holding Companies	Foreign Banks			Domestic Edge Act Corporations	Loan Production Offices	Offices of 4(c)(8) Subsidiaries	Total Offices of Banking Organizations	Savings and Loans	Total Interstate Offices
					Agency	Edge	Branch						
Alabama	1	3							5	74	82	1	83
Alaska	2	15								8	23		23
Arizona	11	417	1	1					5	152	575	15	590
Arkansas									1		1	31	32
California	1	6	5	66	70	1	8	17	33	651	852	28	880
Colorado	8	25							10	157	192	111	303
Connecticut	7	468							8	91	567	1	568
Delaware	17	71	3	3						40	114	9	123
District of Columbia	5	66					2		3	10	81	54	135
Florida	21	982			34	7		19	21	402	1,465	425	1,890
Georgia	6	381			16			1	20	267	685	25	710
Hawaii					2			1	1	32	36	18	54
Idaho	2	107								28	135	28	163
Illinois	15	116				1	49	3	22	197	388	57	445
Indiana	10	322							4	172	498	33	531
Iowa	2	40								37	77	9	86
Kansas	1	1							2	107	110	21	131
Kentucky	10	122							5	70	197	30	227
Louisiana	1	16			1				4	96	117	13	130
Maine	6	350							4	6	360	12	372
Maryland	12	293						2	10	175	480	46	526
Massachusetts	1	1					7	1	20	121	150	14	164
Michigan	5	61			1				6	29	97	21	118
Minnesota	1	3						2	3	55	63	16	79
Mississippi									1	70	71		71
Missouri									4	122	126	54	180

continued on next page

Appendix I continued

State	Bank Offices				Offices Limited to International Transactions				Nonbank Offices				
	Domestic		Foreign		Foreign Banks								
	Out-of-State Banking Organizations Represented	Banking Offices Controlled by Out-of-State Bank Holding Companies	Out-of-State Banking Organizatons Represented	Banking Offices Controlled by Out-of-State Bank Holding Companies	Agency	Edge	Branch	Domestic Edge Act Corporations	Loan Production Offices	Offices of 4(c)(8) Subsidiaries	Total Offices of Banking Organizations	Savings and Loans	Total Interstate Offices
Montana	3	42								20	62	6	68
Nebraska	3	20								35	55		55
Nevada	3	80							1	32	113	8	121
New Hampshire									3	34	37		37
New Jersey	2	205							13	146	364	30	394
New Mexico	1	35							4	38	77		77
New York	4	375	8	39	16	2	43	20	28	373	896	74	970
North Carolina			1	19					7	305	331	1	332
North Dakota	3	63								36	99		99
Ohio	3	156						1	9	426	592	131	723
Oklahoma	3	7							2	100	109	10	119
Oregon	4	318					5	2	4	59	388	85	473
Pennsylvania	5	378					5	1	9	435	828	13	841
Rhode Island	3	56							1	15	72		72
South Carolina	8	391							4	226	621	1	622
South Dakota	3	71								33	104	4	108
Tennessee	7	316							8	127	451	11	462
Texas	6	118			14	8		5	32	325	512	78	590
Utah	5	140								34	174	11	185
Vermont									1	4	5		5
Virginia	7	110						1	7	238	356	22	378
Washington	6	598					10	2	4	153	757	55	812
West Virginia										47	47	2	49
Wisconsin	3	14							2	32	48		48
Wyoming	3	5						1	1	4	11	2	13
TOTAL		7,364		128	154	19	129	79	332	6,446	14,651	1,616	16,267

Appendix 2.
Activities of 4(c)(8) Offices Permitted by Regulation
(as of January 1989)

Making, acquiring, or servicing loans such as would be made by the following companies:
- consumer finance
- credit card
- mortgage
- commercial finance
- factoring

Operating an industrial or Morris Plan bank or other industrial loan company

Performing trust company or fiduciary activities

Investment or financial advising

Full payout leasing of personal or real property

Investments in community welfare projects

Data processing services

Acting as insurance agent or broker primarily in connection with credit extensions

Underwriting credit life, accident, and health insurance

Courier services

Management consulting to depository institutions

Issuance and sale at retail of money orders with a face value of not more than $1,000, U.S. savings bonds, and travelers checks

Real estate and personal property appraisal

Arranging commercial real estate equity financing

Securities brokerage

Underwriting and dealing in U.S. government obligations and money market instruments

Foreign exchange advisory and transactional services

Futures commission merchant

Investment advice on financial futures and options on futures

Consumer financial counseling

Tax planning and preparation

Check-guaranty services

Operating a collection agency

Operating a credit bureau

Activities of 4(c)(8) Offices Permitted by Order

Operating a "pool reserve plan" for loss reserves of banks for loans to small businesses

Operating a savings and loan type business in Rhode Island

Operating certain state stock savings banks

Buying and selling gold and silver bullion and silver coin for the account of customers

Operating an Article XII New York Investment Company

Performing commercial banking functions at offshore locations

Offering NOW accounts

Operating a distressed savings and loan association

Issuance and sale of variably denominated payment instruments (maximum face value of $10,000)

Operating a chartered bank that does not both take demand deposits and make commercial loans

Providing financial feasibility studies for specific projects of private corporations; valuations of companies and large blocks of stock for a variety of purposes; expert witness testimony on behalf of utility companies in rate cases

Providing advice regarding loan syndications, advice in connection with merger, acquisition/divestiture, and financing transactions for nonaffiliated financial and nonfinancial institutions; valuations for nonaffiliated financial and nonfinancial institutions; fairness opinions in connection with merger, acquisition, and similar transactions for nonaffiliated financial and nonfinancial institutions

Executing and clearing futures contracts on stock indexes and options on such futures contracts

Advisory services with respect to futures contracts on stock indexes and options on futures contracts

Credit card authorization services and lost or stolen credit card reporting services

Acting as a broker's broker of municipal securities

Employee benefits consultant

Student loan servicing activities

Offering the combination of securities brokerage services and related investment advice to institutional customers

Printing and selling checks

Cash management services

Acting as agent and adviser to issuers of commercial paper in connection with the placement of such paper with institution purchasers

Underwriting and dealing in, to a limited extent, municipal bonds, mortgage-related securities, consumer-receivable related securities, and commercial paper

Provision of financial office services

Operating a proprietary system for trading put and call options on U.S. Treasury securities

Retention of a thrift after the thrift's parent is acquired by a new bank holding company

Acquisition of a healthy savings bank which qualifies as a commercial bank on the basis of its commercial loans and demand deposits

Permitting a nonprofit tax-exempt college to become a bank holding company and engage in college activities, including fund raising incidental to educational activities, but requiring the college to divest real estate received as gifts

Consulting and management services to insolvent thrifts

Corporate bond trading

Activities Prohibited under Section 4(c)(8)

Insurance premium funding

Underwriting life insurance that is not related to credit extension

continued on next page

Appendix 2 continued

Real estate brokerage

Land investment or real estate development

Real estate syndication

Management consulting

Property management services

Operating a travel agency

Contract data entry services

Underwriting property and casualty insurance

Dealing in platinum and palladium

Engaging in pit arbitrage

Public credit ratings on bonds, preferred stock, and commercial paper

Acting as a specialist in French franc options on the Philadelphia Stock Exchange

Selling title insurance

Sale of certain thrift notes

Oil and gas activities

Timber brokerage activities

Sale of level-term credit life insurance

Acceptance of deposit accounts linked to credit card accounts

Selling auto club memberships

Appendix 3.
Sources for Information

Domestic Banking Offices	Federal Reserve Board of Governors and Federal Reserve Bank of Atlanta. Number of offices as of December 31, 1987. Structure as of September 15, 1988.
Foreign Banking Offices	Federal Reserve Board of Governors. Structure as of June 30, 1988. Number of offices as of December 31, 1987.
Foreign Banks: Agency, Edges, and Branches	Federal Reserve Board of Governors as of June 30, 1988.
Domestic Edges	Federal Reserve Board of Governors as of June 30, 1988.
Loan Production Offices	Federal Reserve Bank of Atlanta 1987 Survey of 200 Largest Banking Organizations in the United States.
Offices of 4(c)(8) Subsidiaries	Federal Reserve Bank of Atlanta survey and Federal Reserve Banks of Minneapolis, St. Louis, Boston, and Atlanta.
Savings and Loan Associations	Federal Home Loan Bank Board as of June 30, 1988.

Notes

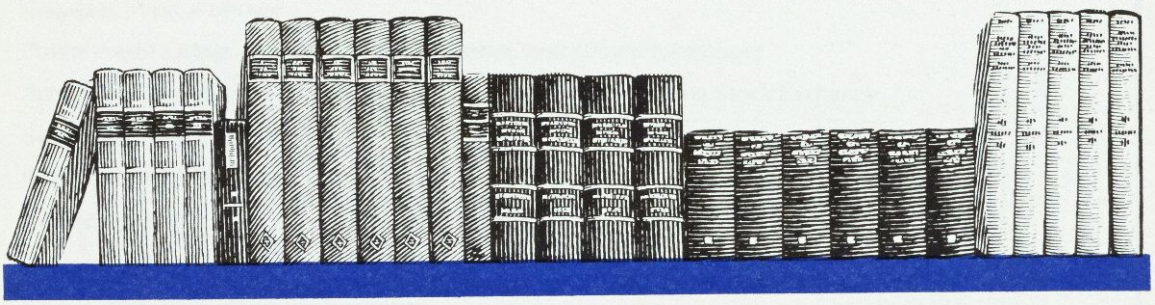
[1] David D. Whitehead, "Interstate Banking: Taking Inventory," Federal Reserve Bank of Atlanta *Economic Review* 68 (May 1983): 4-20. For more detail on the interstate presence of commercial banks in the early 1980s, see David D. Whitehead, *A Guide to Interstate Banking 1983*, Federal Reserve Bank of Atlanta, 1983.

[2] David D. Whitehead, "Interstate Banking: Probability or Reality?" Federal Reserve Bank of Atlanta *Economic Review* 70 (March 1985): 6-19.

Book Review

Migration and Residential Mobility in the United States

by Larry Long
New York: Russell Sage Foundation, 1988.
416 pages. $39.95.



As migration transforms the demographics of a region, its economic and social features shift as well: the education, experience, or number of workers might be altered, thus changing the nature of a local work force. Also, the different saving and spending habits of a new population can significantly modify an area's financial and consumer markets. Research that attempts to quantify migration's impacts can thus be helpful in many different arenas.

One recent contribution to the literature is *Migration and Residential Mobility in the United States* by Larry Long. The publisher's news release heralds this study as "a comprehensive interpretation of the last half century of Americans-in-motion." The publicist's claim notwithstanding, Long's stated goal is less ambitious: to integrate and communicate the trends and patterns of geographical mobility within the United States. His overview of migration revolves around the answers to seemingly simple questions such as *who, where, how much, why,* and *with what effect.*

Long's monograph succeeds most when the migration picture is clear, that is, when the *who, where,* and *how much* questions can be answered with precise and consistent data. The author is at his best in giving an overall numerical perspective and describing what has and has not changed, as measured by national censuses and surveys, over the last 50 years. That he does this well will come as no surprise to migration researchers, since Long is a well-known and highly respected scholar at the Center for Demographic Studies of the U.S. Bureau of the Census. His prolific research is further evinced in the fact that 20 of his own journal articles are cited among the bibliographic entries.

The author also succeeds in pointing out where the migration picture is blurred, that is, where our understanding of the basic determinants and consequences of migration remains incomplete. The obscurity here emerges as Long seeks answers to the *why* and *with what effect* questions about migration. Because these knowledge gaps limit the policy implications that can be inferred authoritatively, Long's contribution is less in the realm of policy prescriptions, which are few and tentative, than in identifying many of the gaps in our understanding of migration's causes and socioeconomic effects.

Following an introduction to the available migration research and data, Long's tome investigates how much geographical mobility has occurred within the United States over short

and long distances, as well as how levels and rates have changed not only recently but also over extended periods. His discussion focuses primarily on the degree of migration between states. Thus, major interstate population movements—where migrants come from and where they go—are identified, and issues including the extent and likelihood of return and repeat interstate migration are explored. Numbers abound in these chapters, although Long presents the data effectively, using visual aids such as maps that are shaded to reflect data or that depict migration streams with arrows.

Among the major movements chronicled in Long's maps, graphs, charts, and tables are the 1930's exodus from the Dustbowl states, the migration of many New Yorkers to the New Jersey suburbs and Florida beaches, and the parade of migrants to California. He also notes many shifting migration patterns of the last half-century, particularly the pre-1970's South-to-North stream that since has experienced a dramatic reversal.

Several of the migration trends that Long discusses were unanticipated when they occurred, and others probably will surprise readers even in retrospect. It is fairly well known, for example, that blacks reversed their net out-migration from the South during the 1970s and several heavily industrialized northern states that compose the Rustbelt lost population to migration during the last decade. Long also presents three findings that will prove novel to most readers: (1) northeastern and midwestern states that experienced net out-migration in the 1970s typically had below-average rates of out-migration among persons at prime migration ages; (2) high-growth states in the West and South typically have had above-average rates of both in-migration *and* out-migration; and (3) the propensity to leave one's state of birth was lower in the late 1970s—when the baby boom began moving through adulthood—than in the late 1950s, a time popularly portrayed as one of stability.

Of course, many of Long's findings are hardly surprising. The notion that "migration begets migration"—that people who move are more likely to move again or that large numbers of in-migrants may increase the mobility potential of nonmigrants—jibes with intuition. Also, his discovery that the incidence of long-distance moving has increased might be expected from certain important developments like post-World War II improvements to the nation's transportation and communications networks, industrial decentralization, and the growing importance of a widely dispersed service sector.

The early chapters also provide some items to interest trivia buffs, although the examples cited here are more than trivial. The reader is informed, for example, that the largest county in the 48 contiguous states is San Bernadino, California, with more than 20,000 square miles, and the smallest is the two-square-mile Washington, D.C., suburb of Falls Church City, Virginia. This type of information is not mere minutiae, though, since Long uses such tidbits to illustrate key points. The "large county-small county" example illustrates an obvious limitation of presenting internal U.S. migration statistics on a county basis. A family in San Bernadino can move 120 miles and not change counties, but a move of a few blocks may entail a change in county residence in the suburbs of the nation's capital.

Chapters 5-7 explore the reasons for and the consequences of migration. The first of these chapters considers the impacts of migration, especially at the regional level. The racial composition of migration streams is addressed, as are some effects of population movements on the levels of education and poverty. Generally, the net effect of migration on the South has been an improvement in its socioeconomic standing relative to the rest of the nation. Chapter 6 considers migration from a metropolitan-nonmetropolitan perspective. Interest in this topic has intensified over the last decade and a half, partly because movement between urban, suburban, and rural areas provided researchers with another migration surprise. From the late 1960s through the 1970s migrants' preferences shifted from favoring metropolitan to nonmetropolitan locales, but around 1980 metropolitan areas experienced a resurgence. Chapter 7 examines the reasons that people give for moving. Long points out the shortcomings of inferring migration's determinants by querying individuals: for example, they simply may not know their motives, or may have reasons for misrepresenting them. Nonetheless, he provides some justification for this methodology by indicating the drawbacks of determining the

motives for moving solely from observed migratory flows. For example, the reasons for retirees' migrating cannot truly be identified and compared except through personal interviews.

In the book's last chapter, Long analyzes international census and survey statistics in order to compare migration in different nations. From this vantage, the author poses a number of important policy questions:

- Do state welfare benefit levels stimulate or inhibit migration?
- Should public policy attempt to influence mobility of the unemployed or low-income persons and families by directly subsidizing migration?
- Is national economic growth constrained by worker reluctance to move in response to industrial restructuring?
- Do national policies regarding home ownership and financing restrict mobility?
- Is greater mobility increasing the need for social services, such as care of elderly people that relatives living nearby would otherwise provide?
- Should the U.S. government enact "place" policies to increase economic opportunities in some areas or "worker" policies to increase mobility toward areas with growing employment opportunities?

Answers to these questions rely on a much more detailed understanding of the causes and consequences of migration than is now available. Research to help define the factors that motivate migration must still be done. It is critical to know, for example, whether economic or noneconomic influences are migration's primary determinants and whether the reasons for moving have changed over the decades. Answers to the causality question will be found only in a properly constructed model that incorporates economic features such as employment opportunities as well as noneconomic influences such as climate, amenities, or quality of life.

Long considers the data on reasons for moving and concludes that life-style amenities play a strong subordinate role in motivating migration, though there is little evidence that they have superseded economic motivation. He also concludes that "there is no firm evidence for or against the widespread notion that reasons for moving have changed over the last several decades in favor of noneconomic motives like a desire to live where the weather is nice, where recreational opportunities are present (whether mountains for skiing or beaches for surfing), or where other quality-of-life amenities exist."

Long does not review in detail the voluminous economics literature relating to migration, but he cites major research contributions, provides relevant bibliographical information, and ably summarizes the strengths and shortcomings of "traditional econometric models." Most importantly, he recognizes that a variety of age, period, and cohort effects influence migration propensities. The author also seems to agree that for theory development a more extensive use of both macroeconomic analyses of migration and microdata is necessary.

These exercises are not carried out in Long's book, but to ignore the work because of its lack of detailed econometric findings by others or statistical analyses based on the author's own work would be a mistake. Rather, economists interested in understanding the *why* and *with what effects* questions surrounding migration will find that this monograph provides valuable, if not essential, background material. (The book contains a wealth of data in tables and appendices for readers who want to review the statistics.)

The *with what effect* question actually has many aspects. Long writes, "[The question] can refer to the effects of migration on places, on people (those who do not migrate as well as those who do), and on more abstract concepts, such as occupational mobility of the labor force, the care of elderly persons whose children have moved away, and the nature of participatory democracy and the role of voluntary activities in a society where individuals' commitment is to career and its advancement through migration rather than to places." Several aspects of this question are inherently complex. Moreover, the census data and the survey information that Long has chosen are not suitable for addressing many of these difficult impact issues. Thus, his concentration on migrants' characteristics and the effects of migration on places is not surprising.

Migration and Residential Mobility in the United States is one in a series of analyses aimed at condensing and analyzing the vast statistical results of the 1980 census. Topics in the series explore major features of our society—ethnic groups (blacks, Hispanics, foreign-born); spatial dimensions (migration, neighborhoods, housing, regional and metropolitan growth and decline); and status groups (income levels, families and households, women). The studies, constituting "The Population of the United States in the 1980s," were planned, commissioned, and monitored by the National Committee for Research on the 1980 Census; this committee was formed by the Social Science Research Council and is sponsored by the council, the Russell Sage Foundation, and the Alfred P. Sloan Foundation, with the collaboration of the U.S. Bureau of the Census. The ultimate objective of the committee and its sponsors has been to produce "a definitive, accurate, and comprehensive picture of U.S. population in the 1980s." Long's monograph represents an important contribution to the studies and surely will be a valuable reference work for researchers in this field.

William J. Kahley

The reviewer is an economist in the regional section of the Atlanta Fed's Research Department.

THE FEDERAL RESERVE BANK OF ATLANTA
will host a conference on

Financing Community Development in the South

The Federal Reserve Bank of Atlanta
Conference Center
October 31 - November 1, 1989

Though Southern states continue to experience rapid growth on average, many areas in the region remain impoverished. This part of the country seems to have reached a plateau in narrowing the average income gap with the nation. Traditional industrial development strategies, which took advantage of the region's historically low wages and taxes, no longer seem to be effective now that developing nations can offer even lower costs.

This two-day conference will look beyond economic growth to consider plans for changing the region's economic structure and methods of financing this transformation. The Atlanta Fed will bring together community leaders, state and local policymakers, scholars, and a range of experts from the banking, business, and planning fields to discuss the following topics:

- **How the South Stacks Up and How It Needs To Change during the 1990s**
- **What Development Strategies Have Worked Elsewhere**
- **Financing Issues**
  - **Public Infrastructure**
  - **State and Local Tax Systems**
  - **Funding Entrepreneurship**
  - **The Role of Financial Intermediaries: Commercial and Investment Banks**
- **How the Community Reinvestment Act Can Foster Local Economic Development**
- **Investing in Human Capital: Educational Reform and Better Jobs**

The conference, which will be held in the Atlanta Fed's new conference center, will offer ample opportunity for interchange among the speakers and the audience.

If you would like more information on this upcoming event, please send your name and address to Linda Donaldson, conference coordinator, or contact her at 404-521-8747.

Economic Review

Federal Reserve Bank of Atlanta
104 Marietta St., N.W.
Atlanta, Georgia 30303-2713

Address Correction Requested

Bulk Rate
U.S. Postage
PAID
Atlanta, Ga.
Permit 292