DALLASFED

During the past decade, predictions of economic
growth and unemployment by major forecasting
groups were quite good, but their predictions of
inflation were poor in that they frequently under-
estimated the actual rates of inflation. These results
were obtained from an analysis of annual surveys of
forecasts of economists and econometric firms
published in *Business Week*. Comparison of the
judgmental forecasts and econometric forecasts
revealed little difference in their accuracy. One
plausible explanation for this similarity in perfor-
mance is that the economists and the managers of
econometric models are influenced by each other's
opinions and, thus, it is impossible to distinguish them
as independent entities.

The courts should consider revising their standards
for antitrust analysis of commercial bank mergers.
The current line-of-commerce approach assumes
each bank competes only with other banks in its
immediate area and that all banks offer one common
"package" of services. Recognition of service
packaging is a useful starting point, but failing to
acknowledge the existence of distinct classes of bank
customers results in misleading assessments of the
effects of proposed mergers.

A Comparison of Judgmental and Econometric Forecasts of the Economy: The Business Week Survey

By Thomas B. Fomby*

People are presumably most interested in forecasts of variables closely affecting their own activities — for example, prices and quantities of the products they sell. But aggregate measures such as inflation, gross national product (GNP) growth, and unemployment play an important role in shaping the expectations and decisions of many executives at the local level, including those managing large, diversified corporations. Thus, expectations about the economy as a whole can significantly affect day-to-day business decisionmaking.

Each year since 1971, *Business Week* magazine has published in its year-end issue a survey of forecasts of inflation, average unemployment, and the percentage change in real GNP for the coming year. The forecasts are grouped into two categories: those made by corporate and consulting economists and those made by firms having econometric models. Since this survey has quite a broad audience and has not previously been analyzed, further study of the forecasts seems warranted.

* *Thomas B. Fomby is an associate professor of economics at Southern Methodist University and a consultant at the Federal Reserve Bank of Dallas.*

The forecasts by the economists may be described as "judgmental" only because they may have been formed without using the tools of statistical analysis or economic models. The econometric-model forecasts are, by and large, the result of specifying multiequation economic models, estimating the economic relationships represented by a model's equations, and using the estimated model to forecast the economic variables. Though a distinction is drawn between these two categories of forecasts, their differences may be more apparent than real. Both groups of forecasters are aware of each other's predictions and may be influenced thereby.

One purported advantage of the econometric method of forecasting over judgmental methods is that a structural econometric model provides a framework by which information can systematically be incorporated into a forecast. Since they make explicit the forecaster's assumptions, econometric models should aid the forecaster in understanding what made a prediction good or bad and, thus, could aid future forecasting accuracy.

With these ideas in mind, several fundamental questions arise with respect to the *Business Week* annual survey.

Table 1
ANNUAL CONSENSUS FORECASTS
BY ECONOMISTS AND ECONOMETRIC MODELS
FOR BUSINESS WEEK SURVEY
(In percent)

Year	Inflation rate			Real GNP growth			Unemployment rate		
		Consensus forecasts			Consensus forecasts			Consensus forecasts	
	Actual rates[1]	By economists	By models	Actual rates[2]	By economists	By models	Actual rates[3]	By economists	By models
1972 ...	4.2	3.4	3.3	7.3	5.7	5.7	5.6	5.4	5.5
1973 ...	7.5	3.4	3.3	3.4	6.0	6.1	4.9	5.0	5.1
1974 ...	11.0	6.1	6.0	-3.5	1.4	1.6	5.6	5.7	5.5
1975 ...	7.5	8.7	9.6	2.4	-1.2	-1.2	8.5	7.3	7.1
1976 ...	4.8	5.9	5.9	4.9	5.9	6.0	7.7	7.8	7.8
1977 ...	6.2	5.5	5.9	5.7	5.0	4.8	7.0	7.1	7.4
1978 ...	8.8	6.0	6.2	4.8	4.3	4.1	6.0	6.7	6.7
1979 ...	8.1	7.5	7.7	1.7	2.3	2.1	5.8	6.5	6.5
1980 ...	9.7	n.a.	n.a.	-.3	n.a.	n.a.	7.2	n.a.	n.a.
1981 ...	8.8	9.6	9.5	.8	.7	1.1	7.6	7.8	7.9

1. Percentage changes in the GNP implicit price deflator, fourth quarter to fourth quarter.
2. Fourth quarter to fourth quarter.
3. January-December averages.
n.a. — Not available.
SOURCES OF PRIMARY DATA: *Business Week.*
        U.S. Department of Commerce, Bureau of Economic Analysis.

• Are the forecasts—both judgmental and econometric—unbiased? That is, are they correct on average?

• Are the econometric forecasts of inflation, real GNP growth, and unemployment superior to the judgmental forecasts?

• How accurate are these published forecasts compared with "naive" forecasting methods that, for example, predict next year's inflation rate to be the same as this year's?

• Are the results of the *Business Week* survey significantly different from results of comparable surveys of macroeconomic forecasts?

This article answers these questions. Among several results reported, two stand out. Inflation is, by far, the least accurately predicted variable in the survey. Both groups of forecasters—judgmental and econometric—did very poorly. Second, there is little difference between the groups in overall forecasting performance. Before the analysis of this study is presented, some details of the *Business Week* survey need discussion.

A closer look at the data

From 1971 to the last survey period, the number of economists participating each year ranged from 24 to 33, while the number of econometric firms ranged from 7 to 11. Of the economists surveyed, there was little continuity in participation. Few participants carried over from year to year, and no one individual participated every year. As a result, it is impossible to evaluate the performance of any individual economist over the entire period. This would have been an interesting inquiry since a recent study found a moderate degree of consistency in the relative performance of individual forecasters participating in the American Statistical Association-National Bureau of Economic Research (ASA-NBER) quarterly macroeconomic survey.[1] Some scored well

1. Victor Zarnowitz, "Expectations and Forecasts from Business Outlook Surveys," NBER Working Paper Series, no. 845 (Cambridge, Mass.: National Bureau of Economic Research, January 1982). The ASA-NBER survey is discussed in more detail later.

above average with respect to several variables. It would have been interesting to ascertain whether certain judgmental forecasters performed significantly better because of greater skills or access to better information.

Similarly, there was substantial turnover in the econometric forecasting participants with the exception of Chase Econometrics, Data Resources, and Wharton Econometric Forecasting Associates, all three of which have participated from the beginning. Thus, the latest survey does allow the inspection of the performances of these econometric firms.

Another inquiry that is precluded by the content of the *Business Week* survey is the effect of the time horizon on the forecasting accuracy of the participants. In previous research studies, it has been found that, in general, forecast errors tend to increase as the target quarter progresses into the future.[2] Unfortunately, reporting of quarterly predictions in the *Business Week* survey only began with the 1980 forecast year, and the scarcity of data prohibits an adequate analysis of the time horizon issue.

For comparisons of judgmental and econometric forecasts, "consensus" forecasts were constructed by taking the averages of the forecasts from the respective groups. Table 1 reports these averages and the actual values for the variables under study. However, the way in which the survey results for 1980 (printed in December 1979) were reported prevents calculation of the consensus forecasts for that year.
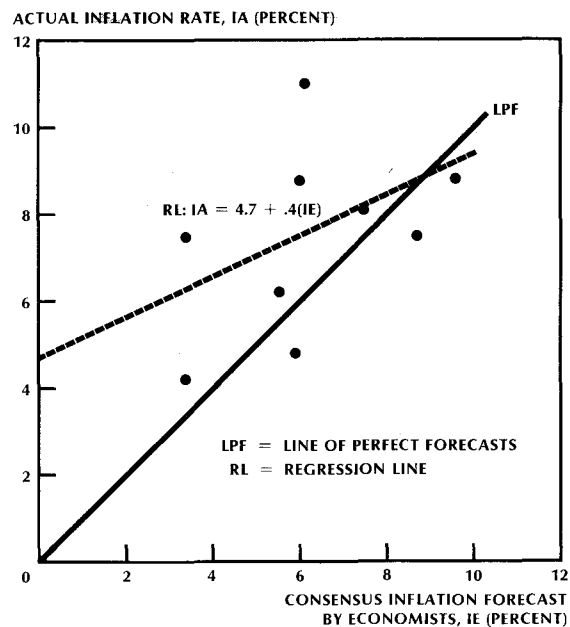
Unbiasedness of forecasts

A question that might be asked is, Are the predictions of economists correct on average? The same question can be asked about the econometric forecasts. A simple and useful graphical way of addressing this issue is a scatter diagram relating predictions to realizations—commonly referred to as the prediction-realization diagram.[3] Figure 1
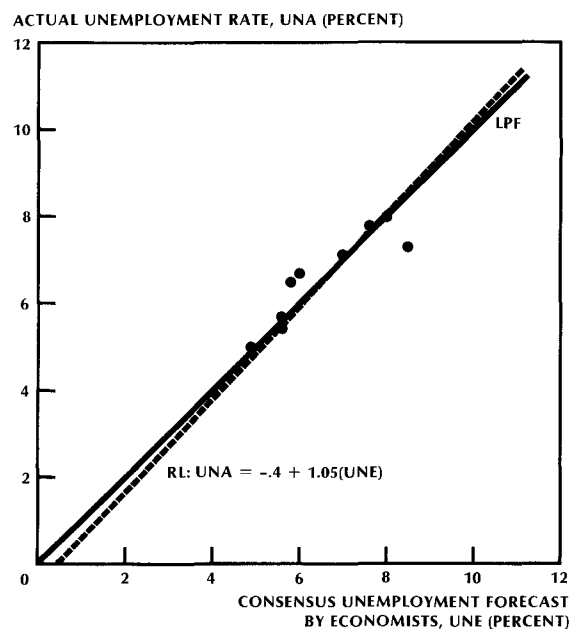
2. Ibid. Also see Stephen K. McNees, "The Forecasting Record for the 1970s," *New England Economic Review*, Federal Reserve Bank of Boston, September/October 1979, pp. 33-53.

3. Diagrams plotting predicted against actual values were first presented by H. Theil, *Economic Forecasts and Policy*, 2d rev. ed. (Amsterdam: North-Holland Publishing Company, 1961).

FIGURE 1

Prediction-Realization Diagram
for Economists' Predictions on Inflation
in the Business Week Survey

ACTUAL INFLATION RATE, IA (PERCENT)

RL: IA = 4.7 + .4(IE)

LPF = LINE OF PERFECT FORECASTS
RL = REGRESSION LINE

CONSENSUS INFLATION FORECAST
BY ECONOMISTS, IE (PERCENT)

Prediction-Realization Diagram
for Economists' Predictions on Unemployment
in the Business Week Survey

ACTUAL UNEMPLOYMENT RATE, UNA (PERCENT)

RL: UNA = -.4 + 1.05(UNE)

CONSENSUS UNEMPLOYMENT FORECAST
BY ECONOMISTS, UNE (PERCENT)

Table 2
RESULTS OF TESTS FOR UNBIASEDNESS
OF BUSINESS WEEK FORECASTS
BY ECONOMISTS AND ECONOMETRIC MODELS

Forecaster	Economic variables		
	Inflation rate	Real GNP growth	Unemployment rate
Economists (consensus) . . .	Biased[1]	Unbiased	Unbiased
	$a = 4.72, b = .43$	$a = .35, b = .81$	$a = -.40, b = 1.05$
	$F = 2.94 (p = .12)$	$F = .21 (p = .82)$	$F = .08 (p = .92)$
Econometric models			
Consensus	Biased	Unbiased	Unbiased
	$a = 5.04, b = .38$	$a = .38, b = .79$	$a = -.06, b = 1.00$
	$F = 3.15 (p = .11)$	$F = .23 (p = .80)$	$F = .08 (p = .92)$
Chase Econometrics	Biased	Unbiased	Unbiased
	$a = 5.42, b = .32$	$a = .60, b = .77$	$a = 1.47, b = .75$
	$F = 3.78 (p = .08)$	$F = .27 (p = .77)$	$F = .58 (p = .58)$
Data Resources	Biased	Unbiased	Unbiased
	$a = 4.64, b = .44$	$a = .22, b = .84$	$a = -.54, b = 1.07$
	$F = 2.95 (p = .12)$	$F = .18 (p = .84)$	$F = .12 (p = .89)$
Wharton Associates	Biased	Unbiased	Unbiased
	$a = 4.91, b = .39$	$a = -.05, b = .92$	$a = -.04, b = 1.00$
	$F = 3.08 (p = .11)$	$F = .12 (p = .89)$	$F = .01 (p = .99)$

1. The results for biasness in this test and the remaining inflation rate tests are only marginal: the probability levels observed range from .08 to .12.
NOTE: The maintained hypothesis of these tests requires that the error terms of the regressions be independent and normally distributed. The assumption of independence was tested by using the Durbin-Watson statistic adjusted for one gap (1980 observation missing). (See N. E. Savin and Kenneth J. White, "Testing for Autocorrelation with Missing Observations," Econometrica 46 [January 1978]:59-67.) In the inflation and GNP regressions, the null hypothesis of independence was accepted at the .05 level. In the unemployment regressions for Chase and Wharton, the null hypothesis of independence was accepted at the .01 significance level; the remaining unemployment equations had Durbin-Watson statistics that yielded inconclusive outcomes. The sample size was judged to be too small to provide a meaningful test for normality by means of traditional asymptotic tests, such as the Kolmogorov-Smirnov test.

presents prediction-realization diagrams for the economists' predictions on inflation and unemployment. The line of perfect forecasts, LPF, is a 45-degree line where the predicted values exactly equal the actual values of the variable under study. Of course, forecasts are rarely perfect, so the scatter of points representing unbiased forecasts will fall somewhere around this line. A regression line, RL, can be fitted through the scatter of prediction-realization points to obtain a visual characterization of the bias inherent in the predictions. As the diagrams show, the inflation predictions of the economists seem to be biased, while the unemployment predictions seem essentially to be correct on average.

To test for unbiasedness statistically, a standard two-variable regression model is specified:

$$A_t = \alpha + \beta P_t + e_t,$$

where A_t is the actual value of the economic variable at time t and P_t is the predicted value for the same period. The error e_t is assumed to be independent and identically distributed as a normal random variable with a zero mean and constant variance. The error term is included to represent the reality that not all predicted and realized values will fall on one line. A test for unbiasedness consists of testing the statistical hypothesis ($\alpha = 0, \beta = 1$). If this hypothesis is accepted, the data support the contention of unbiasedness; otherwise, the predic-

tions are not correct on average, and bias in prediction is implied.[4] Table 2 summarizes the tests for the *Business Week* forecasts.

When considering the prediction of real GNP growth and unemployment, all forecasts—the consensuses of the economists and econometric models as well as the models of Chase, Data Resources, and Wharton—appear to be unbiased. However, with respect to inflation, all forecasters did poorly. Though the price forecasts are only marginally biased at conventional levels of significance, they are certainly biased compared with the real GNP growth and unemployment predictions. By and large, everyone tended to underestimate inflation, especially when inflation moved to post-World War II highs. Such persistent underestimation is the root of many present-day phenomena, such as the recent financial plight of the savings and loan industry.[5] In the accompanying box, some plausible explanations for this phenomenon are assessed.

Precision of the forecasts

Even if forecasts are biased, they may not necessarily be "bad." Forecasts that consistently miss by only a small amount may be better than forecasts that are unbiased but very erratic. The mean square errors of forecasts allow comparisons taking this consideration into account. The mean square error of forecasts is defined by:

$$MSE = \sum_{t=1}^{n} [(A_t - P_t)^2]/n,$$

where A_t and P_t represent the actual and predicted values at time t and $\sum_{t=1}^{n}$ denotes the summation of the squared errors of forecasts for all observations, n.

A more popular measure of the trade-off between bias and variance is the root mean square error of forecasts, defined by:

$$RMSE = \sqrt{MSE}.$$

4. This methodology was proposed in Jacob Mincer and Victor Zarnowitz, "The Evaluation of Economic Forecasts," in *Economic Forecasts and Expectations: Analyses of Forecasting Behavior and Performance*, ed. Jacob Mincer (New York: National Bureau of Economic Research, 1969), pp. 3-46.

5. See Vanessa Bush and Mary Nowesnick, "It's Time to Clean Out Those Old Mortgages," *Savings and Loan News*, October 1981, pp. 44-51, for a description of the effects and history of falling asset values in the savings and loan industry.

The square root of the mean square error is convenient because its scale is the same as the scale of the variable being forecast. Another measure of forecasting accuracy is the mean absolute error:

$$MAE = \sum_{t=1}^{n} |A_t - P_t|/n,$$

where $|A_t - P_t|$ represents the absolute value of the difference. Both RMSE and MAE have been used to judge forecasting accuracy. The choice between these measures should be based on whether large errors should carry more than a proportional penalty. The RMSE criterion penalizes large errors more severely than does the MAE criterion.

In addition to comparing the judgmental (economists') forecasts and the econometric forecasts for accuracy, both of these should be compared with "naive" forecasting methods, which disregard the use of economic information altogether. If the naive methods perform better than the economics methods, the performance of the economics methods must be judged as being very poor. But if the naive methods are inferior, the use of economic insights would seem useful.

The naive methods chosen for this test are (1) the "no-change" rule, (2) the "last-period's-change" rule,. and (3) the "autoregressive" rule. Mathematically, these naive forecasting rules are expressed as:

(1) $P_t = A_{t-1}$;

(2) $P_t = A_{t-1} + (A_{t-1} - A_{t-2}) = 2A_{t-1} - A_{t-2}$;

(3) $P_t = a + b A_{t-1}$.

The no-change rule simply uses last period's value as the best guess. The last-period's-change rule modifies last period's value by the amount of change over the preceding period. The autoregressive rule involves fitting a linear autoregression of the actual value on its immediately preceding value:

$$A_t = \alpha + \beta A_{t-1} + e_t.$$

The coefficients a and b represent the ordinary least squares estimates of α and β.

The no-change rule and last-period's-change rule are easily implemented period by period since these rules do not involve parameters that must be estimated from the data. The autoregressive rule, however, requires estimation of the parameters α and β. The most stringent test using the naive autoregressive rule would be to compute a and b using the actual data from 1972 to 1981 and then compare the root mean square errors with those of

How Expectations of Economists Were Formed

Why did economists and econometric models do so poorly in forecasting inflation rates in the 1970's? During this period, inflation was trending upward while predictions were generally below the actual rates. One explanation might be that economists formed their expectations in an adaptive manner. Let A_t be the actual rate of inflation at time t and P_t be the predicted rate of inflation for the same period. The adaptive expectations hypothesis states that expectations are formed as follows:

$$P_t = \lambda A_{t-1} + (1 - \lambda)P_{t-1}, \quad 0 \le \lambda \le 1.$$

The forecast for period t is formed as a linear combination of last period's actual inflation rate and last period's predicted value. This model leads to the conclusion that forecasters will underestimate the rate of inflation if inflation is accelerating.

To determine whether the adaptive expectations hypothesis is a reasonable characterization of the way in which expectations were formed in the 1970's, the following regression model was specified for the economists' consensus forecasts:

$$P_t = \alpha_0 + \alpha_1 A_{t-1} + \alpha_2 P_{t-1} + e_t.$$

The theory of adaptive expectations specifies that $(\alpha_0 = 0)$ and $(\alpha_1 + \alpha_2 = 1)$. The estimated equation is:

$$P_t = 1.1854 + .6412 A_{t-1} + .0704 P_{t-1},$$
$$\quad (1.1029) \quad (.1304) \quad\quad (.1682)$$

with the figures in parentheses being the standard errors of the estimates. The Durbin h value is 0.39. A joint F test was conducted and was found to support the adaptive expectations hypothesis. Apparently, the economists formed their expectations in an adaptive way.[1]

Assuming that the economists participating in the *Business Week* survey had access to yearly inflation figures for 1948 to the year they were surveyed, it appears that the economists, as a group, did not form their expectations optimally. The adaptive expectations mechanism is rational (unbiased and efficient) if inflation can be characterized as a random walk:

$$A_t = \alpha + A_{t-1} + e_t,$$

where e_t is an independent and identically distributed unobservable random error with zero mean and finite variance.[2] However, statistical tests of the series preceding each survey do not support the contention that inflation should have been perceived as a random walk.

1. Other studies have similarly found that inflationary expectations appear to have been formed in an adaptive manner. For example, in "Inflationary Expectations: Their Formation and Interest Rate Effects," *American Economic Review* 66 (March 1976):124-31, Kajal Lahiri reports the estimates $(\alpha_1 = 0.534)$ and $(\alpha_2 = 0.426)$ for the 1952-70 period. In "Empirical Evidence on the Formation of Price Expectations," *Journal of the American Statistical Association* 65 (December 1970):1441-54, Stephen J. Turnovsky reports the estimates $(\alpha_1 = 0.226)$ and $(\alpha_2 = 0.781)$ for the 1962-69 period. In both studies the sum of the coefficients, $(\alpha_1 + \alpha_2)$, was very close to unity, thus supporting the contention that expectations were formed in an adaptive manner.

2. John F. Muth, "Optimal Properties of Exponentially Weighted Forecasts," *Journal of the American Statistical Association* 55 (June 1960):299-306.

---

the judgmental and econometric forecasts. This test is stringent (that is, the rule is difficult to beat) because the regression chooses the best fit after the fact; information that was not available to a forecaster at the time is used to construct a forecasting rule for comparison. Monday morning quarterbacks always perform well.

A less stringent test would involve estimating the parameters of the autoregressive rule on data preceding the forecast period and applying the estimated rule in forecasting the next period's value. Once the actual value is observed, the autoregressive rule is reestimated and used anew for the following period's forecast. In this less stringent form of the naive autoregressive rule, the forecaster is not allowed perfect hindsight but, rather, continually updates forecasts as new data become available. Because of the lack of hindsight, the root mean square errors of such a procedure are almost surely greater than those in the *ex post* case.

For brevity the former method of implementing the autoregressive rule is called the *ex post* autoregressive rule, and the latter method the *ex ante* autoregressive rule. If judgmental or econometric forecasts beat the *ex post* autoregressive rule, they are most certainly adequate. If they do not, judgment should be reserved until a comparison is made using the *ex ante* rule. Outperform-

Table 3
**RELATIVE ACCURACY OF BUSINESS WEEK FORECASTS
AND NAIVE FORECASTING METHODS**

Method	Inflation rate		Real GNP growth		Unemployment rate	
	Root mean square error	Mean absolute error	Root mean square error	Mean absolute error	Root mean square error	Mean absolute error
Business Week forecasts						
Economists (consensus)...	2.569	2.025	2.394	1.750	.559	.400
Econometric models						
Consensus............	2.646	2.088	2.475	1.850	.637	.488
Chase Econometrics ...	2.820	2.200	2.409	1.988	.933	.700
Data Resources	2.526	2.063	2.368	1.688	.579	.450
Wharton Associates....	2.600	2.213	2.177	1.700	.603	.413
Naive forecasts						
No-change rule..........	2.567	2.325	3.810	3.138	1.210	.925
Autoregressive rule						
Ex post...............	1.717	1.283	2.715	2.133	1.044	.833
Ex ante...............	3.113	2.378	9.356	2.522	1.519	.833

ing the ex *ante* rule indicates the forecasters are better than the autoregressive rule when only previous data are available. The results of analyzing the forecasting accuracy for inflation, real GNP growth, and unemployment are summarized in Table 3. Though the last-period's-change rule was examined, it never performed well. It is not discussed further.

For all the variables — inflation, real GNP growth, and unemployment — the root mean square errors and mean absolute errors for the economists' consensus forecasts were smaller than those of the consensus forecasts of the econometric models. The consensus judgmental forecasts were more accurate than the consensus econometric forecasts. This is not to say that individual econometric firms did not prevail over the judgmental forecasts for selected variables and measures of accuracy. For example, the price forecasts by Data Resources were better than either the consensus forecasts of the economists or the other econometric models when root mean square errors were considered. With regard to mean absolute error, the Data Resources forecasts were worse than the consensus for the economists. However, as found in other studies,

no major econometric forecaster overshadows judgmental forecasters or other econometric forecasters for all or even most macroeconomic variables.[6]

The price forecasts of all economics methods, judgmental as well as econometric, were relatively poor. The no-change rule was as good as any economics forecast — except that by Data Resources — when the RMSE criterion was considered. The ex *post* autoregressive rule surpassed all forecasts on both counts by a substantial margin. However, economists and models did outperform the ex *ante* autoregressive rule and in that sense were successful. All forecasters were slow to learn that record inflation rates were easily achieved and likely to be broken in the near term. With regard to forecasts of real GNP growth and unemployment, forecasters did quite well. All had substantially smaller forecasting errors than any of the naive

6. McNees, "Forecasting Record for the 1970s," and Victor Zarnowitz, "An Analysis of Annual and Multiperiod Quarterly Forecasts of Aggregate Income, Output, and the Price Level," *Journal of Business* 52 (January 1979):1-33.

Table 4
COMPARATIVE ANALYSIS OF CONSENSUS
FORECASTS OF ECONOMISTS PARTICIPATING
IN ASA-NBER AND BUSINESS WEEK SURVEYS

	Tests for unbiasedness of forecasts		
	Inflation rate	Real GNP growth	Unemployment rate
ASA-NBER survey	Biased	Unbiased	Unbiased
	$a = 5.10, b = .39$	$a = -2.02, b = 1.42$	$a = -.45, b = 1.04$
	$F = 3.36\ (p = .09)$	$F = 1.18\ (p = .36)$	$F = 1.10\ (p = .38)$
Business Week survey . . .	Biased	Unbiased	Unbiased
	$a = 4.72, b = .43$	$a = .35, b = .81$	$a = -.40, b = 1.05$
	$F = 2.94\ (p = .12)$	$F = .21\ (p = .82)$	$F = .08\ (p = .92)$

	Report of root mean square errors and mean absolute errors		
	Inflation rate	Real GNP growth	Unemployment rate
ASA-NBER survey	RMSE = 2.614	RMSE = 1.836	RMSE = .354
	MAE = 1.889	MAE = 1.378	MAE = .256
Business Week survey . . .	RMSE = 2.569	RMSE = 2.394	RMSE = .559
	MAE = 2.025	MAE = 1.750	MAE = .400

	Tests for significant differences in forecasting accuracy		
	Inflation rate	Real GNP growth	Unemployment rate
Granger-Newbold statistic (r)	$r = -.319$	$r = .481$	$r = .621$
	$(p = .18)$	$(p = .08)$	$(p = .03)$[1]

1. A low probability value indicates a significant difference in the root mean square errors of the two surveys.

methods. These results coincide with studies elsewhere.[7]

To determine if the differences in the accuracies of the consensus judgmental forecasts and the consensus and individual-firm econometric forecasts are statistically significant, a statistical test is needed. The test adopted here is the Granger-Newbold test of the equality of the population RMSEs of two alternative forecasting methods.[8] Tests for significant differences in the forecasting accuracy of the consensus judgmental and econometric forecasts

found none. The same results hold for comparisons between individual-firm econometric forecasts and other methods. Contrary to what might have been supposed initially, the econometric methods seem to offer no distinct advantages over judgmental forecasting methods.

Comparison with another survey

Among other surveys of economists that provide periodic forecasts of important macroeconomic variables is the ASA-NBER survey mentioned earlier, the quarterly Business Outlook Survey. Comparisons

7. McNees ("Forecasting Record for the 1970s") found that inflation rates are generally more difficult to forecast than other economic variables. Also, as Zarnowitz ("Expectations and Forecasts from Business Outlook Surveys") notes, tests of unbiasedness and accuracy are very unfavorable for expectations of inflation, but they show the forecasts of other variables generally in a much better light.

8. C. W. J. Granger and Paul Newbold, Forecasting Economic Time Series (New York: Academic Press, 1977), p. 281. Strictly speaking, the test by Granger and Newbold applies only to unbiased forecasting methods. Thus, comparisons between inflation forecasts are not strictly valid since they have tentatively been judged to be biased in previous tests here.

between economic surveys are tenuous at best, however. Forecasts published from competing surveys are not always prepared at the same time, may involve differing definitions for data, and may be for different spans of time. Also, the timeliness of surveys differs. The quicker survey results get into print, the more timely they will be.

Even with these caveats, an important question still remains, How accurate are the *Business Week* forecasts—in particular, the economists' forecasts—compared with those of economists polled in other surveys? Comparable forecasts were extracted from ASA-NBER November surveys and examined for 1972 through 1979 and for 1981, with 1980 being intentionally omitted. The forecasts selected were the median forecasts for inflation (measured by the GNP price deflator), real GNP growth, and unemployment. The inflation and real GNP growth projections were calculated from fourth-quarter values, while the average unemployment for a given year was calculated as the average of the unemployment projections for the four quarters of the year.

The results of some comparative analyses are contained in Table 4. In the top panel the tests for unbiasedness of the forecasts are summarized. As in the *Business Week* survey, the ASA-NBER forecasts for real GNP growth and unemployment appear to be unbiased, while the forecasts of inflation were biased and generally underestimated actual inflation. As indicated in the other panels of the table, the ASA-NBER survey seems to have an edge in forecasting real GNP growth and unemployment, with smaller RMSE and MAE measures of forecasting accuracy. The *Business Week* survey performed somewhat better in forecasting inflation as far as RMSE is concerned but worse considering MAE. Applications of the Granger-Newbold test statistically support these contentions.

Given available information, the conditions leading to the difference in performance are not easily pinpointed. Some possibilities are that the ASA-NBER survey may have more respondents and less turnover, or the survey may have a later response date, which would allow its economists additional time to note revisions of Government data. Though the lag between the time the two surveys are conducted and the time they appear in print is uncertain, the *Business Week* survey is likely to be published more quickly because the number of

economic variables it covers is less extensive and, too, the magazine is a weekly publication. Furthermore, the *Business Week* survey is more readily available to the public.[9]

Summary

As evidenced in the *Business Week* annual survey, forecasting inflation is a very difficult task. Both economists and econometric models generally underestimated inflation during the 1970's. This phenomenon is consistent with our finding that economists seem to have formed their expectations in an adaptive manner (see the accompanying box). Though the consensus inflation forecasts in the *Business Week* survey were poor, none of the naive rules examined here proved to be significantly superior. In contrast, our statistical analysis shows the forecasts of real GNP growth and unemployment in a much more favorable light. Both the judgmental and econometric-model forecasts were unbiased and performed better than any naive forecasting method examined.

Our statistical tests show that in the *Business Week* survey, there is no statistically significant difference between judgmental and econometric forecasts. For the period examined the judgmental forecasts were somewhat better. However, in the next decade these results may be reversed, though differences in forecasting performance would not be expected to be great.

The lack of a significant difference in forecasting performance is understandable, given the environment in which these forecasts are made. The distinction drawn here between judgmental and econometric forecasts is probably more apparent than real. Economists participating in the *Business Week* survey may not specifically use econometric models in their forecasting, but they are almost certainly aware of what the econometric models are predicting and are probably influenced. Similarly, the econometric-model forecasters are aware of the judgmental forecasts of the economists and are likewise influenced. The communality between economists and econometric-model forecasters is

9. Currently, the ASA-NBER survey is available in two publications, the *NBER Reporter* and *Amstat News*—both of which are issued irregularly and are not always available in public libraries.

quite strong. For example, in the *Business Week* survey, every time the econometric models predicted an increase or decrease in inflation, the economists as a group predicted the same. An interesting comparison would be one between forecasts of economists who did not have access to econometric-model benchmarks and forecasts of economists who did.

Although there were instances where individual econometric forecasts were superior to the judgmental forecasts, no single econometric forecaster dominated the predictions for all or even most variables. With respect to the *Business Week* survey, econometric models seem to have no advantage when forecasting broad economic aggregates. This conclusion could be different, of course, when considering microeconomic forecasts at the industry and firm levels.

Finally, when the *Business Week* survey is compared with a competitor, the ASA-NBER survey seems to be slightly better in forecasting real GNP growth and unemployment. However, if forecasts of inflation are considered and timeliness and accessibility are important, the *Business Week* survey has certain advantages.

Commercial Banking as a Line of Commerce: Time for Change?

By F. Jay Cummings*

Many commercial banking organizations have grown rapidly in recent years through aggressive programs to acquire additional banks. Since bank mergers are scrutinized using antitrust standards, the manner in which those standards are applied affects the type and amount of growth observed.

In its landmark *Philadelphia National Bank* decision almost 20 years ago, the Supreme Court established the applicability of Section 7 of the Clayton Act to commercial bank mergers.[1] Section 7 requires a merger to be declared illegal if its effect "may be substantially to lessen competition" in any line of commerce (product market) in any section of the country (geographic market). In order to analyze competitive effects, the Court defined a single relevant product market—commercial banking—and described the geographic market as local in character.

The *Philadelphia* decision set the tone for analysis of later bank mergers by supporting the use of structural tests in evaluating anticompetitive effects:

> This intense congressional concern with the trend toward concentration warrants dispensing, in certain cases, with elaborate proof of market structure, market behavior, or probable anticompetitive effects. Specifically, we think that a merger which produces a firm controlling an undue percentage share of the relevant market, and results in a significant increase in the concentration of firms in that market, is so inherently likely to lessen competition substantially that it must be enjoined in the absence of evidence clearly showing that the merger is not likely to have such anticompetitive effects.[2]

The suggested structural test uncovers anticompetitive effects if a sizable market share can be equated with market power and if increased concentration in a highly concentrated market can be equated with a greater likelihood of collusive or oligopolistic behavior.

* F. Jay Cummings is an associate professor of economics at the University of Texas at Dallas and a consultant at the Federal Reserve Bank of Dallas.

1. *United States v. Philadelphia National Bank*, 374 U.S. 321 (1963).
2. *United States v. Philadelphia National Bank*, 374 U.S. 321, 363 (1963).

Using a single product line to examine structural effects requires selection of a proxy, or summary measure, for commercial bank output. In the *Philadelphia* tradition, mergers are generally analyzed today by using commercial bank deposits as output proxies. Mergers involving banks that have sizable shares of a local area's deposits and resulting in an increase in deposit concentration in highly concentrated markets are considered anticompetitive according to the Court's standards.[3] The structural approach has been codified in the merger guidelines of the U.S. Department of Justice.

Despite criticism of its single product market conclusion and its use of a "shortcut" structural test, the Court has since maintained its position in later opinions, most importantly in its *Phillipsburg National Bank* and *Connecticut National Bank* decisions.[4] Although some Federal district courts and bank regulatory agencies have increasingly deviated from the strict standards, the Supreme Court precedent remains: Commercial bank mergers that produce undesirable structural effects are anticompetitive.[5] The structural effects are examined under the assumption that bank services constitute a single line of commerce within local geographic markets.

This article examines the concept that commercial banking can be treated as a single line of commerce defined by a unique cluster of bank services. While increasing competition from nonbank institu-

tions is often correctly cited as a reason for questioning the single-line-of-commerce approach to bank merger analyses, the importance of the Court's concern for clustering of bank services has not been adequately addressed.

Consideration of the way bank services are provided in packages to different classes of customers indicates not one but several distinct lines of commerce in banking, only some of which have significant nonbank competitors. Viewed from this perspective, strict adherence to the Supreme Court's test is likely to lead to approval of some anticompetitive mergers and to disapproval of some harmless mergers. Recent changes in regulations and technology make changes in merger analyses even more important than before.

This article contends that a multiple-line-of-commerce approach, based on disaggregation of bank services by customer groups, is preferable to the single-line-of-commerce strategy. While immediate implementation of the disaggregated approach would certainly burden the regulatory agencies and the merger applicants analytically, the logical merits of the approach argue strongly for moving in that direction.

The question of a cluster of services

In agreeing with the district court in its *Philadelphia* decision, the Supreme Court found "the cluster of products (various kinds of credit) and services (such as checking accounts and trust administration) denoted by the term 'commercial banking' " sufficient to define a distinct line of commerce.[6] The prevailing statutory and regulatory separation of commercial banks from other financial institutions aided the courts in reaching this conclusion, but the courts' acceptance of the distinctiveness of banking requires further examination.

3. Philadelphia National Bank had approximately 21 percent of commercial bank deposits in the Philadelphia standard metropolitan statistical area at the time of the decision, and its market share would have become 36 percent after the merger. The two-firm deposit concentration ratio would increase from 43 percent to 58 percent, and the top four banks would control 77 percent of the deposits as a result of the merger. These structural effects were deemed sufficient to deny the merger.

4. *United States v. Phillipsburg National Bank & Trust Company*, 399 U.S. 350 (1970); *United States v. Connecticut National Bank*, 418 U.S. 656 (1974). Selected criticisms of the Supreme Court position include: Joel M. Yesley, "Defining the Product Market in Commercial Banking," *Economic Review*, Federal Reserve Bank of Cleveland, June/July 1972, pp. 17-31; Richard W. Stolz, "Philadelphia National Bank Case Revisited," *Ninth District Quarterly*, Federal Reserve Bank of Minneapolis, Winter 1977, pp. 5-11; and Michael E. Friedlander and John H. Slayton, "Determination of the Relevant Product Market in Bank Mergers: A Time for Reassessment?" *Business Lawyer* 36 (July 1981):1537-55.

5. Selected examples of these deviations include the following district court decisions: *United States v. Provident National Bank*, 280 F. Supp. 1 (E.D.Pa. 1968); *United States v. First National Bank of Jackson*, 301 F. Supp. 1161 (S.D.Miss. 1969); and *United States v. First National State Bancorporation*, 499 F. Supp. 793 (D.N.J. 1980). Also included are the following Federal Reserve Board decisions: *Northeast Bancorp, Inc.*, 60 *Federal Reserve Bulletin* 375 (1974); *First Bancorp of N.H., Inc.*, 64 *Federal Reserve Bulletin* 967 (1978); and *Republic of Texas Corporation*, 67 *Federal Reserve Bulletin* 57 (1981).

6. *United States v. Philadelphia National Bank*, 374 U.S. 321, 356 (1963).

Perhaps the clearest statement of the judicial meaning of the concept of a "unique" cluster of products and services offered by commercial banks was provided by the district court:

> It is the conglomeration of all the various services and functions that sets the commercial bank off from other financial institutions. Each item is an integral part of the whole, almost every one of which is dependent upon and would not exist but for the other.[7]

The Supreme Court accepted as evidence to support this view a set of examples. These examples were designed to establish that certain banking services were unique to banks as a result of regulation (such as checking accounts), that some services enjoyed cost advantages that insulated banks from competition with other financial institutions (such as small loan companies in the personal loan market), and that other services were insulated by a settled consumer preference (such as savings deposits).[8]

While the importance of the examples cited can be questioned, the Court's approach was intended to demonstrate the distinctiveness of certain specific bank services arising from regulation or the lack of realistic competitive alternatives. The Court did *not* examine whether the appropriate incentives existed for commercial banks to offer or for bank customers to purchase a bundle, or package, of distinctive services. Thus, the Court's examples merely suggest, at best, that the various products a bank offers are exposed to varying numbers of competitive alternatives, ranging from none to a reasonably large number.

In the *Philadelphia* case and later cases, the courts failed to justify their view of banking as a single line of commerce described by the cluster of services provided. However, their concern for bundling of bank services is valid. Defining a relevant product market requires examination of consumer and producer behavior. While highly substitutable products on the demand side should be combined in describing a product market, complementary relationships and conditions of supply cannot be ignored.

For example, the vast majority of consumers consider left and right shoes to be perfect complements, buy them as a package, and would continue to do so even if separate purchases were an alternative. Suppliers react to this complementarity and, perhaps, experience production and distribution cost savings by selling shoes in pairs. Market provision of shoes in pairs suggests that the cost savings of producing and selling the package and the convenience that many consumers derive from pair purchases are sufficient to overcome the value of separate purchases for some consumers.

While the shoe example is extreme, it is instructive. Incentives for packaging, whether supply- or demand-oriented, are important in defining product markets. The extent of bundling or joint purchases should be considered in determining the relevant product lines in bank merger analyses.

Three alternative economic explanations for supplying a cluster of bank services appear plausible. First, a commercial bank may produce what can be viewed as joint products. (Examples of some commonly known joint products are milk and butter, beef and hides, and cotton and cottonseed oil.) Increasing the output of one product implies an increase in the output of the other. Second, a commercial bank may be a multiproduct firm.[9] Although a bank supplies a variety of products, it may have economic incentives, perhaps spurred by the statutory and regulatory environment, to sell its products and services as a package to customers.

A third explanation for packaging is based on demand factors rather than supply conditions. Even if a bank sells its services separately, joint purchases by customers may occur as a result of cost savings in carrying out multiple transactions at one location.

9. The commercial bank as a multiproduct firm is a premise of many economic models of bank behavior. See, for example, Bernard Shull, "Commercial Banks as Multiple-Product Price-Discriminating Firms," in *Banking and Monetary Studies*, ed. Deane Carson (Homewood, Ill.: Richard D. Irwin, 1963), pp. 351-68, and David A. Alhadeff, "Monopolistic Competition and Banking Markets," in *Monopolistic Competition Theory: Studies in Impact*, ed. Robert E. Kuenne (New York: John Wiley & Sons, 1967), pp. 357-78.

7. *United States v. Philadelphia National Bank*, 201 F. Supp. 348, 363 (E.D.Pa. 1962).
8. *United States v. Philadelphia National Bank*, 374 U.S. 321, 356-57 (1963).

The variety of bank service packages and prospects for change

The image of "full-service" banks is suggestive of service bundling, but what incentives do banks have to package their services? One incentive results from the prohibition of interest payments on demand deposits and the ceilings on interest rates payable on savings and time deposits (Federal Reserve Regulation Q).[10] To attract interest-free demand deposits or low-interest time deposits, banks may offer a bundle of services and cut the prices of existing or new services to make the package price attractive enough to overcome customer resistance to holding assets in a non-interest-bearing or low-interest-earning form. Preference for depositor-borrowers or the formal requirement of compensating balances fits this evasion-of-price-controls argument.[11] A different twist on the same explanation may prevail when market loan rates exceed usury ceilings; those ceilings can be evaded by requiring interest-free compensating balances.[12]

Motivation for bank service bundling may be derived from cost savings in providing the package. Flexibility in resource use across numerous activities may lead to more efficient resource use and resulting cost savings. The volume of bank activity when numerous services are provided may lead to reduced cost of information storage and processing through application of more sophisticated computer technologies that become economical only after the bank becomes large. Economies in marketing and advertising may be realized in offering a set of related services to customers and potential

customers.

In a different vein, statutory restrictions on price competition cause banks to compete on a nonprice basis. Offering a range of services may provide the convenience of one-stop banking to customers and reduce their search and commuting costs. Convenience considerations may lead consumers to buy a package of services at one location. From a bank's perspective, lengthy and stable customer relationships may be promoted by offering a set of complementary financial services.

Today's changing economic and regulatory environment undoubtedly has prompted substantial changes in bundling and joint purchase incentives. Traditional packages for some customers may be disappearing, but other packages may be developed in response to improved technologies and new competition.

Various provisions of the Depository Institutions Deregulation and Monetary Control Act of 1980 will apparently reduce many of the sources of incentives for service bundling previously discussed. The expansion of interest-bearing checking deposits to all depository institutions, the phaseout of Regulation Q, and the easing of usury limits will decrease the importance of packaging to evade price controls. Preferences for depositor-borrowers will not be motivated by interest rate restrictions on deposits. Incentives for banks to price individual services are being furthered by the requirement that the Federal Reserve System price the services it provides banks.

The dramatic growth of financial services of nonbank institutions, such as money market funds, began in the late 1970's and is continuing. These institutions, unhindered by reserve requirements and geographic limitations on the scope of their activities, have discovered that interest-sensitive

---

10. Following an evasion-of-price-controls rationale for tying arrangements, this argument is developed, for example, by Franklin R. Edwards, "Tie-in Sales in Banking and One Bank Holding Companies," *Antitrust Bulletin* 14 (Fall 1969):587-605. The evasion-of-price-controls argument and other motivations for tying arrangements are presented in M. L. Burstein, "A Theory of Full-Line Forcing," *Northwestern University Law Review* 55 (March-April 1960):62-95, and in Eugene M. Singer, *Antitrust Economics: Selected Legal Cases and Economic Models* (Englewood Cliffs, N.J.: Prentice-Hall, 1968), pp. 187-95.

11. Consistent with this incentive is the observation by Garvy and Blyn that prior to 1933, banks generally paid interest on demand deposits and the requirement of compensating balances in the conventional sense was not widespread. See George Garvy and Martin R. Blyn, *The Velocity of Money* (New York: Federal Reserve Bank of New York, 1970), p. 32.

12. The loan-deposit packaging may also be explained by a risk-sharing argument for tying arrangements. For example, a reduction in the loan rate and the requirement that borrowers hold demand deposits may enable a bank and its borrowers to share in the risks associated with changing market loan rates. If market rates rise during the term of the loan, a bank would lose less on funds previously committed at lower loan rates when it also requires the borrower to hold interest-free deposits. In contrast, the size of the gain to a bank that is derived from having loans outstanding at higher interest rates when market rates fall is smaller when compensating balances are required since the requirement is conditioned on a lower loan rate at the time the loan is arranged.

"bank" customers can be attracted to the individual financial services provided, especially with high market interest rates. The rising commuting and search costs of customers and the inconvenience of "multistop" banking are increasingly offset by the ability of nonbank institutions to develop instruments offering attractive yields.

While many of the incentives for packaging services or joint purchase differ today, traditional bundling of services for certain customers may still persist. Thrift institutions have been granted additional deposit and asset powers by the Deregulation and Monetary Control Act. However, the increased asset powers may not be extensively exercised in the near future if thrifts do not have the requisite expertise and resources to exploit their new opportunities. Most important for the discussion here, increased thrift competition will primarily affect the retail side of the financial services market. Commercial banks retain their role as sole suppliers of demand deposit services for commercial customers and have broader lending powers in satisfying the demands of this group.[13] To the extent that commercial banks retain a distinctive role with respect to commercial customers, or a segment of that class of customers, traditional service packaging may continue.[14]

While evasion of price controls cannot be used as an explanation for compensating balance requirements if interest rate restrictions are eliminated in the future, banks may still favor established depositors in their lending decisions. Customers who hold deposits and purchase other bank services may provide banks valuable information on potential borrower characteristics through these relationships. Lower information costs in dealing with established

customers may lead banks to retain a preference for this group.

Banks may develop new packages of services in response to improved technologies and the increased competition from packages of nonbank financial services. Nonbank organizations are adding banking services to the sets of services they have historically provided. A number of banks, often assisted by securities firms, are responding by offering various types of asset management accounts that provide a single product to handle checks, savings deposits, credit-card transactions, stocks and bonds, and sometimes more.

Possible economies of scale in electronic fund transfer (EFT) systems may lead banks to offer new bundles of household services to encourage greater EFT usage with its associated cost advantages. In response to volatile interest rates and in an attempt to secure a stable source of funds in an increasingly competitive environment, banks in the future may tie revolving credit terms to the size and length of deposit relationships with customers or may offer large depositors other services at reduced prices.

Many banks, through the holding company device, are beginning to emphasize special packages of services. Banc One's processing of cash management accounts, First Tennessee National Corporation's overnight check processing, and Hawkeye Bancorporation's farm management services for absentee owners are examples. The perceived desire by many banking organizations to find specialized "niches" in the changing financial services industry is stimulated by the growth of the wide range of services offered by nonbank financial institutions.

13. The Deregulation and Monetary Control Act provided one exception to the commercial bank "monopoly" over commercial-customer demand deposit services in allowing Federal mutual savings banks to accept demand deposits in connection with a business loan relationship. The few Federal mutual savings banks may now hold up to 5 percent of their assets (with geographic restrictions) in commercial loans. While Federal savings and loan associations may engage in limited investment in commercial paper, they face an important constraint. Savings and loan associations must maintain at least 82 percent of their total assets in, and derive at least 75 percent of their income from, qualified assets (primarily mortgages) to receive favorable tax treatment under the Internal Revenue Code.

14. This observation is consistent with David Alhadeff's contention 15 years ago: "Significantly the tie-in sales are restricted to those services (business loans and transactions deposits) for which banks are the dominant or sole suppliers whereas bank services that nonbanks also supply (such as home-mortgage loans, consumer loans, and savings accounts) can usually be negotiated separately" ("Monopolistic Competition and Banking Markets," pp. 364-65).

Recent evidence also provides support for this observation. David D. Whitehead, "The Sixth District Survey of Small Business Credit," Economic Review, Federal Reserve Bank of Atlanta, April 1982, pp. 42-48. Interpreting survey responses, Whitehead reports "that business firms use multiple services of local banks. This is consistent with the view that businesses perceive commercial banks as offering a cluster of services" (p. 47).

An alternative framework for bank merger analyses

Clearly, no one package adequately describes the bank-customer relationship. Different banks concentrate on different types of packages, and different customers of the same bank often purchase entirely different packages. Lumping the various bundles together in merger analyses cannot be justified. Some sort of disaggregation is required.

Advocating disaggregation of commercial bank services is not based on a presumption that such an approach will lead to a smaller or larger number of merger approvals. Rather, the mix of merger approvals and disapprovals will likely be different. Most important, merger disapprovals will be based on a clearer identification of undesirable structural, and presumably anticompetitive, effects in relevant markets.

To illustrate the point that different structural effects will often be observed when the disaggregated product line approach is used, consider a hypothetical example with eight banks in a local market. Assume that in this market area, banks with sizable deposits concentrate a bigger portion of their business lending activity on loans to large businesses and that banks with smaller total deposits focus their lending activity on small businesses. It follows, then, that the ratio of small business loans to bank deposits declines as bank deposit size increases. Small business lending may be described as a local market (a proposition discussed later), whereas there may be a variety of financing alternatives for large business borrowers elsewhere.

The structural effects of mergers within the local market area can be examined from two perspectives: one using deposits as an output proxy for the single line of commerce and the other using small and large business loans as disaggregated product lines. If two banks, A and C in Table 1, merge, the acquiring firm's deposit market share rises from 33

percent to 47 percent, and the four-firm deposit concentration ratio increases 10 percentage points. The new bank's market position is strengthened by the merger according to the single product line approach, and the deposit-size distribution of banks becomes more unequal after the merger.[15]

Examination of small business loans as a relevant line of commerce in evaluating the merger of Banks A and C leads to quite different results. Market concentration does not change, the merged firm becomes a more significant small competitor in the small business loan market, and the inequality of market shares is reduced as a result of the merger.[16] Because of the many alternatives for large business borrowers, increased concentration in the relevant market for large business loans may be only trivially affected.

Similarly, if Banks F and G merge, the structural effects measured by deposits as a proxy for the single product line are not troublesome. Analysis of the merger based on small business loans as a line of commerce reveals a merger with structural effects that may be subject to challenge.

As a first step in defining relevant product markets, disaggregation of bank services is necessary. Disaggregation of commercial bank services could be accomplished by dividing services according to *product type* (demand deposit services, time deposit services, trust services, commercial and industrial loans, residential mortgage loans, and so forth) or according to *customer type* (individuals, partnerships, small corporations, large corporations, governments, and so forth). If practical concerns dictated the choice, the product-type approach seems appealing for two reasons. First, existing data may be more easily tailored to product breakdowns. Available balance sheet data provide information on

15. In examining structural effects of bank mergers, courts and regulatory agencies often describe the impact of the merger on the inequality of bank sizes. The coefficient of variation in market shares—the standard deviation of market shares divided by the mean market share—measures the variability of bank market shares. Thus, the magnitude of the coefficient of variation offers one measure of the inequality of bank sizes.

16. The courts will probably continue to consider that the market structure at the time of a merger and the changes in that structure are of utmost importance. However, structural effects may not always coincide with anticompetitive effects. In this example, even though Banks A and C are not active small business lenders, an increase in interest rates on small business loans may lead either or both of these banks to expand their lending and small business loan market shares, even without the merger. Consequently, examination of small business loan shares at the time of the merger would understate the competitive importance of these banks in the market for small business loans.

Table 1
STRUCTURAL EFFECTS OF TWO HYPOTHETICAL BANK MERGERS

Summary measure	Four-firm concentration ratio (Percent) Before merger	After merger	Acquiring bank Market share (Percent) Before merger	After merger	Market rank Before merger	After merger	Coefficient of variation in market shares Before merger	After merger
Merger of Bank A and Bank C								
Total deposits	78	88	33	47	1	1	.82	1.08
Small business loans	71	71	2	10	8	6	.53	.35
Merger of Bank F and Bank G								
Total deposits	78	78	7	10	6	5	.82	.70
Small business loans	71	83	17	31	3	1	.53	.73

Market data	Bank A	Bank B	Bank C	Bank D	Bank E	Bank F	Bank G	Bank H
				Millions of dollars				
Total deposits	100	60	40	35	30	20	10	5
Small business loans	.2	1.2	.8	2.1	2.1	1.8	1.5	.8

asset and liability categories useful for defining different types of products. Second, identification of bank "competitors" appears simpler, particularly since statutory and regulatory provisions reveal the types of institutions permitted to provide various products.

However, disaggregation of bank services by customer type is a preferable starting point because it allows attention to be focused on consumer service substitutability as a key determinant of market definitions.[17] The types of bank services purchased, as well as the range of competitive alternatives, differ among various customer classes. Negotiable order of withdrawal (NOW) accounts issued by thrift institutions may be substitutes for deposits with banks for individual customers; yet commercial customers are not permitted to hold funds on deposit with thrifts.

Similarly, short-term business loans may not be a single product market. The terms of identical loans to two customers may vary according to perceived risks, the nature and extent of customer relationships with a bank, and the competitive alternatives available to the two customers. One firm's greater mobility, reflected in a wider geographic range of available bank lenders and a broader set of borrowing sources, may (for purposes of merger analyses)

be sufficient to distinguish its borrowing activity from that of a less mobile firm. Thus, whether two bank services or a bank service and a nonbank service are substitutes, complements, or unrelated functions depends on the customer class examined.

Focusing on customer types in defining the lines of commerce has the advantages of identifying the group or groups of customers most likely affected by a particular bank merger as well as maintaining the spirit of the Supreme Court's concern for

17. While economists agree that highly substitutable consumer products should be included in the same product market, no commonly accepted cutoff on the degree of substitutability necessary for distinguishing various products has been established. Complicating matters is the fact that producer substitutability cannot be ignored, but it may conflict with consumer substitutability as a criterion for product market definitions. A common compromise is to emphasize demand-side substitutability in defining product markets and in examining structural effects. Supply considerations, then, enter the analysis as they affect the description of potential sources of entry.

See, for example, Brown Shoe Co. v. United States, 370 U.S. 294 (1962), and Kenneth W. Clarkson and Roger LeRoy Miller, Industrial Organization: Theory, Evidence, and Public Policy (New York: McGraw-Hill Book Company, 1982), pp. 52-58. This approach is also consistent with the spirit of U.S. Department of Justice, "Merger Guidelines," June 14, 1982, pp. 4-10.

packages of bank services. For example, the provisions of the Deregulation and Monetary Control Act will lead to more "new" competitive alternatives for individuals than for businesses. Business customers cannot benefit from the range of interest-bearing demand deposits now available for individual bank customers and may experience only limited advantages from the increased lending powers given to thrifts.

The customer-type disaggregation stresses the range of competitive alternatives available to different groups. This approach highlights the possibility that locally limited businesses, generally thought to be small businesses, will be the group most affected by a particular bank merger.

Among the reasons for presuming small firms to be less mobile are that their earnings experience may be shorter, earnings are often more variable, information on their operations is less precise, and their bargaining power is weaker than for larger companies. In addition, profitability of many small businesses may be closely tied to local economic conditions, which an out-of-town lender may find costly to evaluate. At the same time, the costs of transacting borrowing and other bank business with distant banks may be high for firms with localized operations. Larger firms are also more likely to establish relationships with more than one bank.[18] The importance of bank service bundling to small businesses in defining a relevant market for merger analyses can be examined if disaggregation by customer type is adopted.

Implementation of the new approach

Disaggregation of commercial bank services by customer class is conceptually appealing. However, a number of practical problems will arise in implementing the approach. Immediate application would be hindered by the unavailability of data assembled by customer group. Presumably, over time, this difficulty could be overcome, particularly if mandated by the courts.

It is tempting to use data problems as a reason for adopting a revised approach to bank merger analyses that is less demanding. For example, increasing thrift competition in a local area could be taken into account by reducing, or "shading," commercial bank deposit market shares and concentration ratios.[19] The rationale for this approach is that thrift services are partial, but not complete,

substitutes for bank services. However, shading retains the single-line-of-commerce convention. This approach cannot reveal the nature of competition between banks and thrifts since the extent of competition differs across customer types.

The problem of data availability is not unique to commercial bank mergers. In industrial mergers, market shares and market concentration are often based on modifications of existing data on product lines or geographic areas that are broader than desirable. Even available data on bank deposits, the proxy normally used for the single line of commerce, are faulty since they are based on bank office location rather than depositor location. In analyzing bank mergers, the cost of acquiring precise data for specific customer classes may necessitate use of data based on more broadly defined classes. Moving in the direction of disaggregation by customer classes, even in a limited way, remains logically preferable to adhering to the single-line-of-commerce approach.

18. Various surveys support these observations. See, for example, Clifton B. Luttrell and William E. Pettigrew, "Banking Markets for Business Firms in the St. Louis Area," *Review*, Federal Reserve Bank of St. Louis, September 1966, pp. 9-12; "Bank Markets and Services: Summary of Three Surveys of Bank Customers," *Business Conditions*, Federal Reserve Bank of Chicago, May 1967, pp. 6-10; Robert D. Bowers, "Businesses, Households, and Their Banks," *Business Review*, Federal Reserve Bank of Philadelphia, March 1969, pp. 14-19; and Cynthia A. Glassman and Peter L. Struck, "Survey of Commercial Bank Lending to Small Business," in *Studies of Small Business Finance*, A Report to Congress Prepared by the Interagency Task Force on Small Business Finance (Washington, D.C., February 1982).

The St. Louis survey reveals, for instance, that only 13 percent and 23 percent of small firms (net worth of less than $1 million) used more than one bank for loans and checking accounts, respectively. On the other hand, 44 percent and 68 percent of the large firms surveyed used more than one bank for loans and checking accounts, respectively. The recent Interagency Task Force survey examined banks' reasons for rejection of small business loan applications. Rejection criteria used at least occasionally by the surveyed banks included: too little owner's equity in the business (100 percent of the banks), insufficient quality of collateral (93 percent), poor earnings record (92 percent), new firm with no established earnings record (79 percent), questionable management ability (73 percent), and no established deposit relationship (53 percent).

19. The Supreme Court rejected "shading" as a relevant approach in the *Connecticut* case, *United States* v. *Connecticut National Bank*, 418 U.S. 656, 662 (1974).

The extent of disaggregation should reflect expected differences among customer groups in the types of bank services purchased, in the range of competitive alternatives, and in relationships between banks and customers. However, the degree of customer disaggregation must be tempered with practical data concerns. One possible classification would distinguish households, small businesses, and large businesses. A more comprehensive and demanding classification might separate individual customers other than businesses; small and large nonfinancial, nonagricultural businesses; small and large agricultural firms; financial institutions; and Federal, state, and local governments.

Once a decision on the classification of customers is made, establishing relevant product markets for each class requires judgment on the degree of substitutability of the various products typically purchased by members of the class. For example, do household customers regard bank-card loans as close substitutes for personal loans? Since answers to such specific questions are highly speculative, broad categorization is appealing. In the case of households, perhaps distinguishing deposits and broad loan categories, such as installment credit and mortgages, would suffice. Alternatively, examination of the largest source of business that a bank derives from each customer class could provide the basis for the analysis.

In cases where it is expected that bundling or joint purchase exists, there appears to be no reason to separate the services in defining relevant markets. On the other hand, there may be competitive alternatives for portions of the bundle. The possibility of purchasing components of the bank bundle separately suggests a need to analyze individual products. For example, while loan and deposit service packaging may still exist to a substantial degree for locally limited businesses, finance companies can offer loans (although not deposit services) to this group. In these instances, it is appropriate to examine two alternative product market definitions, based on (1) the dollar volume of loan and deposit services to small businesses at local commercial banks only and (2) borrowing by small businesses from commercial banks, finance companies, and other sources.

While such estimates are useful, additional quantitative and qualitative information should be sought. Surveys of banks and bank customers may help determine the importance of various customer characteristics as they might influence relevant product and geographic market definitions. Cost considerations may limit the desirability of extensive surveys, but banks themselves may collect a variety of useful data. For example, computerized records of business loans may include information on borrower characteristics, such as borrower size and location and loan terms. Analysis of these data can assist in identifying distinct classes of business customers. Examination of activities of corporate calling officers and of loan production offices may provide further insight into product and geographic definition issues.

Overall evaluation of competitive effects

Undesirable structural effects may very well be observed in certain product lines and desirable effects seen in others in using the new approach. In these instances, weights must be assigned to the various product lines to evaluate the merger's overall impact. Subjectivity is bound to enter such analyses, but subjectivity is not undesirable when hard-and-fast rules based on arbitrarily designated market shares and concentration ratios do not reflect market realities.
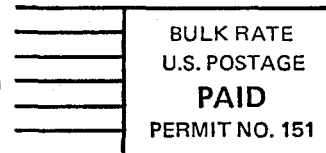
From a legal perspective, determining a merger's overall impact by weighing effects in several product lines is not necessary. Increased concentration in *any* product line, upon close reading of the Clayton Act (Section 7), would suggest that the merger should be declared illegal—on the grounds of a lessening of competition "in any line of commerce in any section of the country." This wording of the Clayton Act does, however, ease the task of employing the recommended approach. The bank regulatory agencies and the courts do not have to examine every conceivable customer class and product offered in analyzing the competitive effects of bank mergers.

If large corporations in an area transact their banking business on a nationwide basis and if individuals have available an increasing range of competitive financial services provided by nonbank financial institutions, there may be no reason to expect that a merger within a state of even two banking organizations of substantial size will have an impact on these groups. On the other hand, locally limited business customers may be affected. The change in the concentration of loans to small

businesses in a local geographic market, or the combined bank loan-deposit package, may provide useful insight into the merger's effect. If anti-competitive effects are found for this customer class, denial of the merger should be considered. If no negative effects are discovered, approval of the merger is more appealing.

FEDERAL RESERVE BANK OF DALLAS

STATION K, DALLAS, TEXAS 75222

ADDRESS CORRECTION REQUESTED