Economic Review

FEDERAL RESERVE BANK OF ATLANTA   NOVEMBER/DECEMBER 1990
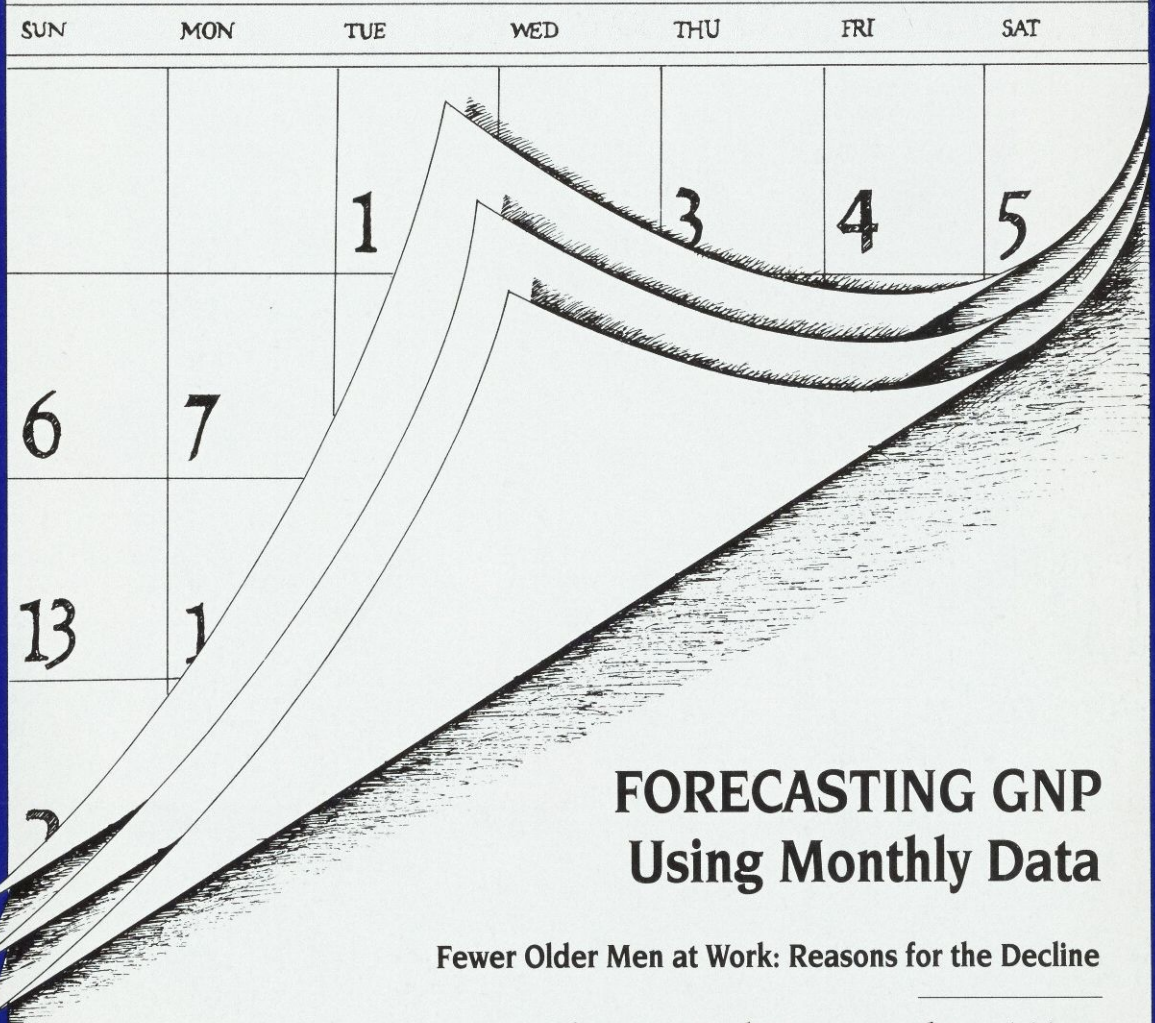
January
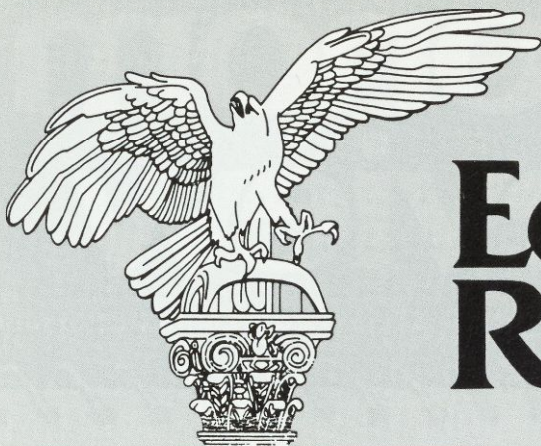
SUN	MON	TUE	WED	THU	FRI	SAT
		1		3	4	5
6	7					
13	1					

FORECASTING GNP
Using Monthly Data

Fewer Older Men at Work: Reasons for the Decline

Southeastern Bank Mergers and Acquisitions

WEBSTER '90

Economic Review

Forecasting U.S. GNP at Monthly Intervals with an Estimated Bivariate Time Series Model

Peter A. Zadrozny

January

SUN	MON	TUE	WED	THU	FRI	SAT
		1		3	4	5
6	7					
13	1					

WEBSTER '90

Conventional methods of constructing econometric models are typically fraught with problems of using available data efficiently. The author presents a new method that resolves difficulties with mixed frequency and temporal aggregation of data. (This method also deals with data measurement errors and transmission delays, but these problems are not discussed in this article.) Using monthly total nonfarm employment data and quarterly U.S. real GNP data from January 1958 to December 1978, the author applies a Kalman filter, often used in engineering and scientific applications, to construct a bivariate time series model for forecasting employment and GNP. The Kalman filter in a state-space setting is powerful in handling complex samples because it separates the model's characteristics from those of the sampling scheme.

Economic analysts continually attempt to comprehend large amounts of new and revised data to gauge the current state and future direction of the economy. The data come in many varieties. Reporting frequencies vary. Some data, like bond prices, pertain to an instant while other data, like quarterly gross national product (GNP), represent sums over periods of time. Some data are newly reported; some are revised, often more than once. Methods of collecting the data differ, and expected variation from the true value of the item differs widely among data series.

Analysts approach this variety of data in many ways: with quantitative tools such as

econometric models; with rules of thumb, intuition, or subjective judgments; or with some combination of these methods. Whatever their method, analysts face the same basic problem of using all information in the available data. Statisticians call this "efficient" use of data, by which they mean that the data's information is not wasted or distorted. The issue of efficiency arises in almost every empirical economic analysis because the form of relationships derived from economic theory does not match the way the variables are actually measured and reported.

For example, suppose that one wishes to forecast U.S. real GNP at monthly intervals. Economists' theory of production indicates that GNP is in part determined by total employment. Monetary theory indicates that GNP is influenced by some monetary aggregate like M1 or M2. But even the most carefully articulated production or monetary theory will not provide sufficient guidance about how to use the actual observations on GNP, employment, and money. A particular problem, with which this article is concerned, is that GNP, employment, and money are available at quarterly, monthly, and weekly intervals, respectively.

This type of problem—efficient use of mixed-frequency data—confronts virtually every economic analyst who uses more than one time series at a time. The difficulty is typically compounded by at least three other data problems: some variables are observed at a particular time while others are added up or averaged over time intervals; observations are subject to measurement errors; and observations are available only after a delay to allow for collecting, processing, and transmitting them.

Although these four data problems have been recognized for some time, they have not been satisfactorily addressed by conventional methods used to estimate forecasting models.[1] These methods allow very little flexibility in distinguishing between the forms of data as they are specified by economic theories and the ways in which the data are actually measured. Previously suggested solutions involved either wasting or distorting sample information.

As a response to these problems, this article briefly describes and applies a time series estimation and forecasting method that has the ability to resolve the problems simultaneously in a satisfactory way.[2] In view of the novelty of the method, the application in this article addresses only the first two data problems—mixed frequencies and data summed or averaged over time (temporal aggregation). The remaining problems of measurement errors and transmission delays will be addressed in a future study.

The method is applied to forecast U.S. real GNP at monthly intervals with the aid of monthly observations on U.S. total employment. It is used to estimate directly a monthly time-series forecasting model—a vector autoregressive moving-average (VARMA) model—with monthly employment and quarterly GNP data.

Problems Caused by Mixed Frequencies and Temporal Aggregation

Methods conventionally used to estimate time series models and forecast economic time series, such as ordinary least squares regression and its variants, are designed for data at a single frequency. Consequently, they generally do not handle mixed-frequency data satisfactorily. Attempting to use one of these regression methods with mixed-frequency data, economists have resorted to one of three options: transforming the data to a single frequency by adding up, averaging, or skip sampling the higher-frequency observations (like monthly employment) to the lower-frequency observations (like quarterly GNP); transforming the data to a single frequency by interpolating lower-frequency observations to the highest frequency in the data; or restricting the model so that the regression method can be directly applied.[3] None of these options is entirely satisfactory.

The first method—adding up, averaging, or skip sampling the higher-frequency observations—throws away sample information. Changing from monthly to quarterly observations in this way involves losing two-thirds of the observations in the monthly series.

On the other hand, when the goal is to estimate a model, for forecasting or other purposes, interpolation is at best a nuisance and at worst a source of distortion in the data. In the context of a sample of monthly and quarterly observations, for instance, interpolation means estimating the missing monthly values of the quarterly observations. Commonly used interpolation methods do not use all the available sample information. Statisticians commonly interpolate each time series separately, using only its relationship to its own previous values (see, for example, A.C. Harvey and R.G. Pierse 1984). Economists typically interpolate with regression methods, which primarily emphasize contemporaneous relationships between variables. However, even the most sophisticated regression-interpolation method will generally not be able to exploit all significant relationships within and among data series.[4] Moreover, to the extent that the interpolation method is inconsistent with the economic model being estimated, the estimates of the model and subsequent forecasts will be biased.

Imposing restrictions on a model by limiting certain relationships among its variables is another, more complex way to handle mixed-frequency data with single-frequency methods. In the example involving monthly employment and quarterly GNP, these restrictions eliminate all feedback from (quarterly) GNP to (monthly) employment and reduce feedback from lagged GNP to GNP to quarterly intervals. Most likely these restrictions on the model will inhibit its effectiveness. Because virtually any collection of aggregate economic time series will have significant feedbacks across all variables and all frequencies, models with arbitrarily restricted feedbacks of lower-frequency variables will generally produce suboptimal forecasts.

Mixed-frequency problems are often compounded because some variables are only observed as sums or averages; that is, they are temporal aggregates.[5] The problems caused by aggregation over time are twofold. First, to estimate a model with data that are added up or averaged, one must correspondingly add up or average the equations of the model. However, when this calculation can be done, it also requires imposing restrictions like those discussed in the previous paragraph. Moreover, if two or more variables have different degrees of temporal aggregation that are not multiples of each other, then the equations of the model cannot be aggregated into a form compatible with the data. In such cases, conventional methods break down and cannot be applied at all. Second, although estimating the model may pose no problems, with a conventional method one can only produce forecasts of variables in an aggregated or averaged form; there is no option to forecast some or all variables in original disaggregated form.

In sum, conventional regression methods, designed for single-frequency disaggregated data, are inadequate for estimating and forecasting with multivariate economic time series models when the data involved are mixed frequency or temporally aggregated. By contrast, the method outlined in the next section in principle overcomes these problems in a completely satisfactory way.[6] Unlike single-frequency methods, it does not require that the data be aggregated to the lowest frequency or interpolated to the highest frequency or that the model be arbitrarily restricted.

An Overview of Estimating and Forecasting with the Kalman Filter

The key to the present method is the use of a device called a Kalman filter. This device, often applied to engineering and other scientific problems, allows routine handling of the sampling complications of mixed frequencies and temporal aggregation. The basic insight that brings the Kalman filter into play is to view the model as operating at the highest frequency in terms of temporally disaggregated measurements of all the variables. In the employment-GNP context discussed above, the fact that GNP is observed only at quarterly intervals and only as an aggregate over each

quarter presents fewer observations of GNP than actually come from the process that generates the data. If the underlying process generates disaggregated monthly values of both employment and GNP, then quarterly GNP leaves one with "missing values" relative to the observations that could be obtained.

The Kalman filter inherits its ease of handling these and other sampling complications from the state-space representation of the model being considered. Indeed, the Kalman filter can handle any sampling complications—including measurement errors and data transmission delays—as long as they can be described as linear transformations of the model's variables. This flexibility derives from a clean separation, in a state-space formulation, of the characteristics of the model from the characteristics of the sampling scheme into two unconnected parts.

Therefore, the first step in applying the Kalman filter with a given model is to write the model in state-space form. A state space formulation of a time series model has two parts—a state vector, together with its law of motion, and an observation equation. The state vector is a list of variables that summarizes the relevant "state of the world" in the model. For a given model, there are many ways to formulate a state vector and its law of motion. In essence, however it is set up, a state vector consists of current and lagged values of the variables and disturbances of a model. The state vector and its law of motion summarize the dynamics of a model, as an entity separate from the way the variables in the model are observed. The observation (or measurement) equation tells how observations are made in terms of the state vector. In particular, it appropriately maps linear combinations of state variables and optional measurement errors into the vector of observations.

Kalman filtering state-space methods derive their power for handling complex samples from the separation of observation characteristics from the model equations. Setting up the observation equation to handle such complications is a simple matter. Then, given a state-space representation that incorporates the sampling scheme, the Kalman filter can be applied to formulate a likelihood function to estimate a model or

to produce forecasts with an estimated model.[7]

This powerful advantage of state-space Kalman filtering methods is contrasted with the severe limitations of conventional regression methods for handling sampling complications. The absence of a separate observation equation in conventional regression methods forces an attempt to transform the model equations to conform to the sampling scheme. Even when such a transformation is possible, it usually requires additional assumptions not motivated by economic theory.

Application to Forecasting GNP at Monthly Intervals

The following empirical application demonstrates the usefulness of Kalman filtering as a means for efficiently producing short-term economic forecasts with up-to-date values of different economic data that arrive at different frequencies. In the application, a VARMA model was estimated with monthly observations of U.S. total nonfarm employment and quarterly observations on U.S. real GNP for the period January 1958 to December 1978. The model was then used to forecast employment and GNP, from one to twelve months, for the period January 1979 to December 1988.

The employment data used are produced and published in *Employment and Earnings* by the U.S. Bureau of Labor Statistics. Based on monthly surveys of payrolls, they represent totals of all types of employees in nonfarm establishments. The GNP data are produced and published in *Survey of Current Business* by the U.S. Bureau of Economic Analysis. They are based on numerous surveys and are produced and revised in a lengthy process.[8] Both data series were obtained in seasonally adjusted form.

The data are historical series that mix revised early values with preliminary recent values. They have been treated as being of the same degree of revision (an approach that is justified because they are mostly final revisions).

Before the data were used they were transformed into annualized percentage growth

form. This transformation was done for two reasons. First, because GNP and the other aggregate economic variables are typically forecast in annualized percentage growth form, this data transformation makes the present forecasts comparable to forecasts of other economists. Second, putting data in this form makes them covariance stationary—in particular, removing their trends. Because trends dominated variations in the untransformed data, a failure to remove them would have resulted in inefficient use of the sample information.[9]

Chart 1 suggests that monthly employment figures contain useful information for forecasting GNP at monthly intervals.[10] Clearly, variations of employment and GNP follow each other quite closely; somewhat less clearly, within-year variations in employment appear to slightly lead variations in GNP. Of course, other monthly series may contain additional useful information for forecasting GNP, but, given the limited objectives of the present study, other series were not considered.[11]

That the two series contain useful information for forecasting is more precisely revealed by serial and cross-serial correlation coefficients that are summary measures indicating the presence of lagged linear relationships (feedbacks). A serial correlation at lag k is a measure of the degree of linear relationship between a variable and values of itself k periods in the past. A cross-serial correlation at lag k is a measure of the degree of linear relationship between a variable and another variable k periods in the past. Like all correlations, serial and cross-serial correlations range from -1 to $+1$: -1 indicates a perfect negative linear relationship; 0 indicates no linear relationship; and $+1$ indicates a perfect positive linear relationship. Once significant feedbacks are indicated by inspecting serial and cross-serial correlations, the precise nature of the feedbacks is determined by estimating a model.[12]

Cross-serial correlations of employment and GNP are displayed in Charts 2 and 3. In these charts individual correlations are significantly different from zero at about the 95 percent confidence level if they are greater than or equal to .20 in absolute value. At this level of confidence, Charts 2 and 3 show significant feedbacks between employment and GNP at lags of up to about six months. Charts 4 and 5

Chart 1.
Percentage Growth of Employment and GNP
(January 1958 to December 1988)
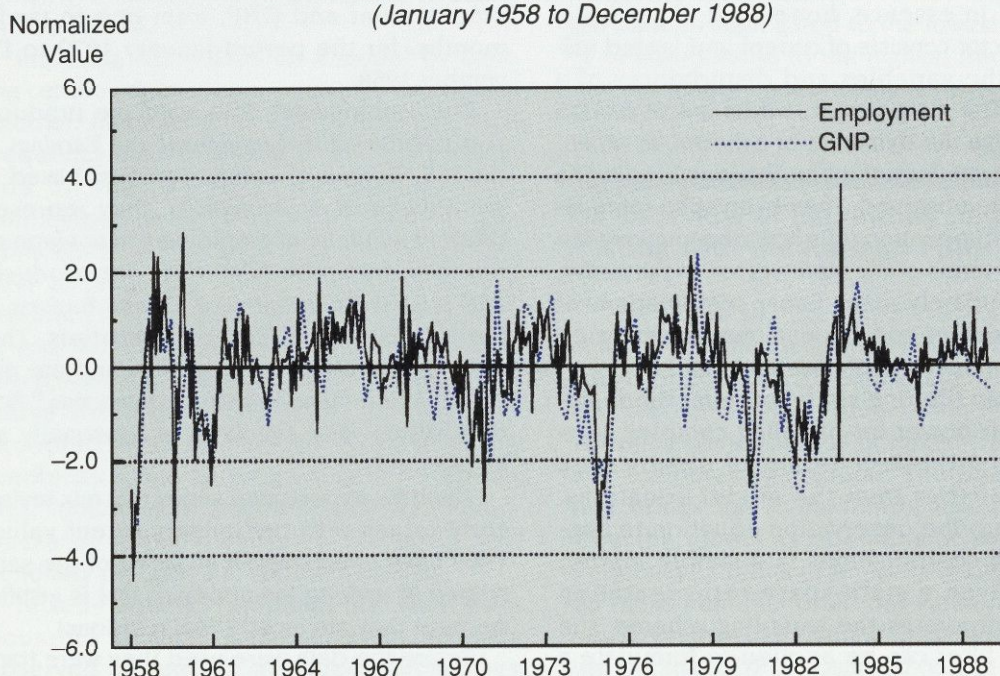
Chart 2.
Monthly Cross-Serial Correlations of Percentage Growth of Current GNP and Lagged Employment

C(k)

[Line chart showing monthly cross-serial correlations. Y-axis ranges from -1.0 to 1.0. X-axis shows Monthly Lags from 0 to 36. Values start near 0.45, peak at ~0.6 at lag 1, decline to around 0.0 by lag 6, then fluctuate between 0.0 and -0.3 for the remaining lags.]

Monthly Lags
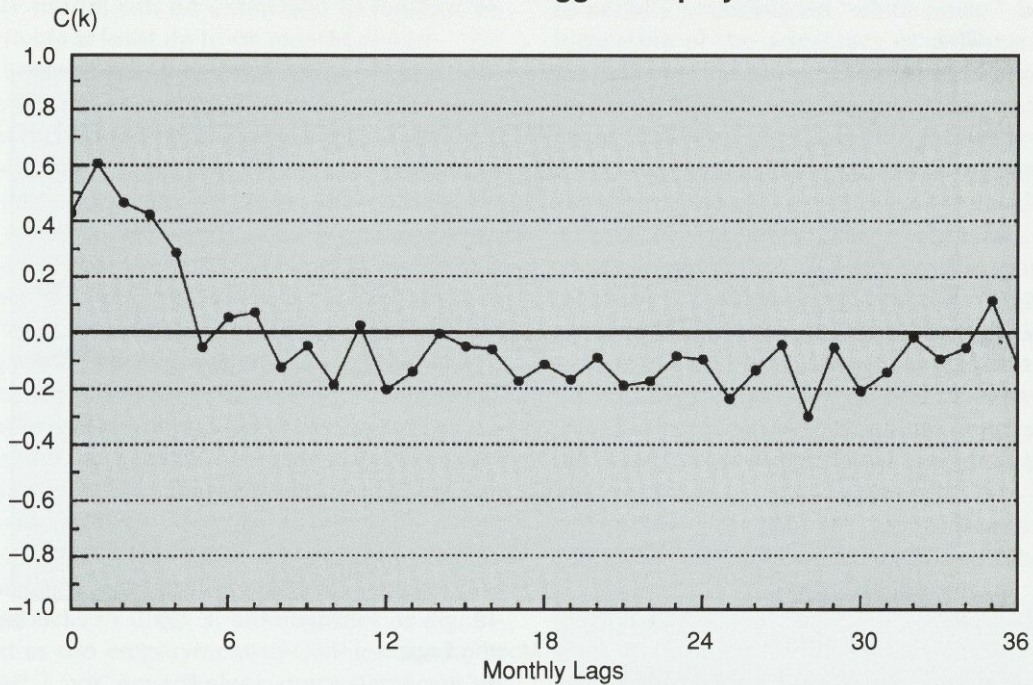
Chart 3.
Monthly Cross-Serial Correlations of Percentage Growth of Current Employment and Lagged GNP

C(k)

[Line chart showing monthly cross-serial correlations. Y-axis ranges from -1.0 to 1.0. X-axis shows Monthly Lags from 0 to 36. Values start near 0.4, gradually decline through the lags, crossing zero around lag 16-17, and fluctuating around -0.1 to -0.2 for later lags.]

Monthly Lags
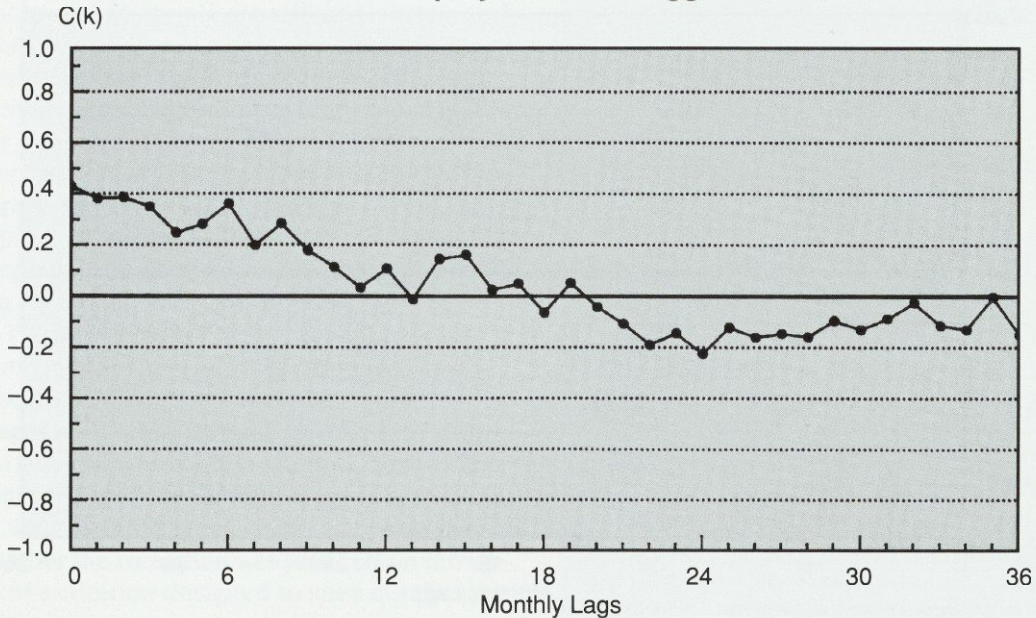
Chart 4.
Monthly Serial Correlations of Percentage Growth of Employment
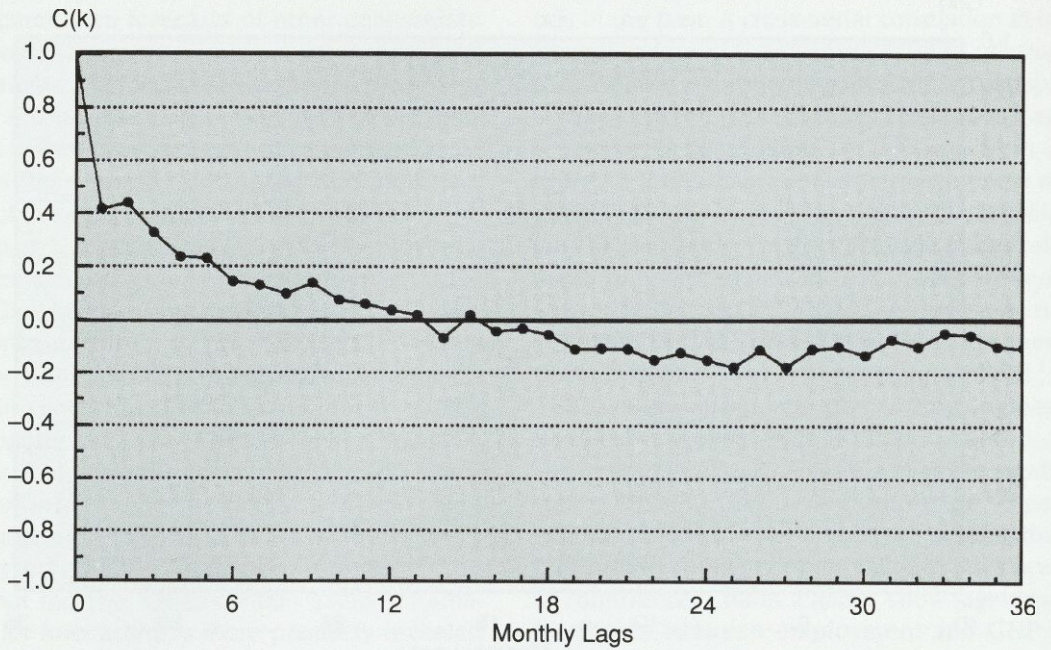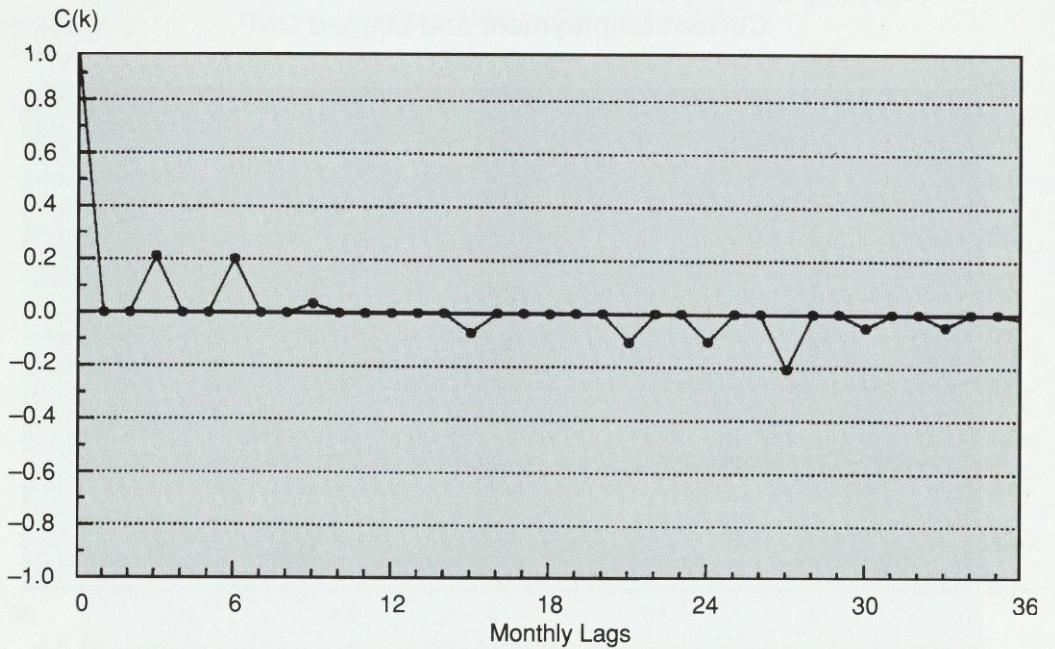
C(k)

Monthly Lags

Chart 5.
Monthly Serial Correlations of Percentage Growth of GNP

C(k)

Monthly Lags

demonstrate that employment is significantly serially correlated and GNP is only marginally serially correlated in this range. These four charts suggest that a monthly employment-GNP model can be estimated to forecast effectively at least up to six months ahead.

Because monthly employment fluctuations are the sample information which permits estimation and forecasting at monthly intervals, it would seem that the most critical correlations in this respect are the cross-serial correlations of current GNP and lagged employment in Chart 2, which reflect feedbacks from employment to GNP. Indeed, these are the most significant correlations in Charts 2 to 5. In the regression approach to this application, in the absence of interpolated monthly GNP values, employment would have to be treated as exogenous. As a result, GNP-to-employment and GNP-to-GNP feedbacks would not be used in producing GNP forecasts. Although the GNP-to-GNP feedbacks reflected in Chart 5 are only marginally significant, the GNP-to-employment feedbacks in Chart 3, although not as significant as the employment-to-GNP feedbacks in Chart 2, are, nevertheless, quite significant up to about the eight-month lag.

When all variables in a multivariate system feed back on each other and on themselves, as employment and GNP do, efficient forecasting requires that all own and cross feedbacks—that is, all significant serial and cross-serial correlations—be taken into account. Efficiently forecasting GNP several months ahead with the present data requires not only exploiting significant feedbacks from employment to employment and to GNP, as regression methods do, but also exploiting significant feedbacks from GNP to GNP and to employment, as the present method allows but regression methods cannot.

The search for the best model began with a fairly complex model (a bivariate ARMA [3,1] model) and whittled it down in a series of stages to a simpler form (a bivariate ARMA [1,1] model). At each stage, the weakest parameters were eliminated, the reduced model was reestimated, and it was decided whether the reduction was justified on the basis of a criterion designed to seek out the true data-generating process.[13] The resulting estimated model is presented in Table 1.[14]

Considering that the data are in growth form, the model appears to fit the data fairly well.[15] But, because the major goal of time series estimation is to reduce the residuals of the model to serially uncorrelated "white noise," better indicators of the adequacy of an estimated model are serial and cross-serial correlations of residuals. When these correlations are insignificantly different from zero, the model is adequate because it has captured all systematic variation in the data over time: all residual variation is unpredictable white noise. In fact, time plots of correlations and cross-correlations of residuals show only a few isolated marginally significant correlations but do not show any systematic patterns that would indicate that the model should be modified or extended (P.A. Zadrozny 1990). This finding confirms that the estimated model in Table 1 is adequate in this respect.[16]

The next step is to evaluate forecasts produced with the best overall model (see Table 2),

---

Table 1.
Best Overall Estimated Bivariate Monthly Employment–GNP Model: Unrestricted ARMA (1,1)

Maximum Likelihood Estimates

$$e(t) = .799e(t-1) + .417g(t-1) + 2.37\varepsilon_1(t)$$
$$\quad\quad (1.75) \quad\quad\quad (.377) \quad\quad\quad (2.54)$$

$$-\ .615\varepsilon_1(t-1) - .697\varepsilon_2(t-1)$$
$$\quad (-.204) \quad\quad\quad (-.112)$$

$$g(t) = .203e(t-1) + .353g(t-1) + .634\varepsilon_1(t)$$
$$\quad\quad (1.19) \quad\quad\quad (1,228.) \quad\quad (.140)$$

$$+\ 1.34\varepsilon_2(t) + 1.72\varepsilon_1(t-1) - .613\varepsilon_2(t-1)$$
$$\quad (.222) \quad\quad (.441) \quad\quad\quad (-.078)$$

Model-Fit Summary Statistics

Variable	Standard Error	R^2
Employment	2.83	.228
GNP	2.82	.472

The estimation period is January 1958 to December 1978; e(t) = percentage growth of employment; g(t) = percentage growth of GNP; $\varepsilon_1(t)$, $\varepsilon_2(t)$ = disturbances; t ratios are in parentheses.

over the post-estimation period, January 1979 to December 1988. The forecasts were generated as described in Zadrozny (1990) and were evaluated in several ways in terms of their root mean squared errors (RMSE).[17]

First, Theil U statistics were examined. For a given variable and a given number of forecasting steps ahead, a Theil U statistic is the RMSE of the forecasting method being evaluated divided by the RMSE of naive forecasts, where the naive forecast of a variable any number of steps ahead is the most recent observation of the variable at the time the forecast is made. Theil U statistics provide an internal test of the competitiveness of the model's forecasts. If the model's forecasts, which cost something to be produced, have greater errors than the naive forecasts—that is, they yield Us greater than one—then they are inferior to the costless naive method. Table 2 shows U statistics all significantly less than one. The forecasts of the model in Table 1 are thus quite competitive in this respect, especially in short-horizon forecasting.[18]

The claim made above that all feedbacks among the variables reinforce each other and thus should be used in forecasting is tested next by comparing the Theil U statistics of the model with those of best univariate models for employment and GNP. The criterion for choosing the "best" univariate model was the same as was used in choosing the best bivariate model in Table 1. The message from comparing RMSEs and Theil U statistics of the bivariate model with those of the univariate model is that the accuracy in forecasting employment improves only very slightly when GNP is used in forecasting employment but improves dramatically, especially in short-horizon forecasts, when employment is used in forecasting GNP.

Having demonstrated that all available information should be used in forecasting, the next test is to compare the forecasts of the best overall model with those of the best model obtained with a conventional regression method, shown in Table 3.[19] All serial and cross-serial correlations of the model's residuals were insignificantly different from zero. Table 4 reports monthly forecast-evaluation statistics of the regression mod-

Table 2.
Forecast-Evaluation Statistics of Model in Table 1

Employment Forecast-Evaluation Statistics

Months Ahead	RMSE	Theil U
1	2.33	.576
2	2.40	.592
3	2.56	.632
4	2.69	.664
5	2.77	.684
6	2.83	.699
7	2.87	.710
8	2.93	.724
9	3.02	.747
10	3.10	.767
11	3.18	.786
12	3.22	.796

GNP Forecast-Evaluation Statistics

Months Ahead	RMSE	Theil U
1	2.51	.437
2	2.79	.485
3	3.34	.582
4	3.72	.647
5	3.85	.670
6	3.97	.692
7	4.20	.731
8	4.14	.721
9	4.15	.723
10	4.24	.738
11	4.25	.740
12	4.30	.749

The forecast-evaluation period is January 1979 to December 1988. RMSE means root mean-squared forecast error. Theil U Statistic = RMSE of forecast ÷ RMSE of the naive forecast, where the naive forecast is the most recent observation at the time the forecast is made.

el, which can be directly compared with those in Table 2. The comparison shows that the forecasting performance of the present method is uniformly better—marginally for employment and significantly for GNP—than the performance of the regression method, especially for forecasts up to five months ahead.

Table 3.
Best Standard-Method Bivariate
Monthly Employment–GNP Model:
Restricted AR(1)

Maximum Likelihood Estimates

$$e(t) = .938e(t-1) + 2.31\varepsilon_1(t)$$
$$\quad\quad (8.40) \quad\quad (3.49)$$

$$g(t) = .387e(t-1) + 1.11\varepsilon_1(t) + 1.78\varepsilon_2(t)$$
$$\quad\quad (1.60) \quad\quad (.452) \quad\quad (1.72)$$

Model-Fit Summary Statistics

Variable	Standard Error	R^2
Employment	2.84	.227
GNP	3.24	.304

The estimation period is January 1958 to December 1978; $e(t)$ = percentage growth of employment; $g(t)$ = percentage growth of GNP; $\varepsilon_1(t)$, $\varepsilon_2(t)$ = disturbances; t ratios are in parentheses.

Table 4.
Forecast-Evaluation Statistics
of Model in Table 3

Employment Forecast-Evaluation Statistics

Months Ahead	RMSE	Theil U
1	2.36	.583
2	2.42	.598
3	2.58	.637
4	2.70	.667
5	2.79	.689
6	2.85	.704
7	2.91	.718
8	2.96	.732
9	3.05	.753
10	3.11	.769
11	3.18	.786
12	3.22	.796

GNP Forecast-Evaluation Statistics

Months Ahead	RMSE	Theil U
1	2.96	.516
2	3.17	.552
3	3.51	.612
4	3.79	.660
5	3.91	.680
6	4.01	.698
7	4.24	.738
8	4.18	.727
9	4.18	.728
10	4.27	.744
11	4.27	.744
12	4.32	.752

The forecast-evaluation period is January 1979 to December 1988. RMSE means root mean-squared forecast error. Theil U Statistic = RMSE of forecast ÷ RMSE of the naive forecast, where the naive forecast is the most recent observation at the time the forecast is made.

All the tests so far were "internal" in the sense that all the forecast-evaluation statistics were based on the same employment–GNP data set. To put the results in Tables 1 to 4 in a wider perspective, some comparisons of the forecasting performance of the best overall model with the forecasting performances of models developed by others are now conducted. The internal comparisons are rigorous comparisons of forecasting performance because they are all based on the same variables and model-estimation and forecast-evaluation periods. By contrast, the set of variables and estimation and forecasting periods were different in each study in the external comparisons. Moreover, only one of the outside studies used monthly data and made monthly GNP forecasts. Therefore, any conclusions drawn from the external comparisons are weakened by these data disparities. Nevertheless, as Table 5 shows, even after taking these differences into account the forecasts of the best model produced with the present method appear to be competitive—indeed, generally better for forecasts up to three months. In particular, the present forecasts are impressive in these comparisons because they were produced with a model involving substantially fewer variables.[20]

Table 5.
Comparison with Other Forecasts
for Comparable Periods

Quarterly GNP Forecast RMSEs

Quarters Ahead	Table 2	Roberds	McNees
1	3.34	3.87	4.25
2	3.97	3.85	3.61
3	4.15	4.15	3.78
4	4.30	4.34	3.82

Monthly GNP Forecast RMSEs

Months Ahead	Table 2	Trehan
1	2.51	1.81
2	2.79	3.34

Sources: "Roberds" data are from Table 2, part 1, column 2 in Roberds (1988) and reflect forecasts over quarter 1, 1977, to quarter 1, 1987. "McNees" data are the average of columns 1 to 4 in Table 2 in McNees (1986) and reflect forecasts over quarter 2, 1980, to quarter 1, 1985. "Trehan" data are from Table 2, part 2, column 3 in Trehan (1989) and reflect forecasts over quarter 4, 1978, to quarter 4, 1988.

Conclusion

This article has demonstrated the feasibility and usefulness of applying a Kalman filtering method to directly estimate a monthly model of employment and GNP, when employment is observed monthly and GNP is observed quarterly, and of using the estimated model to produce monthly forecasts of GNP. The forecasts were evaluated by internal tests and by comparisons with forecasting performance of models developed by others. The tests and comparisons showed that the method is able to produce competitive GNP forecasts using only employment as the source of monthly information. The method is computationally much more demanding than standard regression methods; however, with some experience, an application of the scope of the one presented in this article could be conducted in an afternoon on a personal computer.[21]

As noted in the introduction, the Kalman filtering method can routinely handle the sampling complications of mixed frequencies, temporal aggregation, measurement errors, and data transmission delays. In the application mixed frequencies and temporal aggregations were fully accounted for, measurement errors were somewhat accounted for, and transmission delays were ignored.[22] It would be useful to extend the present application to fully account for measurement errors and data transmission delays and to use, in addition, other sources of monthly information such as industrial production and retail sales.

Measurement errors and transmission delays are intimately connected. Consider the case of GNP data. The first observations on GNP for a given quarter are available fifteen days after the end of the quarter. As more survey data come in and are processed, more accurate estimates of GNP are released. In fact, there is a succession of releases of GNP, each presumably more accurate than the previous one. The next two releases, after the so-called fifteen-day release, are forty-five and seventy-five days after the end of the quarter.[23] A plan for future application of the method discussed here is to extend the present application to take such revisions—that is, measurement errors and transmission delays—fully into account, not just in GNP but also in any other monthly variables used.[24]

Finally, although the application in this article was a purely statistical exercise in the sense that no economic theory was directly used to specify and restrict any equations in the model, the method could equally well be used to estimate, forecast, and make inferences with models directly motivated by an economic theory. That is, the method could be applied to a wide range of economic models such as dynamic simultaneous equations models and rational expectations models.[25]

Notes

[1]Some examples of forecasting in which conventional single-frequency methods are applied to mixed-frequency data are Corrado and Greene (1988), Fitzgerald and Miller (1989), and Trehan (1989).

[2]The present article is drawn from a working paper (Zadrozny 1990) that describes the method and application in much greater detail and sophistication. This working paper is available from the Public Information Department of the Federal Reserve Bank of Atlanta.

[3]Skip sampling means reducing a time series by picking out nonconsecutive values, that is, at greater time intervals than those in the original series. For example, to reduce a monthly employment series to quarterly values by temporal aggregation one would average the three monthly values of each quarter. However, to reduce the employment series by skip sampling one would pick the first (or second or third) monthly value of each quarter.

[4]Litterman's (1983) method, which exploits two levels of serial correlations in residuals, is the most sophisticated regression-interpolation method. Nevertheless, it cannot exploit all possible (potentially significant) cross-serial correlations between low-frequency variables being interpolated and high-frequency variables serving as interpolators.

[5]Yet another problem may be cross-sectional aggregation; that is, some observations are sums of several different variables. In fact, some observations could be simultaneously cross-sectionally and temporally aggregated. The method presented here can simultaneously handle cross-sectional and temporal aggregation.

[6]Obviously, the present method also has its limitations. First, compared with regression methods, it is computationally much more demanding so that the sizes of models that it can handle are much more limited. Second, regardless of how it is estimated, a model must satisfy a set of identification conditions. Identification conditions for estimating VARMA models with single-frequency data that are not temporally aggregated are well known; their extensions to mixed-frequency temporally aggregated data are as yet unknown. Fortunately, in a given application there are informal empirical ways of checking whether identification conditions are satisfied.

[7]For a detailed discussion of setting up a state-space representation of a VARMA model for a sample involving mixed frequencies and temporal aggregation and then using the Kalman filter with the obtained state-space representation to estimate the model, see Zadrozny (1990).

[8]Carson (1987) and Young (1987) discuss in detail the compilation of GNP data.

[9]Data are said to be (covariance) stationary when their mean values (central tendency) and variances (dispersion) are constant over the sample. A failure to put data into stationary form prior to using them usually results in inefficient estimates of the model and biased test statistics. Economic data are typically nonstationary because of the presence of strong trends. In the application, trends were removed by taking monthly and quarterly differences in the transformation to percentage growth form. There has been some discussion lately about whether trends should be removed by differencing or by linear regression (see, for example, Whiteman and Roberds 1990). A generalization of differencing, which recognizes that variables may share common trends, is to treat them as cointegrated (see, for example, Engle and Granger 1987 and Fountis and Dickey 1989). The model can also legitimately be estimated with undetrended data if appropriate restrictions are placed on the parameters. This technique has been used especially in Bayesian analyses (see, for example, Doan, Litterman, and Sims 1984, Litterman 1986, and Roberds 1988). Data from 1947 to 1957 were also considered, but time plots analogous to Chart 1, extending from 1947 to 1988, revealed that the transformed data display notably greater variations from 1947 to 1957 than in later years. Therefore, to be sure of a stationary sample, only data from 1958 onward were used.

[10]Hereafter, "employment" and "GNP" mean employment and GNP in transformed annualized percentage growth form.

[11]For example, Fitzgerald and Miller (1989) used total hours worked (average hours worked times total employment) in lieu of total employment; Trehan (1989) used industrial production and retail sales in addition to employment.

[12]The reported correlations were computed with the usual formulas for sample variances and covariances, modified only to skip terms in summations which could not be computed because of missing values of GNP. Each of the summations were normalized by the actual number of terms that were computed. In Chart 5 correlations at nonquarterly intervals (1, 2, 4, 5, . . .) were set to zero to indicate that they could not be computed with purely quarterly data.

[13]Parameters were considered "weakest" when their estimates and t ratios (estimates divided by estimated standard errors) were both less than .1 in absolute value. The criterion for selecting the "best" model was the Akaike (1973) information criterion, modified to correct for finite-sample bias as suggested by Hurvich and Tsai (1989).

[14]In Tables 1 and 3 the moving-average parts in the estimated equations have been normalized so that the disturbances, ε_1 and ε_2, have unit variances and are uncorrelated with each other and that the current value of the second disturbance, $\varepsilon_2(t)$, is absent from the employment equation. The normalization is necessary to solve an identification problem (eliminate redundant parameters) and, therefore, does not restrict the model. Without further identifying assumptions, $\varepsilon_1(t)$ and $\varepsilon_2(t)$ cannot be viewed as respective "pure" employment and GNP shocks; they can only be viewed as unspecified linear combinations of such pure shocks.

[15]An R^2 goodness-of-fit statistic of .228 for employment is fair and an R^2 of .472 for GNP is good. Although R^2 is not a reliable statistic here, it nevertheless is useful because it summarizes the fit of the individual equations

in the model. As in ordinary least squares (OLS) estimation, R^2 is computed as 1 minus the variance of residuals divided by the variance of observations. But, unlike in OLS, the present estimation criterion is not entirely geared toward maximizing the fit of individual equations. Therefore, it can happen that the R^2s of some equations can be negative. In OLS, R^2s necessarily satisfy $0 \leq R^2 \leq 1$.

[16]One source of skepticism about the adequacy of the model might be that only three of the parameters are in the range of statistical significance by conventional standards. An estimated parameter is significantly different from zero at about the 90 percent confidence level when its t ratio is greater than about 1.6 in absolute value. (Note that the improbably high t value of 1,228 is undoubtedly the result of numerical errors inherent to the numerical method for calculating standard errors which was used.) The generally low t values of estimated parameters are admittedly something to be concerned about and will be given attention in future extensions of this work. The only thing to do in this respect is to try to come up with restrictions on parameters that result in higher t ratios without reducing the fit—and forecasting ability—of the estimated model. In fact, the role of Bayesian VAR or BVAR estimation is to do just that. See, for example, Doan, Litterman, and Sims (1984), Litterman (1986), and Roberds (1988).

[17]Let RMSE(k) denote the RMSE of k-months ahead forecasts of a variable. Then, for example, in the case of employment, RMSE(k) is the square root of the average value of $[e(t + k) - e(t + k|t)]^2$ over the range $t = S + 1$ to $t + T - k$, where $e(t + k)$ denotes the observation on employment in month $t + k$, $e(t + k|t)$ denotes the forecast of employment in month $t + k$ made in month t, $S + 1$ denotes January 1979, and T denotes December 1988. Of course, in the case of GNP, with $g(t + k)$ and $g(t + k + t)$ defined analogously to $e(t + k)$ and $e(t + k|t)$, $[g(t + k) - g(t + k|t)]^2$ could only be computed for values of $t + k$, which represent third months of quarters.

[18]Although the principal interest is in evaluating GNP forecasts, employment forecasts are also reported be-

cause the method treats the two variables symmetrically.

[19]To make the comparison, a search was made for a best model subject to the type of restriction that would be necessary in estimation with a conventional regression method. Although the maximum likelihood estimation program of the present method was actually used, the imposition of restrictions delivers results equivalent to what regressions can produce.

[20]For example, Roberds (1988) uses a total of eight variables: real GNP, implicit GNP deflator, unemployment rate, business fixed investment, monetary base, yield on three-month Treasury bills, Atlanta-Fed dollar index, and commodity price index. Trehan (1989) uses a total of four variables: real GNP, total employment, industrial production, and retail sales.

[21]The application was carried out with a FORTRAN program written by the author. The program was compiled with the Lahey Computer Systems FORTRAN 77 compiler (version 3.00) and was run on a 386 personal computer operating at 20 MHZ clock speed with a numerical coprocessor. It took less than twenty minutes to do all of the computations underlying the figures and tables reported in Zadrozny (1990).

[22]The extent to which measurement errors were accounted for is discussed in Zadrozny (1990, 22-24).

[23]Revisions continue, with the next one being annual. Every July the previous three years' releases are revised to be consistent with annual survey data.

[24]Using state-space methods, Conrad and Corrado (1979), Howrey (1984), and Scadding (1987) accounted for the different precisions of preliminary and revised data but did not consider transmission delays. On the other hand, Bordignon and Trivellato (1989), also using state-space methods, studied the effects of timeliness of data but did not simultaneously work with preliminary and revised data.

[25]See, for example, Judge et al. (1980) and Hansen and Sargent (1989) for discussions of dynamic simultaneous equations models and rational expectations models.

References

Akaike, H. "Information Theory and an Extension of the Maximum Likelihood Principle." In *Second International Symposium on Information Theory*, edited by B.N. Petrov and F. Csaki, 267-81. Budapest: Akademia Kiado, 1973.

Bordignon, S., and U. Trivellato. "The Optimal Use of Provisional Data in Forecasting with Dynamic Models." *Journal of Business and Economic Statistics* 7 (April 1989): 275-86.

Carson, C.S. "GNP: An Overview of Source Data and Estimating Methods." *Survey of Current Business* 67 (July 1987): 103-26.

Conrad, W., and C. Corrado. "Application of the Kalman Filter to Revisions in Monthly Retail Sales Estimates." *Journal of Economic Dynamics and Control* 1 (1979): 177-98.

Corrado, C., and M. Greene. "Reducing Uncertainty in Short-Term Projections: Linkage of Monthly and Quar-

terly Models." *Journal of Forecasting* 7 (April-June 1988): 77-102.

Doan, T.A., R.B. Litterman, and C.A. Sims. "Forecasting and Conditional Projections Using Realistic Prior Distributions." *Econometric Reviews* 3, no. 1 (1984): 1-100.

Engle, R.F., and C.W.J. Granger. "Cointegration and Error Correction: Representations, Estimation, and Testing." *Econometrica* 55 (March 1987): 251-76.

Fitzgerald, T.J., and P.J. Miller. "A Simple Way to Estimate Current-Quarter GNP." Federal Reserve Bank of Minneapolis *Quarterly Review* 13 (Fall 1989): 27-31.

Fountis, N.G., and D.A. Dickey. "Testing for a Unit Root Nonstationarity in Multivariate Autoregressive Time Series." *The Annals of Statistics* 17 (1989): 419-28.

Hansen, L.P., and T.J. Sargent. "Formulating and Estimating Dynamic Linear Rational Expectations Models." *Journal of Economic Dynamics and Control* 2 (1980): 7-46.

Harvey, A.C., and R.G. Pierse. "Estimating Missing Observations in Economic Time Series." *Journal of the American Statistical Association* 79 (March 1984): 125-31.

Howrey, E.P. "Data Revision, Reconstruction, and Prediction: An Application to Inventory Investment." *Review of Economics and Statistics* 66 (August 1984): 386-93.

Hurvich, C.M., and C.L. Tsai. "Regression and Time Series Model Selection in Small Samples." *Biometrika* 76, no. 2 (1989): 297-307.

Judge, G.G., W.E. Griffith, K.C. Hall, and T.C. Lee. *The Theory and Practice of Econometrics*. New York: John Wiley, 1980.

Litterman, R.B. "A Random Walk, Markov Model for the Distribution of Time Series." *Journal of Business and Economic Statistics* 1 (April 1983): 169-73.

_____. "Forecasting with Bayesian Vector Autoregressions—Five Years of Experience." *Journal of Business and Economic Statistics* 4 (January 1986): 25-38.

McNees, S.K. "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts." *Journal of Business and Economic Statistics* 4 (January 1986): 5-15.

Roberds, W.T. "A Quarterly Bayesian VAR Model of the U.S. Economy." Federal Reserve Bank of Atlanta Working Paper 88-2, May 1988.

Scadding, J. "The Nature of GNP Revisions." Federal Reserve Bank of Cleveland Working Paper, 1987.

_____. "Getting the Noise Out: Filtering Early GNP Estimates." Federal Reserve Bank of Cleveland *Economic Review* 24 (Summer 1988): 24-31.

Trehan, B. "Forecasting Growth in Current Quarter Real GNP." Federal Reserve Bank of San Francisco *Economic Review* (Winter 1989): 39-51.

Whiteman, C.H., and W.T. Roberds. "Monetary Aggregates as Monetary Targets: A Statistical Investigation." Federal Reserve Bank of Atlanta Working Paper 90-7, October 1990.

Young, A.H. "Evaluation of the GNP Estimates." *Survey of Current Business* 67 (August 1987): 18-42.

Zadrozny, P.A. "Estimating a Multivariate ARMA Model with Mixed-Frequency Data: An Application to Forecasting U.S. GNP at Monthly Intervals." Federal Reserve Bank of Atlanta Working Paper 90-6, August 1990.

Fewer Older Men in the U.S. Work Force: Technological, Behavioral, and Legislative Contributions to the Decline

Jon R. Moen

Older people are becoming a larger share of America's population, yet their participation rate in the labor force has shrunk since the late nineteenth century. This study examines two schools of thought that attempt to explain this decline. The author finds little evidence to support the view that changing attitudes toward older workers account for their dwindling numbers in the work force. He finds more valid the theory that credits the decline to the nation's shift from an agrarian to a manufacturing economy. In addition, since the 1930s government policies such as New Deal employment legislation and Social Security have created incentives that work to institutionalize retirement as a career stage.

S ince the late 1800s the labor force participation rate among American men aged sixty-five and older—the number actively employed or seeking employment as a percentage of the total population of American men in that age group—has continually declined. As the nation's population includes more older people, it is increasingly important to consider how social and economic policies influence the number of older Americans in the work force, the number of hours they work, and their skills; future labor force size and quality will be influenced by the effects of such poli-

cies. These policies in turn depend to some extent on decisionmakers' understanding of the factors that have influenced the secular (that is, long-term) decline in labor force participation by men aged sixty-five and older.

Two broad explanations, with differing policy implications, have attempted to account for the decline in labor force participation by older men up to the 1930s. The first is a "new history" or behavioral approach, which asserts that the labor force participation rate has declined among older persons because changing attitudes have resulted in both age-based job discrimination and loss of status, independent of any changes in the economy or in the economy's existing production technology. Second, a more traditional view emphasizes factors in the economy such as technological change or occupational shifts.[1] Explanations of the steep decline after the 1930s have centered on economic factors, particularly gov-

ernmental actions, like the institution of Social Security, and the accelerating development of private pension plans.

The broad-based nature of the decline is demonstrated clearly by the data presented in Table 1. These data show that between 1860 and 1900 the overall participation rate decline occurred among rural, nonfarm residents. Since 1900 the decline in labor force participation has been steep and systematic and has occurred across all household types. The fact that the decline in participation rates of older men has been both widespread and pronounced during the twentieth century has been viewed by disciples of the behavioral approach as prima facie evidence that older workers are being systematically excluded from the work force. This view draws much of its intellectual support from a historical perspective that has developed over the past few decades of the

place of older workers and the aged in American society. This article is a review of some of the arguments economic historians have advanced to support each of these views. It is also an examination of the role twentieth-century legislation and governmental regulation may have played in the decline.

Modernization Theory—The Technological (Structural) Shift View

Most historical interpretations of the secular decline in labor force participation rates among workers aged sixty-five and older before 1930 can be loosely categorized under the heading of modernization theory (see, for example, Abraham Epstein 1928). Stated simply, modernization theory contends that the secular decline can be readily explained by

Table 1.
Labor Force Participation Rates of American
Men Aged Sixty-Five and Older, 1860 to 1980
(*percent*)

Household Type	1860	1900	1950	1980
All	75.59	69.98	46.99	25.60
Farm	79.78	80.34	61.81	57.64
Nonfarm	70.60	62.84	43.97	24.49
Urban	64.19	65.30	—	25.02
Rural	77.62	73.29	—	27.06
Nonfarm, rural	73.74	61.20	—	22.96

Sources: Data for 1860 are from the Batemen-Foust Sample of Northern Households, 1860 (Ann Arbor, Mich.: Interuniversity Consortium for Political and Social Research [ICPSR]) and from the Sample of Older Males in Four Large Cities, 1860, available from the author. Data for 1900 are from the 1900 Public Use Sample, Stephen Graham, *The 1900 Public Use Sample Users Handbook*, no. 7825 (Ann Arbor, Mich.: ICPSR, 1981). Data for 1950 are from the 1950 Public Use Sample, *Census of Population, 1950: Public Use Microdata Sample, Technical Documentation (Preliminary)*, U.S. Bureau of the Census (Washington, D.C.: U.S. Government Printing Office, 1984). Data for 1980 are from the 1980 Public Use Microdata Sample, U.S. Bureau of the Census.

the shift in the United States from an agrarian economy to one based on manufacturing and mass production. The emphasis on efficiency in this new economic order made older workers obsolete.

Agrarian societies characteristically allow older workers to set their own pace and to turn over gradually the more strenuous farming tasks to younger household members. In addition, owning land brings older persons respect and social status and a degree of control over their children through the prospect of withdrawing inheritances. According to the modernization view, the shift to manufacturing and the loss of control over property have diminished older persons' status in the eyes of the young. Stated differently, the shift resulted in the falling marginal productivity of older members of the labor force and the substitution of younger for older workers. Real changes in the economy are seen as the ultimate causes of the decline in importance of older persons in the work force.

Brian Gratton's (1986) study of elderly workers in Boston provides, on the surface, some support for the modernization theory. Gratton suggests that the increasing importance of the lower-participation manufacturing sector relative to the high-participation farming sector was responsible for the decline in the overall labor force participation rate of older men at the beginning of the twentieth century, even though participation rates changed little within either sector.

The Shift Away from Agriculture. The declining number of farms and the increase in the urban population in the United States have long been regarded as important contributors to the fall in the labor force participation rate among men sixty-five and older. As noted above, it is believed that farming allowed older workers to set their own pace or let younger workers undertake the more strenuous tasks. On the other hand, factory jobs were thought to have permitted older workers fewer employment alternatives as

their ability to work at a continuous pace diminished. Unlike farming, manufacturing separated the home and workplace. Rather than work part-time or at less demanding jobs, older men working in manufacturing were more likely to retire completely.

Although the link between the declining importance of agriculture and the decline in labor force participation rates for older workers is conceptually simple, the mechanism through which the shift affected the overall participation rate is undoubtedly more complex. Gratton intuitively described how the shift may have affected overall participation: "As the economy industrialized, the relatively low participation of older men in the industrial work force began to dominate their general participation rate" (1986, 21).

Gratton is correct in pointing out that structural changes in the economy could have caused participation rates of older men to fall in the aggregate, independent of workers' behavior within either the manufacturing or farming sector. However, closer scrutiny of the data reveals that labor force participation rates did decline in some household types.

In 1860 roughly 65 percent of the men aged sixty-five and older lived in farm households. By 1980 this share had fallen below 3 percent. At the same time, the percentage of older men living in rural nonfarm households remained about 25 percent, while the percentage of men living in urban areas increased from about 11 percent to approximately 71 percent. As can be seen in Table 1, labor force participation rates of men in farm and urban households remained virtually constant between 1860 and 1900. They probably declined little through the 1930s. Gratton (1986) has shown that the participation rate of men sixty-five years old and older in Boston was constant at about 64 percent through 1920 and had fallen only a few points by 1930.[2] Most of the fall in farm and urban participation rates occurred after the 1930s. For nonfarm households in rural areas and small towns, the decline was steady from 1860 through 1980 and was greater than that of farm and urban households.

Table 1 shows that because much of the decline in labor force participation occurred in the nonfarm, rural labor force, a farm/nonfarm

distinction reveals more about the sources of the decline than the urban/rural distinction does. This farm/nonfarm distinction is also apparent in the participation rates estimated from a sample collected by Fred Bateman and James Foust for northern households for 1860, although their sample covers only rural areas.[3] About 45 percent of the rural households covered by Bateman and Foust's sample were not located on farms. These households contained laborers, craftsmen, and various other workers who helped support the rural economy.

Structural Shifts in Occupations and Labor Force Participation Rates. The effect of shifting structure of occupations on older men's labor force participation can be examined in more detail by dividing the rate into household types and assigning each a weight equal to that group's importance in the population. In that way the influence of component changes on the overall labor force participation rate of men aged sixty-five and older can be assessed. (See the appendix for a formal derivation.) This decomposition reveals that overall labor force participation can change if certain occupations are becoming more or less important in the economy, even if the underlying propensity for individuals to remain at work in given occupations does not change. Although it cannot be interpreted as a behavioral model, the equation can nevertheless be useful in revealing trends and relationships not immediately apparent in aggregate series of labor force participation rates.

As the data in the table in the appendix indicate, between 1860 and 1900 about one-third of the total participation decline came from the increasing weight attached to nonfarm households in the overall average. The rest of the decline was primarily because of falling participation rates in nonfarm rural households. Of the three periods examined in the table in the appendix, the fall in overall participation was smallest—about 7 percent—between 1860 and 1900.

The importance of declining participation rates among men aged sixty-five and older increased between 1900 and 1950 as they accounted for about 80 percent of the overall decline of 33 percent. Falling participation rates of older workers appear even more significant

between 1950 and 1980, accounting for 83 percent of the overall decline of 46 percent.

To some extent this decomposition analysis could be viewed as providing evidence against a structural shift explanation of the decline in participation rates before the middle of the twentieth century. However, for most of this period about 75 percent of the older male population lived in either farm or urban households, where their participation rate remained fairly constant between 1860 and 1930. The remaining 25 percent lived in households with declining participation rates. The reasons for the decline are not yet clear, although across time the older portion of male population aged sixty-five and older in nonfarm rural households increased fairly rapidly, at least before 1950. The increasing proportion of older men may have reduced the overall participation rate of men aged sixty-five and older, even though their age-specific rates changed little in rural nonfarm households.

The insights provided by the decomposition of the overall participation rate support Gratton's claim that the structural shift away from agriculture was the major cause of the decline in the labor force participation rate of men sixty-five and older; his findings appear to be at least 75 percent correct. The decline in the participation rate of older men in nonfarm rural households, however, dominates the effect of shifting weight between farm and urban households, at least in an accounting identity sense. If the aging argument given above proves to be correct, the overall decline could well be more the result of structural and demographic changes than of changes resulting from changes in attitudes toward older workers.

Distributions of Jobs by Age— Behavioral Evidence

As compelling as the modernization theory might be, it is by no means immune to criticism. Clarence Long (1958) anticipated what is perhaps the most convincing argument against the modernization theory more than three decades ago. In his important study of the U.S. labor force, he points out that while it is possible that industrialization and technological progress may have made obsolete those older workers who could not keep pace in the changed work environment, the new machinery may have also lessened the need for sheer physical stamina. Thus, by making many tasks easier, older workers could have stayed on the job longer. Generally, however, another approach gained currency after 1970. In particular, the behavioral or "new history" view perceives older workers as having been systematically discriminated against simply because they were old, which in turn resulted in their increasing loss of access to jobs in manufacturing. This connection is hard to rationalize using standard economic theory. That is, modernization theory argues that if the value of marginal product of older male workers declined below the given wage rate, it would be expected that firms would substitute younger, more productive workers for the now expensive older workers. The new history view makes no such assumptions about the value of marginal product of older workers.

Several researchers have examined the loss of access to manufacturing and other industrial jobs by older men. Using the distribution of jobs by age for specific industries or cities as an indicator of discrimination, these researchers have indeed shown that older men were not distributed evenly across occupations.

Gratton (1986) has produced age distributions by occupation for Boston men between 1890 and 1950 from decennial census data. His results show an increasing share of men aged sixty-five and older in service (perhaps "servant") occupations across time. The proportion of men sixty-five and older who were professionals and higher-level white-collar workers was fairly stable, and the proportion who were blue-collar workers declined only slightly. Although one might view this data as supporting the behavioral or attitudinal explanation of the participation decline, Gratton actually concludes from his overall analysis of this data that rapid occupational and industrial change rather than discrimination based on age were most important in contributing to the loss of occupations among older men. Older workers who lost jobs tended to have trouble finding new ones, while workers who stayed with a firm did so because they had developed skills

the firm valued. Again, mature industries that were not changing rapidly tended to have older work forces. Gratton emphasizes that looking only at cross-sectional data can result in misleading interpretations of age distributions of workers by occupation either by misinterpreting age cohort effects or by assuming that a particular cross section reflects a turning point in attitudes toward older workers.

Roger Ransom and Richard Sutch (1986) examine several cross sections from late nineteenth-century state and national censuses and from the Michigan furniture industry. Like Gratton, they point out that histories of workers' career paths are more useful than cross sections in determining the extent and evolution of "deskilling" or "retirement on the job," two terms they use to describe how firms shifted older workers to less strenuous and less skilled occupations rather than eliminating them altogether. Ransom and Sutch's age distributions show a disproportionate number of older men in such unskilled jobs as janitors or guards, a fact they cite as evidence that older workers tended to move down the occupational ladder within a firm. Their interpretation of the evidence is questionable, however, because the workers' career histories are not known. These older workers in unskilled jobs could have come from other firms, or they could have been unskilled workers all their lives.

N. Sue Weiller (1989) looks at age distributions from the 1925 New York State Census and an employment survey conducted by the National Civic Federation in 1926. With these data she constructs the ratio of men sixty-five and older currently holding a particular job to those of the same age group who had held the job at an earlier time. She interprets a high ratio as indicating that older men were moving into the job later in life. Occupations with the highest ratios include elevator operators, guards, and janitors. Weiller found a low ratio, which indicates that fewer men remained in the occupation later in life, among factory workers, industrial laborers, and machinists. It should be pointed out, however, that a high ratio means that more older workers were coming into the occupation, not that a higher proportion of older men were employed in the particular occupation than the average for all industries. For example, factory workers ac-

counted for 12.5 percent of employed men aged sixty-five and older, and elevator operators accounted for 1.1 percent (Weiller 1989, 75).

Although the studies by Gratton (1986), Ransom and Sutch (1986), and Weiller (1989) focus on the distribution of jobs by age for different time periods, they also show that older workers were distributed unevenly across occupations. It is hard to discern a pattern, however, that would attribute the unevenness to age discrimination or at least show increasing discrimination across time.

Weiller's ratios of current to former occupations show that there were several jobs in the group that older workers could take if they had lost a better, higher-paying position. Such "retirement" jobs do not, however, account for a majority of the jobs held by employed men sixty-five and older. Indeed, jobs with high ratios tended to employ a lower share of the older male work force than did more standard jobs like factory operatives, craftsmen, or foremen (see Weiller 1989, Table 2). While it is clear that the age distributions show that some jobs held by older workers were "retirement" jobs, it is nevertheless the case that a large share of the older male work force was employed in traditional occupations in the mid-1920s.

The age distributions presented by Ransom and Sutch (1986, 22) from the decennial censuses for 1870, 1880, and 1900 show that within each year older workers were not evenly distributed across occupations and that the distribution of older workers across occupations changed little between 1870 and 1900. The number of workers aged sixty and older as a share of workers sixteen and older increased slightly in the total labor force and in most categories. Even in categories like "iron and steel workers" or, more broadly, "industrial workers" the representation by older men in the work force did not decline. This finding is surprising because these types of jobs are the ones identified by modernization theory as the ones most likely to have been cut off from older workers by the rapid pace of technological change and industrialization.

Gratton's (1986) age distributions, defined as the percent of men sixty-five and older relative to all men in certain occupations, reveal no overwhelming evidence that older workers

were increasingly being denied access to particular jobs. Their representation as iron and steel workers increased slightly between 1890 and 1930, for example. Even though older men made up an increasingly larger share of janitors, elevator operators, and guards, Gratton (1986, 83) points out that such jobs accounted for only about 7 to 8 percent of the jobs held by older men between 1920 and 1950. Gratton's study, along with Ransom and Sutch's (1986) evidence from the decennial censuses, does not indicate that older men were increasingly losing representation in the nonagricultural work force. Apparently within industries there was some sorting by age across occupations, but the sorting remained stable into the 1930s. In the work of Gratton, Ransom and Sutch, and Weiller it is hard to see the effect of increased discrimination by age as posited by the attitudinalists and some modernists.

Additional Age Distribution Evidence. The evidence discussed so far provides an incom-

plete summary of the occupational distribution of older workers across time. Several samples drawn from the manuscript schedules of the U.S. census can help to build a picture of the types of jobs older men were most likely to hold between 1860 and 1980. The addition of certain evidence from 1950 and 1980 is particularly important because some current research has suggested that government policy starting in the New Deal is responsible for hastening the decline in the labor force participation of men aged sixty-five and older (for example, Gratton 1986 and Jon R. Moen 1987).

Tables 2-4 present the age distributions of male workers in several broad categories of occupations between 1860 and 1950. The categories reflect the characteristics of occupations that may have been more or less favorable to the continued employment of older workers. For example, farmers were self-employed and could vary their work effort as

Table 2.
Age Distributions of Occupations
Rural Men, Northern United States, 1860*
(*percent*)

Age	Farmer	Professional	Skilled	Service	Labor	Other	None	All
15-19	4.1	3.9	4.9	23.3	29.9	3.8	53.9	18.0
20-24	9.4	15.0	14.6	18.4	26.9	24.7	18.7	15.8
25-34	25.9	32.0	34.1	26.9	22.7	38.7	10.1	23.9
35-44	24.2	24.3	22.6	15.3	9.9	22.1	4.1	17.6
45-59	24.7	20.2	17.1	12.1	7.2	7.2	4.0	16.3
60-69	8.3	3.7	4.5	3.8	2.5	2.6	3.3	5.5
70-79	2.9	0.9	1.9	0.3	0.8	0.4	3.7	2.3
80+	0.5	0.1	0.4	0.0	0.2	0.4	2.2	0.7
Number in Sample	15,129	1,126	3,500	365	7,829	235	4,963	33,147

* *Northern states for this sample are Connecticut, Illinois, Indiana, Kansas, Maryland, Michigan, Minnesota, New Hampshire, New Jersey, New York, Ohio, Pennsylvania, Vermont, and Wisconsin.*

Source: Estimated from the Bateman-Foust Sample.

Table 3.
Age Distributions of Occupations
All Men, United States, 1900
(*percent*)

Age	Farmer	Professional	Skilled	Service	Labor	Other	None	All
15-19	1.3	1.8	7.5	12.9	22.5	9.8	47.5	14.8
20-24	7.2	7.1	13.4	20.1	20.7	9.5	12.3	14.4
25-34	23.9	26.6	29.6	29.5	25.0	26.2	6.8	24.5
35-44	24.3	26.5	25.1	18.6	15.7	26.2	4.1	19.4
45-59	28.7	25.4	17.5	14.2	11.8	19.1	6.8	17.4
60-69	10.2	9.0	5.2	3.8	3.4	6.9	7.3	6.1
70-79	3.8	3.3	1.6	1.0	0.9	2.0	10.3	2.7
80+	0.7	0.3	0.1	0.0	0.1	0.4	4.8	0.7
Number in Sample	7,057	2,801	5,564	3,378	10,630	550	3,087	33,067

Source: Estimated from the 1900 Public Use Sample.

Table 4.
Age Distributions of Occupations
All Men, United States, 1950
(*percent*)

Age	Farmer	Professional	Skilled	Service	Labor	Other	None	All
15-19	2.1	1.0	5.2	9.1	19.6	—	39.8	13.0
20-24	5.7	5.0	12.0	12.1	16.7	—	9.3	10.7
25-34	20.3	24.8	27.9	23.3	21.3	—	7.2	21.6
35-44	25.9	30.7	26.5	21.4	17.2	—	4.7	21.1
45-59	30.0	29.7	21.6	23.8	17.1	—	8.7	20.8
60-69	12.1	7.2	5.9	8.4	6.6	—	11.8	8.1
70-79	3.5	1.6	0.8	1.8	1.4	—	13.3	3.6
80+	0.4	0.2	0.1	0.2	0.2	—	5.4	1.1
Number in Sample	2,606	3,943	10,172	4,307	4,051	—	5,466	30,545

Source: Estimated from the 1950 Public Use Sample.

they grew older. Laborers include unskilled workers, outdoor laborers, and unskilled factory workers. Service workers encompass store clerks, janitors, elevator operators, and the like. Skilled workers include factory operatives as well as self-employed craftsmen, although it is not possible to determine if a worker was self-employed. Professionals and proprietors include managers, plant foremen, and those in professional occupations such as lawyer, physician, and teacher.[4]

"Labor" and "service" occupations consistently employed a higher proportion of younger workers between 1860 and 1950 than did other occupations. Nevertheless, the shares of younger workers in these occupations declined over time, possibly because more younger men were spending increasing time in school (John Pencavel 1986, 19-21). On the other hand, occupations like "farmer," "professional," and "skilled" tended to have a higher proportion of older men in their work forces. In the case of farmers, the proportion of older men has grown somewhat since 1860, as has the share of older skilled workers. The

category "skilled workers," however, by 1950 had the lowest proportion of older workers (aged sixty and older) of all the occupational categories. The early effects of pensions and Social Security may be responsible for this decline. Skilled factory workers made up one of the more widely covered employee groups in the early years of Social Security retirement insurance payments, while farm laborers and the self-employed were not yet covered.

Tables 5-7, which display the distribution of workers in a particular age group across occupations, present a different viewpoint from that of Tables 2-4, which show the age distribution of workers within an occupational category. For example, in Table 5, 63 percent of men aged thirty-five through forty-four in 1860 were farmers, while 4.7 percent were professionals or proprietors. Reading down a column in Tables 5-7 allows comparison of an occupation's importance across cohorts. In 1860 only 3.5 percent of men aged thirty-five to forty-four had no occupation, while 22.5 percent of men in the seventy to seventy-nine cohort had none.

Table 5.
Shares of Cohorts Employed in Different Occupations
Rural Men, Northern United States, 1860
(*percent*)

Age	Farmer	Professional	Skilled	Service	Labor	Other	None	Number in Sample
15-19	10.5	0.7	2.9	1.4	39.3	0.2	44.9	5,950
20-24	27.0	3.2	9.7	1.3	40.0	1.1	17.7	5,252
25-34	49.4	4.5	15.1	1.2	22.4	1.2	17.7	7,935
35-44	63.0	4.7	13.6	1.0	13.3	0.9	3.5	5,817
45-59	69.4	4.2	11.1	0.8	10.5	0.3	3.6	5,390
60-69	68.4	2.3	8.5	0.8	10.7	0.3	9.0	1,831
70-79	57.3	1.3	8.6	0.1	8.0	0.1	22.5	752
80+	36.8	0.5	5.9	0.0	6.4	0.5	50.0	220

Reading across a row gives the distribution of occupations for a particular age group. Reading down a column gives the share in each cohort that is in a particular occupation.

Source: Estimated from the Bateman-Foust Sample.

Table 6.
Shares of Cohorts Employed in Different Occupations
All Men, United States, 1900
(*percent*)

Age	Farmer	Professional	Skilled	Service	Labor	Other	None	Number in Sample
15-19	1.8	1.0	8.5	8.9	48.8	1.1	29.9	4,900
20-24	10.7	4.2	15.7	14.2	46.2	1.1	8.0	4,772
25-34	20.8	9.2	20.4	12.3	32.8	1.8	2.6	8,084
35-44	26.7	11.6	21.8	9.8	26.0	2.3	2.0	6,413
45-59	35.2	12.4	16.9	8.4	21.7	1.8	3.7	5,751
60-69	35.7	12.5	14.3	6.3	18.2	1.9	11.2	2,017
70-79	29.5	10.1	10.0	3.6	10.0	1.2	35.5	898
80+	22.4	3.5	3.0	0.4	5.6	0.9	64.2	232

Reading across a row gives the distribution of occupations for a particular age group. Reading down a column gives the share in each cohort that is in a particular occupation.

Source: Estimated from the 1900 Public Use Sample.

Table 7.
Shares of Cohorts Employed in Different Occupations
All Men, United States, 1950
(*percent*)

Age	Farmer	Professional	Skilled	Service	Labor	Other	None	Number in Sample
15-19	1.4	1.0	13.3	9.9	19.9	—	54.6	3,983
20-24	4.6	6.0	37.3	15.9	20.7	—	15.6	3,273
25-34	8.0	14.8	43.0	15.2	13.0	—	6.0	6,603
35-44	10.5	18.8	41.8	14.3	10.8	—	3.9	6,448
45-59	12.3	18.5	34.7	16.2	10.9	—	7.4	6,337
60-69	12.7	11.5	24.3	14.7	10.9	—	26.0	2,474
70-79	8.4	5.7	7.2	6.9	5.2	—	66.6	1,095
80+	3.0	1.8	1.8	2.1	2.4	—	88.7	332

Reading across a row gives the distribution of occupations for a particular age group. Reading down a column gives the share in each cohort that is in a particular occupation.

Source: Estimated from the 1950 Public Use Sample.

Between 1860 and 1950, a decreasing share of each successively older age group is employed in every occupational category. Several categories, however, show that the proportion of each successive cohort in the category fell off less rapidly than in other categories. Skilled workers and farmers show fairly steady shares of workers across most cohorts in 1860. The same is true for 1900. Both years showed a decline in the share of each successive cohort after age twenty-four employed as laborers. It was probably easy to enter such jobs early in life and then to leave as various skills were acquired. Nevertheless, a large share of workers aged sixty and older were still employed as laborers in 1860 and 1900. By 1950 laborers and farmers included a much smaller share of older workers. Indeed, in all job categories in 1950 a sharp drop-off appears between age groups sixty to sixty-nine and seventy to seventy-nine in the share of the cohort employed. That the decline appears in all categories suggests that something common across occupations rather than characteristics specific to the occupations themselves was affecting employment after the early 1900s.

Like the evidence presented earlier by other researchers, the age and occupation distributions in Tables 2-7 do not immediately support the contention that age discrimination increased in the late nineteenth and early twentieth centuries. This conclusion does not necessarily imply that older workers were not discriminated against, either because of decreased productivity or simple prejudice against the aged. Some occupations do have older workers disproportionately represented. Rather, it can be inferred that whatever discrimination or prejudice may have existed did not increase over time. Not until 1950 is there any evidence of systematic disappearance of older workers from the labor force, and then the decline is apparent across all occupations.

Economic and Governmental Factors Leading to the Decline

The labor force participation rate of men aged sixty-five and older remained fairly constant at about 60 percent between 1910 and 1930. Then, after at least twenty years of relative stability, it fell by one-third in twenty years. Certainly the Great Depression was immediately responsible for some of the decline. Perhaps as early as 1940, however, and definitely by 1950, participation was falling across all types of households, farm as well as nonfarm rural and urban. The decline has continued steadily through today. A changing mix of occupations does not account for much of the decline after 1950.

Economists have offered many theories, but none seems to explain the decline adequately. William G. Bowen and T. Aldrich Finegan's (1969, 374) catalog of potential sources of the labor force participation decline between 1948 and 1965 shows increases in "other income," including Social Security benefits, accounting for around half. However, the effect of Social Security and private pensions on retirement is a subject of continuing debate among economists. (See Edward Lazear 1986, 325-30, for a survey of the literature.) For example, Gratton (1988) has shown that in 1950 relief payments for the aged under Title I of the Social Security Act may have affected labor force participation more than payments under Title II of the act did.

Whatever in fact initiated the latest decline, changed attitudes toward older workers do not appear to be a significant factor. As hinted at earlier, the federal government may have had an important role in changing the status of older workers through New Deal legislation. The Social Security Act of 1935 is usually cited as the key piece of social legislation responsible for beginning the massive withdrawal of older men from the work force. In addition, several other pieces of legislation may have equally affected the status of older workers, even if unintentionally. The National Industrial Recovery Act, which authorized the National Recovery Administration in 1933, the National Labor Relations Act in 1935, and the Fair Labor Standards Act of 1938 all made a difference.

Social Security was designed to aid older workers directly, either by providing aid to those who had lost jobs or, as William Graebner (1980) and Carol Haber (1983) have argued, by drawing them out of the labor force

to open up jobs for younger workers. The effects of Social Security, however, could not have been felt immediately because, under the 1935 act, payments were not to have begun until 1942. (Later revisions moved payments forward to 1940.) The effects of the National Industrial Recovery Act, the National Labor Relations Act, and the Fair Labor Standards Act could have been felt much sooner.

The National Industrial Recovery Act, passed in 1933, authorized the National Recovery Administration to allow industries to collude to set output quotas and raise prices to increase profits. To diminish what were then thought to be free competition's most destructive aspects, industries were instructed to establish codes of fair competition to prevent individual firms from undercutting each other's profits. The industry codes approved by the National Recovery Act also contained clauses that established minimum wages and maximum weekly hours, after which overtime was to be paid. The intent was to spread a given number of hours of work among more workers while limiting output and maintaining or raising real wages. The number of industrial employees working in an industry covered by a code grew rapidly, and within a year almost 95 percent of industrial workers may have come under the protection of an approved code.

By limiting hours of work and at the same time supporting or raising wage rates, the National Recovery Act codes gave employers an incentive to release their least productive workers and retain those who could work more intensively. In other words, employers now had a reason to discriminate against less productive workers, including older workers, who previously had been able to obtain employment because competition had kept their wages low. The legislation allowed industries to develop additional codes, and some adopted rules under which older workers could be defined as handicapped on the basis of age alone and receive a wage lower than the minimum specified for most workers (Graebner 1980, 209). In effect, federal government policy had made older workers potentially obsolete. Prejudice against the aged, as described by David Hackett Fischer (1978) and others, may have had little effect on old-

er workers before the New Deal, but the National Recovery Act codes gave employers a reason to discriminate on the basis of age.

Although the National Recovery Administration was declared unconstitutional in May 1935, the National Labor Relations Act, established that same year, extended most of the National Recovery Administration's provisions on unionization and collective bargaining. In 1938 the Fair Labor Standards Act extended most of the provisions on hours of work and minimum wages that had been established in the National Recovery Administration's industrial codes.

Even if these pieces of legislation did change employers' attitudes toward older workers, the relevant question is, how much did they actually affect the employment of older workers? Charles Roos, the director of research for the National Recovery Administration, pointed out that the wage and hour provisions decreased the employment of several demographic groups, including older workers (1937, 174-75, 193). The minimum wage is not a binding constraint in many industries today, but Gavin Wright has shown that during the 1930s, establishment of the minimum wage raised wages for some workers in the South, although both the National Recovery Administration and the Fair Labor Standards Act added to unemployment of blacks, another demographic group significantly, if indirectly, affected by the wage and hour provisions (1986, 220, 223-25). Michael Weinstein (1980) has shown that the National Industrial Recovery Act significantly affected income distribution and hours of work in the overall economy. Increases in real wages also contributed to overall unemployment, although he does not distinguish any evidence for different age groups.

The effect of the National Recovery Administration and subsequent legislation on older men's employment and hours of work is difficult to quantify in the aggregate for the 1930s; however, the anecdotal evidence makes it clear that the employment of older men was not advanced by the National Recovery Administration or the Fair Labor Standards Act. Pencavel (1986, 13) has noted that the overall decline in hours worked per week by all men between 1929 and 1940 was caused in part by

Fair Labor Standard Act overtime pay requirements. Since at least 1955, average weekly hours worked and the labor force participation rate of men sixty-five and older have both fallen, while hours worked per week by younger men have stayed fairly constant or have risen slightly (Pencavel 1986, 16). Little or no direct evidence on work hours by age for industries in the 1930s is available that would help identify clearly the effect of the National Recovery Administration on the labor force participation rate of older men. Current research indicates that restrictions on work hours, minimum wages, and overtime pay provisions have not been significant factors recently in the decline of the labor force participation rate of men sixty-five and older. Nevertheless, such provisions during the 1930s may have helped establish the precedent that withdrawal from the labor force later in life was an appropriate and expected part of one's career.

Conclusion

The evidence presented and reviewed in this article does not support the view that older male workers became increasingly cut off from work during the late nineteenth and early twentieth centuries because of systematic discrimination or adverse changes in attitudes toward older members of society. For fairly broad occupational categories, the age and occupation distributions show little change in the representation of older male workers. Although it is clear that some occupations were increasingly becoming identified as jobs for older workers, such positions were not a significant proportion of all occupations held by men sixty-five and older.

Even if attitudes toward older workers worsened, the most significant development was the employment policies of the New Deal, which set the stage for the virtual disappearance of older men from the work force. Their sudden sharp decline from all occupations beginning around 1950 demonstrates that retirement is a modern aspect of work. While the National Recovery Administration and Fair Labor Standards Act, perhaps inadvertently, may have provided employers with a reason to furlough older workers, a farther-reaching result of these pieces of legislation may have been the institutionalization of retirement as a standard stage in the life of one's career. As the United States faces a shrinking supply of younger workers—and a growing population aged sixty-five and older who could contribute to the work force—in the next few decades, policies less encouraging to full retirement may become increasingly desirable.

Appendix

The force participation rate, p, is defined as

$$p = L/N, \qquad (1)$$

where L is the labor force, that is, the summed total of those employed and seeking employment, and N is the relevant total population. In the context of this article all variables refer to men aged sixty-five and older.

The participation rate in equation (1) can be rewritten in a manner that allows an examination of changes in subsector participation rates as well as shifts in the shares of the older male population in farming and nonfarming. If the labor force, L, is broken into a farm, L_f, and non-farm, L_n, component, equation (1) can be rewritten as

$$p = (L_f + L_n)/N. \qquad (2)$$

A similar decomposition of the population variable, N, yields, after some manipulation,

$$p = (N_f/N)\,(L_f/N_f) + (N_n/N)\,(L_n/N_n), \qquad (3)$$

which can be rewritten as

$$p = a_f p_f + a_n p_n. \qquad (4)$$

In equation (4) a_f and a_n are the shares of the

older male population in farming and nonfarming, and p_f and p_n are the participation rates of older men in the farming and nonfarming sectors. Equation (4) expresses the labor force participation rate as a weighted average of labor force participation in farming and nonfarming, where the weights are the shares of the older male population in farming and nonfarming.

If the shares of the older male labor force in farming and nonfarming are denoted by l_f and l_n, respectively, and equation (4) is rewritten in percentage rate of change form, the dynamic behavior of the labor force participation rate can be examined. Performing these calculations yields

$$\dot{p} = l_f(\dot{a}_f + \dot{p}_f) + l_n(\dot{a}_n + \dot{p}_n), \qquad (5)$$

with the dot indicating the derivative of the natural logarithm of the particular variable, which is also the instantaneous rate of change of the variable. As in equation (4), equation (5) expresses the overall change in labor force participation, $\dot{p}$, as a weighted average of the changing shares of men living in different household types as well as changing labor force participation rates within those sectors.

One difficulty with using identities such as equation (5) with historical census data is that the census enumerators were not instructed to record an individual's previous occupation if he was recorded as currently having no occupation. The application of equation (5) to historical census data assumed that if a man with no occupation was living on a farm, his previous occupation was in farming. Likewise, for men living in nonfarm households with no current occupation, it is assumed that their previous occupation was in the nonfarming sector.

With the aid of the decomposed participation rate, the relative importance of declining labor force participation and shifting weights between farm and nonfarm households from 1860 through 1980 can be examined specifically. Estimated values of the variables in equations (4) and (5) are given in the table below for the periods 1860-1900, 1900-50, and 1950-80. These estimates were derived from the data displayed in Table 1.

The Shift of American Men Aged Sixty-Five and Older from Farm to Nonfarm Households, 1860-1980

$$p = a_f p_f + a_n p_n$$

$$\dot{p} = l_f (\dot{a}_f + \dot{p}_f) + l_n (\dot{a}_n + \dot{p}_n)$$

p_i = LFP rate of older males in household type i

a_i = share of older male population in household type i

l_i = share of older male labor force living in household type i (mid-period value)

$\dot{p}_i$ = average annual rate of change in the LFP rate of older males in household type i

$\dot{a}_i$ = average annual rate of change in the share of older male population living in household type i

i = f, n (farm, nonfarm)

Values of Variables for Households, 1860-1900

	a_f	a_n		p	p_f	p_n
1860	.54	.46		.76	.80	.71
1900	.41	.59		.70	.80	.63
	$\dot{a}_f$	$\dot{a}_n$		$\dot{p}$	$\dot{p}_f$	$\dot{p}_n$
	−.007	.007		−.002	.000	−.003

	l_f	l_n
	.521	.479

Continued

Values of Variables for Households, 1900-1950

	a_f	a_n		p	p_f	p_n
1900	.41	.59		.70	.80	.63
1950	.16	.84		.47	.62	.44
	$\dot{a}_f$	$\dot{a}_n$		$\dot{p}$	$\dot{p}_f$	$\dot{p}_n$
	−.018	.007		−.008	−.005	−.007

	l_f	l_n
	.342	.657

Values of Variables for Households, 1950-1980

	a_f	a_n		p	p_f	p_n
1950	.16	.84		.47	.62	.44
1980	.03	.97		.26	.58	.25
	$\dot{a}_f$	$\dot{a}_n$		$\dot{p}$	$\dot{p}_f$	$\dot{p}_n$
	−.051	.005		−.020	−.002	−.019

	l_f	l_n
	.146	.853

Notes

[1] The "new history" viewpoint originated in the work of Fischer (1978) and was further developed by Achenbaum (1978), Graebner (1980), and Haber (1983), among others.

[2] This finding is consistent with the estimate for 1860 of 64.2 percent made by Moen (1987) based on a sample of men aged sixty-five and older living in New York, Boston, Philadelphia, and Chicago.

[3] The Bateman-Foust Sample of Northern Households, 1860, covers 21,118 rural households in 102 randomly selected townships in the north and north-central states. The sample is drawn from manuscript schedules of the U.S. Census of 1860 and is available through the Interuniversity Consortium for Political and Social Research in Ann Arbor, Michigan.

[4] Complete listings of occupations in each category are available from the author.

References

Achenbaum, W. Andrew. *Old Age in a New Land: the American Experience Since* 1790. Baltimore: The Johns Hopkins Press, 1978.

Bowen, William G., and T. Aldrich Finegan. *The Economics of Labor Force Participation*. Princeton, N.J.: Princeton University Press, 1969.

Epstein, Abraham. *The Challenge of the Aged*. New York: Vanguard Press, 1928.

Fischer, David Hackett. *Growing Old in America*. New York: Oxford University Press, 1978.

Graebner, William. A *History of Retirement: The Meaning and Function of an American Institution, 1885-1978.* New Haven, Conn.: Yale University Press, 1980.
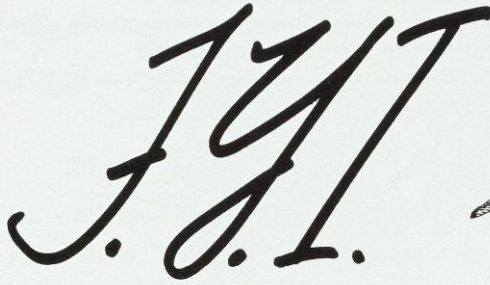
Gratton, Brian. *Urban Elders*. Philadelphia: Temple University Press, 1986.

_____. "The New Welfare State: Social Security and Retirement in 1950." *Social Science History* 12 (Summer 1988): 171-96.

Haber, Carol. *Beyond Sixty-Five: The Dilemma of Old Age in America's Past*. New York and London: Cambridge University Press, 1983.

Lazear, Edward. "Retirement from the Labor Force." In *The Handbook of Labor Economics*, edited by Orley Ashenfelter and Richard Layard, vol. 1, 305-56. New York: North Holland, 1986.

Long, Clarence. *The Labor Force under Changing Income and Employment*. Princeton, N.J.: Princeton University Press, 1958.

Moen, Jon R. "Essays on the Labor Force and Labor Force Participation Rates: The United States from 1860 through 1950." Ph.D. dissertation, University of Chicago, 1987.

Pencavel, John. "Labor Supply of Men: A Survey." In *The Handbook of Labor Economics*, edited by Orley Ashenfelter and Richard Layard, vol. 1, 3-102. New York: North Holland, 1986.

Ransom, Roger, and Richard Sutch. "The Labor of Older Americans; Retirement on and off the Job, 1870-1937." *The Journal of Economic History* 46 (March 1986): 1-30.

Roos, Charles. NRA *Economic Planning*. Bloomington, Ill.: Principia Press, 1937.

Weiller, N. Sue. "Industrial Scrap Heap: Employment Patterns and Change for the Aged in the 1920s." *Social Science History* 13 (Spring 1989): 65-88.

Weinstein, Michael. *Recovery and Redistribution under the NIRA*. New York: North Holland, 1980.

Wright, Gavin. *Old South, New South*. New York: Basic Books, 1986.

Southeastern Interstate Banking and Consolidation: 1984-89

Robert E. Goudreau and Larry D. Wall

As late as 1982 interstate mergers and acquisitions were effectively prohibited in all fifty states and the District of Columbia.[1] A flurry of changes in state laws has occurred since 1982, however, and only three states currently prohibit interstate acquisitions entirely. States have relaxed restrictions on intrastate and interstate banking largely because of increases in the level and volatility of interest rates and changes in technology that have allowed foreign and nonbank firms offering close substitutes for traditional bank products to lure banks' customers away.

These changes in state banking laws have permitted a restructuring of the banking system that may have a number of public policy implications. For example, large-scale consolidation may complicate deposit insurance reform if the "too-big-to-fail" doctrine is maintained by increasing the number of banking organizations that become too big to fail.

However, if this doctrine is eliminated, consolidation could further reform by increasing the number of organizations subject to discipline by savvy institutional investors. Additionally, if greater consolidation boosts the efficiency of U.S. banks in global financial markets, the banking system (and perhaps U.S. nonfinancial corporations) could become more competitive. The U.S. experience may also provide examples to regulators and bankers in other parts of the world, such as the European Community, where national barriers to bank mergers and acquisitions are being relaxed.

Although it is too early to know what final changes consolidation will effect in the banking system, an examination of merger and acquisition activity to date may provide valuable insight about the direction in which the banking industry is headed. This study is a review of bank mergers and acquisitions in the Southeast during the period 1984 through 1989, the year preceding and roughly five years following enactment of initial interstate banking laws for the region in July 1985.[2] The

The authors are, respectively, an assistant economist in and the research officer in charge of the financial section of the Atlanta Fed's research department. They thank David L. Verlander and Sherley Wilson for valuable research assistance.

analysis below starts with a brief look at changes in the relevant state and federal laws limiting intrastate and interstate banking. A variety of interstate banking laws that make up the "Southeast Compact" are examined. Following a review of the economic reasons banks might choose to participate in mergers and acquisitions, there is a discussion of trends in four categories of mergers and acquisitions that have affected southeastern depository institutions: intrastate transactions, interstate transactions within the region, the acquisition of out-of-region banks by southeastern institutions, and the acquisition of regional banks by out-of-region organizations. The concluding section contains a review of the most significant bank holding company participants in interstate acquisitions for the Southeast and an examination of their apparent expansion strategies.

Legal Limitations

Intrastate Limits in the Southeast. Consolidation of existing banking organizations can take several legal forms. First, bank holding company acquisitions (without the subsequent merger of the acquired institution into a holding company subsidiary) may provide additional bank subsidiaries for the acquirer. Another approach could be for a bank holding company to structure the transaction solely as a merger of one of its subsidiary banks with a subsidiary bank of another holding company or as a merger with an independent bank. Third, independent banks (those not affiliated with a holding company) may merge, with the surviving bank possessing the assets, deposits, and branch offices of the combined organization.

State and local laws constrain these means of intrastate expansion in various ways. Branching laws for all southeastern states except Georgia were recently eased to allow statewide branching, but they remained restrictive for most of the 1984-89 period, which is the focus of this analysis.[3] During the six years all of the southeastern states permitted statewide branching through mergers, and banking offices of the merged (target) firm became branch offices of the surviving organization. After 1985 all southeastern states except Mississippi allowed expansion through multiple bank holding company acquisitions.

Local laws and general laws of local application written by county officials governed branching rights for Alabama-chartered banks from 1984 to 1989. The effect of having local laws for banks in a particular county was to limit the establishment of de novo branch banks to the confines of the county. State banking statutes in place for most of the period for Florida, Georgia, Louisiana, and Tennessee explicitly enumerated de novo branching rights and typically constrained branching to the parent bank's home county. (Louisiana, Florida, and Tennessee statutes permitted statewide branching in 1988, 1989, and 1990, respectively.) Banking laws in Mississippi, where multibank holding companies are prohibited, allowed de novo branch establishment within a steadily expanding geographic radius from the parent bank's home office until statewide branching was ultimately permitted in July 1989.[4]

Interstate Limits—A National Perspective. National banking laws have also constrained banking organizations in individual states and the District of Columbia from acquiring banks and thrifts in other states.[5] In 1927 the McFadden Act (and later its 1933 amendments) effectively limited national banks' branching activity to a single state at most. Although the legislation left open the possibility of interstate bank acquisitions through the bank holding company structure, it was not broadly used for interstate expansion before passage of the Bank Holding Company Act of 1956. Nonetheless, concerns about the potential widespread use of this vehicle provided sufficient reason to prompt adding the Douglas Amendment, which barred bank holding companies that had not already done so from acquiring banks outside their home state unless the other state explicitly allowed such acquisitions.[6]

The McFadden Act and the Bank Holding Company Act's Douglas Amendment effectively closed the door to full-service interstate banking from 1956 to 1982, when New York passed a nationwide reciprocal banking law and Massachusetts lawmakers took the first

step in establishing the New England banking compact by approving a regional, reciprocal interstate banking statute.

The momentum for interstate banking grew, and by 1985 a number of states across the nation had passed or were contemplating passage of their individually tailored interstate laws. In addition to such state legislation, a number of federal statutes—the Garn-St Germain Act of 1982, the Competitive Equality in Banking Act of 1987, and the Financial Institutions Reform, Recovery, and Enforcement Act of 1989—have been enacted, allowing bank holding companies to acquire out-of-state banks and thrifts.

Southeast Compact. Although individual state laws vary, southeastern interstate banking statutes typically encompass the states of Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia, as well as the District of Columbia. Not surprisingly, interstate banking laws for states situated on the outskirts of the region include neighboring states.

Florida and Georgia lawmakers were the first in the Southeast to approve regional, reciprocal interstate banking statutes. They did so in 1984. (See the box on page 52 for a description of interstate banking laws for southeastern states.) During the same year Kentucky, North Carolina, and South Carolina legislators crafted their individual versions of regional, reciprocal interstate laws. Tennessee, Maryland, Virginia, and the District of Columbia lawmakers followed suit in 1985, and by 1986 Alabama, Louisiana, and Mississippi had their own interstate laws. Alabama and Louisiana statutes, however, did not become effective until July 1987, and interstate transactions for Mississippi banks could not take place until July 1988. Arkansas legislators were the last in the Southeast to approve interstate banking statutes. Approval in Arkansas came during 1988, but interstate transactions involving Arkansas banks could not occur until January 1989.

Although most southeastern state laws on interstate banking remain regional and reciprocal in scope, a few states have dropped their initial regional restriction and now operate under statutes that are national and recip-

rocal. For example, Kentucky and Louisiana adopted national, reciprocal statutes in July 1986 and January 1989, respectively, after initially adhering to provisions requiring reciprocity with southeastern partner states. District of Columbia law calls for reciprocity with eleven southeastern states but allows acquisitions of District of Columbia banks by bank holding companies outside its southeastern region. Acquirers from outside the Southeast, however, must make certain commitments regarding their investments and operations in the District of Columbia. West Virginia's interstate banking law, which became effective January 1988, has always been reciprocal and national in scope (*Banking Expansion Reporter* 1990, 7-10).

Economic Incentives

State and federal laws determine which mergers and acquisitions are legally permissible, but economic incentives decide which ones actually occur. A variety of motivations have been suggested. First, the target bank's managers and shareholders may want to improve its performance; the acquirer simply may possess better management. Alternatively, the target's managers and owners may be risk averse and inclined to turn down potentially profitable loans, while a larger, perhaps publicly traded, organization could more comfortably make such loans because of its greater diversification or willingness to tolerate added risk.

A second reason for bank acquisitions may be to diversify both the funding sources and earnings of the acquirer. Large acquirers that rely on purchased funds may be especially interested in buying banks with significant core deposit funding bases. Although the deregulation of deposit interest rates has reduced cost advantages of relying on core deposits, they are still highly valued because of their greater stability relative to purchased funds.[7] Similarly, diversification of earnings, both geographically and by customer type, can reduce the overall credit riskiness of a bank's consumer and commercial loan portfolio.

Aside from these reasons, some acquirers believe that as a larger organization they could take advantage of available economies of scale or scope. Other organizations may want to expand enough to preclude being swallowed by a potential acquirer. Furthermore, managers may move to increase their bank's size so they can enjoy the higher salaries and managerial perquisites larger banks typically provide. Finally, an acquirer may wish to increase market power and thereby reduce competition. If a less competitive environment translates into higher profits or reduced risk for acquirers, the bank's shareholders can benefit. However, these gains typically come at the expense of the public, which must pay higher prices for bank services.[8]

Merger and Acquisition Activity—Yearly Changes

After a number of large regional banks consolidated to build their interstate franchises, the yearly dollar volume of intraregional transactions declined steadily. However, the number of such transactions doubled and spread across all southeastern states. Out-of-region banks were very active in purchasing regional banks throughout the six-year period. Southeastern banks acquired a number of depository institutions outside the region.

The most significant participants in southeastern acquisitions were large bank holding companies. These large holding companies took control of substantial asset holdings and usually acquired a sizable number of banking firms. With few exceptions, they employed three main strategies. First, many holding companies sought to acquire large banks or thrifts in nearby or distant states and become major competitors from the start. Later they increased their market presence by acquiring smaller banks as opportunities arose. Second, some companies unwilling or unable to acquire large-scale targets achieved a major presence through a series of smaller acquisitions. Third, some larger bank holding companies chose to purchase small or mid-sized banks in contiguous or nearby markets. They

essentially appended the acquired bank's operations to their own.

This study demonstrates the desire, perhaps pent-up demand, for interstate banking. Many banking organizations, large and small, sought to expand their banking operations immediately after laws opened state borders to interstate transactions. Profitable home-state acquisition targets presumably had become scarce. The primary motivations for interstate expansion appear to be greater market influence, economies of scale and scope, and earnings diversification.

Table 1 provides an overall picture of merger and acquisition activity in the Southeast during the years 1984 through 1989. Annual summary figures for mergers and acquisitions approved by federal regulators appear in the table. The statewide summary figures presented concern the total number of transactions and total dollar volume of transactions federal regulators approved during each of the six years under study. Total asset values reported are for the acquired banking organization (target bank) as of December 31 of the year preceding approval. (For example, for transactions approved by federal regulators during the year 1988, the total asset values for the acquired banking firms as of December 31, 1987, are used in compiling summary figures.) The table also provides yearly totals for the region. The different types of transactions reviewed are (1) intrastate transactions for banking firms located within each southeastern state, (2) intraregional transactions among depository institutions headquartered within the Southeast, (3) acquisitions of banking firms outside the region by southeastern banking institutions, and (4) acquisitions of southeastern banking firms by out-of-region acquirers.

Several patterns are apparent in the statewide summary figures. Local economic conditions appear to be an important determinant of merger activity. Out-of-state acquirers were particularly attracted by strong growth in Florida. Acquirers from within the region and from North Carolina and New York made substantial acquisitions.[9] Georgia and Tennessee, both of which experienced significant growth for at least part of the period, also attracted considerable out-of-state interest.

Table 1.
Mergers and Acquisitions in Southeastern States*

State	Intrastate Trans-actions	Intrastate Total Assets	Intraregional Acquirer Located In Trans-actions	Intraregional Target Located In Trans-actions	Intraregional Target Located In Total Assets	Regional Acquirer/Nonregional Target Trans-actions	Regional Acquirer/Nonregional Target Total Assets	District of Columbia Trans-actions	District of Columbia Total Assets	New York Trans-actions	New York Total Assets	North Carolina Trans-actions	North Carolina Total Assets	Virginia Trans-actions	Virginia Total Assets	Total Trans-actions	Total Total Assets
1984																	
Alabama	8	319															
Florida	26	4,658								1	1,945	1	1,695			2	3,640
Georgia	28	1,962															
Louisiana	5	2,164															
Mississippi	8	693															
Tennessee	23	609															
Southeastern Total	98	10,405															
1985																	
Alabama	14	1,706															
Florida	14	476		4	13,239							3	5,493			3	5,493
Georgia	27	1,177	4									2	8,247			2	8,247
Louisiana	18	5,190															
Mississippi	7	381															
Tennessee	16	1,468				1	166										
Southeastern Total	96	10,398	4	4	13,239	1	166										
1986																	
Alabama	18	726															
Florida	13	1,736	2	4	341					1	592	4	1,474			5	2,066
Georgia	26	933	5	2	464	3	3,003					5	3,707			5	3,707
Louisiana	15	1,183															
Mississippi	10	1,495															
Tennessee	17	972		1	5,062	1	167							1	589	1	589
Southeastern Total	99	7,045	7	7	5,867	4	3,170										

State	Intrastate Trans-actions	Intrastate Total Assets	Intraregional Acquirer Located In Trans-actions	Intraregional Target Located In Trans-actions	Intraregional Target Located In Total Assets	Regional Acquirer/ Nonregional Target Trans-actions	Regional Acquirer/ Nonregional Target Total Assets	District of Columbia Trans-actions	District of Columbia Total Assets	New York Trans-actions	New York Total Assets	North Carolina Trans-actions	North Carolina Total Assets	Virginia Trans-actions	Virginia Total Assets	Total Trans-actions	Total Assets
1987																	
Alabama	11	766	6														
Florida	11	2,787		7	1,228							7	419			7	419
Georgia	17	973	4					1	1,600			3	265			4	1,865
Louisiana	18	725															
Mississippi	11	757															
Tennessee	16	991		3	261							1	9	10	4,903	11	4,912
Southeastern Total	84	6,999	10	10	1,489												
1988																	
Alabama	5	128	5	1	40	3	66										
Florida	11	1,618	4	7	809					1	629	2	1,033			3	1,662
Georgia	6	242	5	4	598												
Louisiana	13	924		1	259												
Mississippi	3	338	1														
Tennessee	11	915		2	53									5	514	5	514
Southeastern Total	49	4,165	15	15	1,759	3	66										
1989																	
Alabama	4	83	3	2	57	1	84										
Florida	6	137	1	8	527			1	36			3	9,373			4	9,409
Georgia	8	338	7	2	116							1	35			1	35
Louisiana	28	1,510		1	1,041	10	401										
Mississippi	2	138	1	2	497												
Tennessee	4	113	3			1	41										
Southeastern Total	52	2,319	15	15	2,238	12	526										

* Total assets are given in millions of dollars as of December 31 of the preceding year.

Sources: Compiled by the Federal Reserve Bank of Atlanta from the *Federal Reserve Bulletin*, Board of Governors of the Federal Reserve System, 1984-89; *Merger Decisions*, Federal Deposit Insurance Corporation, 1984-89; and *Quarterly Journal*, Office of the Comptroller of the Currency, 1984-89.

The largest Georgia acquisitions made by organizations outside the six-state region originated in North Carolina and the District of Columbia, whereas the largest out-of-region acquirers of Tennessee banks came from Virginia.

Mississippi and Alabama generally had slower growth, which encouraged far fewer out-of-state acquisitions of their banks. The low rate of interstate acquisitions for the two states also reflects later trigger dates and an initial contiguous-state limitation for Mississippi's allowable region. Louisiana experienced by far the worst economic conditions in the region, but substantially more Louisiana assets were purchased by out-of-state organizations than in either Alabama or Mississippi. Although it is possible that the larger size of the Louisiana transactions aroused greater interest by outside acquirers, it is far more likely that the transactions resulted from an increased willingness on the part of Louisiana banks to seek financial and managerial support from larger out-of-state organizations. It is important to note that no Alabama, Louisiana, or Mississippi banks were acquired by out-of-region bank holding companies during the 1984-89 period.

Another pattern in the data demonstrates that the degree of home-market consolidation is an important consideration affecting an organization's absorption by an out-of-state institution or its ability to acquire banks outside its state. Florida organizations, for example, had not made substantial acquisitions of other Florida banks before the advent of interstate banking. Thus, the state's large banks had not yet consolidated their holdings into efficient statewide networks. These banks were in a weak competitive position to look for acquisition targets outside Florida and were themselves relatively easy targets for outside acquirers experienced in running large banking organizations. In contrast, North Carolina banks have operated under a statewide branching law since the nineteenth century. North Carolina banks were well positioned to spread throughout the Southeast with efficient, consolidated operations in their home markets and the requisite expertise needed to run large branch operations in states where they made acquisitions. Somewhere between Florida and North Carolina

banks' experience, Georgia's organizations have had less time and opportunity to develop statewide networks than North Carolina banks but have consolidated more operations than Florida banks. As a result, North Carolina and Georgia banks tended to acquire Florida institutions, and North Carolina banks dominated transactions with Georgia banking organizations.[10]

The effect of interstate banking laws can also be seen in the limited number of organizations outside the region acquired by southeastern banks. Only one large out-of-region acquisition was completed by a regional bank during the 1984-89 period: Citizens and Southern Corporation of Georgia acquired Citizens and Southern Corporation of South Carolina.[11] However, nineteen small banks outside the six-state region were acquired by southeastern bank holding companies during the six-year period, and total assets for these acquisitions equaled $1.1 billion. Tennessee bank holding companies expanded into Kentucky in 1985 and 1986. In 1988 Alabama bank holding companies entered Texas and South Carolina. Alabama and Louisiana banks expanded or commenced operations in Texas, and Tennessee banks continued to purchase Kentucky institutions during 1989. During that year Hibernia Corporation alone acquired ten small Texas banks with total assets of $401 million.

A third apparent trend for most states is that large acquisitions tend to be made early. The peak year for total volume of assets in interstate transactions involving southeastern organizations was the first full year of the compact, 1985, when $27 billion of Florida and Georgia banking assets were acquired. Similarly, the biggest years for volume of Tennessee assets acquired were 1986 and 1987 (interstate acquisition of Tennessee banks was first permitted in the latter half of 1985). Explanations for this trend lie in the motives of both acquirers and targets. Large acquirers wish to take over organizations with substantial market share in the target state, and they want opportunities to realize significant organizationwide economies of scale or scope. Moreover, banks that are probable targets try to find acceptable acquirers before less desirable ones make offers that are too good for

the target's shareholders to refuse. Thus, larger banks interested in being an acquirer or target have an incentive to find a suitable partner within the first few years after a change in interstate banking laws.

A fourth trend also appears in the data. The marked overall increase in interstate banking transactions apparently reduced bankers' interest in intrastate transactions. Although total transactions were approximately 100 per year (except for 1988 and 1989 when only about fifty transactions occurred), total dollar volume of southeastern intrastate mergers and acquisitions equaled $10 billion in 1984 and 1985 and then fell steadily to $2 billion in 1989. This trend can be accounted for in part by the fact that during the latter part of the sample period a number of potential in-state acquirers were themselves acquired by out-of-state organizations. Thus, some of the transactions labeled interstate acquisitions probably would have occurred as intrastate deals if interstate banking had not been an option.

Figures in Table 2 illustrate the dramatic transfer of state-controlled banking assets to out-of-state bank holding companies from December 1984 to December 1989. In 1984 only three Florida banks with $4 billion in total assets were controlled by out-of-state institutions at year's end. These holdings equaled 5.2 percent of total assets for Florida commercial banks and 1.8 percent of total southeastern assets. Out-of-state banks did not control banks in any other regional states as of December 1984.

Out-of-state ownership of southeastern banks increased sharply during the next five years, especially in Florida, Georgia, and Tennessee. In contrast, out-of-state bank holding companies moved to gain control of only a modest amount of banking assets in Alabama, Louisiana, and Mississippi. Accordingly, the proportion of statewide banking assets held by out-of-state firms was 35 percent for Florida, 26 percent for Georgia, and 31 percent for Tennessee, and equaled or fell well below

Table 2.
Summary of Out-of-State Owned Banks, Southeastern States

State	1984			1989		
	Number of Institutions	Total Assets[1]	Percent of Total Assets	Number of Institutions	Total Assets[2]	Percent of Total Assets
Alabama	0	0	0	3	105	0.3
Florida	3	4,095	5.2	55	49,122	35.0
Georgia	0	0	0	20	16,858	26.0
Louisiana	0	0	0	2	1,394	3.8
Mississippi	0	0	0	2	523	2.6
Tennessee	0	0	0	23	13,931	30.7
Southeastern Total	3	4,095	1.8	105	81,933	23.8

[1] Commercial bank total assets as of December 31, 1984, in millions of dollars.

[2] Commercial bank total assets as of December 31, 1989, in millions of dollars.

3.8 percent for Alabama, Louisiana, and Mississippi. In December 1989 out-of-state holding companies controlled 24 percent of banking assets throughout the Southeast.

Significant Participants

Interstate merger and acquisition activity of the various large acquirers reflects both the opportunities available to different banking organizations and the organizations' evolving strategies. An ex post review of merger and acquisition activity for the major acquirers may lend some insight into the various strategies employed by the most active organizations.

The most noteworthy bank holding company participants in regional merger and acquisition activity during the 1984-89 period are presented in Table 3. Southeastern holding companies qualify for inclusion if they engaged in at least three interstate acquisitions of banking firms located anywhere in the nation or made at least one purchase of a banking institution holding assets exceeding $1 billion. Bank holding companies headquartered outside the region qualify for inclusion if they acquired at least three southeastern banking organizations or purchased at least one regional depository institution with assets greater than $1 billion.

Acquisition activity presented in Table 3 can be grouped into three broad categories based on the apparent strategy of the acquirers. First, a bank holding company may take over a large banking institution that already has a strong presence in another state's major banking markets. This large target institution may compete in a distant or a nearby state. After making one or more purchases of large banking firms in another state, the acquiring bank holding company may then advance its competitive presence by acquiring smaller banks in desirable markets as opportunities arise. By purchasing a large institution, the acquiring holding company may reap the benefits of becoming a major competitor from the start, realizing available economies of scale or scope, and diversifying its earnings. Earnings diversification tends to reduce the credit riskiness of the acquiring bank's loan portfolio because the target bank's current and potential borrowers usually provide variety to the bank's portfolio through additional sources of income and employment. A drawback of buying large banks is that added demands are placed on bank management in controlling the combined institution's expanded assets. Another disadvantage of this particular interstate strategy is the likelihood that a large target bank's unanticipated or current financial problems may depress the parent organization's earnings. Examples of organizations that appear to be following this competitive strategy include SunTrust in Florida and Tennessee, First Wachovia in Georgia, Deposit Guaranty in Louisiana, and Sovran in Tennessee.

A second strategy is to buy a series of small or mid-sized organizations until a major statewide presence is obtained. This approach may be necessary when a potential acquirer is unable to reach an agreement with potential large targets in a state where it wishes to be a major player. This strategy shares many of the advantages of the first. The disadvantage of this approach is that it takes longer to obtain the benefits of being one of a state's large operators, and any number of factors can intervene to slow the development of a statewide operation. The advantages of this approach relative to an immediate big acquisition are twofold: (1) the demands on the acquirer's management to merge the two cultures are reduced (the typically smaller target can more easily conform to the acquirer's culture) and (2) the probability that the acquirer will be purchasing hidden problems is reduced. An example of this strategy is NCNB's acquisitions in Florida. Although several of NCNB's acquisitions exceeded $1 billion in assets, a sizable combination of transactions was needed to provide a statewide presence in this heavily populated state. After acquiring Atlantic Bancorporation's $3.8 billion in total assets in 1985, First Union appeared to be following in NCNB's footsteps, with numerous smaller acquisitions until 1989. In 1989 First Union dramatically increased its Florida presence by purchasing Florida National Banks, an $8 billion organization.

A third strategy for a bank holding company is to build its interstate presence by purchasing banks, frequently small organizations, in

adjacent banking markets of other states. By doing so the acquirer may gain some economies of scale or scope because of the propinquity of these markets and can tap the market knowledge possessed by the target banks' managers. Moreover, management of banking operations for smaller banks at near-by locations should not pose serious problems for the parent organization. On the other hand, earnings diversification benefits for the acquiring firm are minor because loans held by small banks add minimal variety to the acquiring bank's portfolio. A clear example of this strategy is Synovus Financial Corp's acquisitions in Alabama. Synovus's new Alabama banking markets are either in the Alabama portion of the Columbus, Georgia, banking market or within 100 miles of the company's headquarters in Columbus. Furthermore, Synovus acquired several banking organizations in the nearby panhandle of Florida. Another example is the acquisition by Union Planters Corporation of Memphis, Tennessee, of two Mississippi bank holding companies located close to company headquarters. The two target banks held $500 million in total assets combined and competed in northern Mississippi banking markets situated within seventy-five miles of Memphis.

Conclusion

The relaxation of restrictions on interstate banking in the Southeast led to a marked restructuring of the region's banking industry. Prior to 1985 when the U.S. Supreme Court affirmed the constitutionality of regional banking compacts, banks nationwide were generally constrained to merging with or acquiring home-state institutions. The desire to expand through interstate banking is evidenced by the many southeastern banks purchased and the large volume of assets transferred to the control of banking institutions in other states during the Southeast's first five years of interstate banking.

Table 3.
Southeastern Merger and Acquisition
Activity-Significant Participants[*]

Acquirer	1984	1985	1986	1987	1988	1989
Alabama						
SouthTrust Corporation Birmingham, AL ($7,769)				Central Bank of Volusia County South Daytona, FL 5/22/87 - $60	First National Bancshares, Inc. Jacksonville, FL 1/12/88 - $32	Sentry Bancshares Corporation Roswell, GA 5/9/89 - $57
				First Bancshares, Inc. Marianna, FL 11/13/87 - $66	Melton's Bank Liberty, TN 2/25/88 - $0.3	Florida Central Banks, Inc. Chipley, FL 8/14/89 - $14
				Bank of Florida Corporation St. Petersburg, FL 11/13/87 - $116	Latta Bank & Trust Company Latta, SC 5/27/88 - $13	Florida Community Banks, Inc. Bonifay, FL 8/14/89 - $21
				VistaBanks, Inc. Ormond Beach, FL 11/13/87 - $109		
				Gulf/Bay Financial Corporation Tampa, FL 11/19/87 - $35		
Central Bancshares of the South, Inc. Birmingham, AL ($4,524)					Weslayan Bancshares, Inc. Houston, TX 1/15/88 - $46	City National Bank of Plano Plano, TX 11/9/89 - $84

*The acquiring bank holding company's consolidated total assets, in millions of dollars, are displayed in parentheses. The total asset value presented is as of December 31, 1989. The figure displayed for the acquired (target) banking organization is that firm's total assets in millions of dollars as of December 31 of the year preceding the year in which the merger or acquisition received approval from federal regulators.

Acquirer	1984	1985	1986	1987	1988	1989
Alabama (continued)						
AmSouth Bancorporation Birmingham, AL ($8,572)				First Mutual Bank Pensacola, FL 9/29/87 - $789	First National Bank of Crosby Crosby, TX 11/17/88 - $7 Gulf First Holding Corporation Panama City, FL 2/18/88 - $73 First National Bank of Destin Destin, FL 4/28/88 - $79	
Florida						
Barnett Banks Inc. Jacksonville, FL ($29,007)			First City Bancorp, Inc. Marietta, GA 11/28/86 - $457		First Fulton Bancshares, Inc. Palmetto, GA 1/11/88 - $42 ANB Bankshares, Inc. Brunswick, GA 12/5/88 - $212 First Federal Savings and Loan Association of Columbus Columbus, GA 12/27/88 - $254	Investors Trust Financial Corporation Duluth, GA 6/9/89 - $58

Table 3 (continued)

Acquirer	1984	1985	1986	1987	1988	1989
			Florida (continued)			
					FMB Financial Holdings, Inc. Fayetteville, GA 12/27/88 - $91	
			Georgia			
Citizens and Southern Corporation Atlanta, GA ($23,348)		Landmark Banking Corporation Fort Lauderdale, FL 7/29/85 - $3,661	Citizens and Southern Corporation Columbia, SC 2/4/86 - $2,843	Southern Bank Corporation Tallahassee, FL 10/5/87 - $53		The Ocean State Bank Neptune Beach, FL 10/23/89 - $124
			The Farmers & Merchants Bank Walterboro, SC 9/25/86 - $73			Liberty Federal Savings and Loan Association Port Richey, FL 8/11/89 - $60
			Andrews Bank & Trust Company Andrews, SC 10/31/86 - $87			
			Bank of the Islands Sanibel, FL 11/4/86 - $102			
			Community National Bank Kissimmee, FL 11/4/86 - $14			
			First National Bank, Seminole County Longwood, FL 11/4/86 - $14			

Acquirer	1984	1985	1986	1987	1988	1989
			Georgia (continued)			
			First National Bank, Winter Park Winter Park, FL 11/4/86 - $211			
SunTrust Banks, Inc. Atlanta, GA ($31,044)		Trust Company of Georgia & Sun-Banks, Inc. Atlanta, GA & Orlando, FL 1/8/85 - $15,523 (Holding Company Formation) Hernando Banking Corporation Brooksville, FL 3/28/85 - $239 Pan American Bank of Sarasota Sarasota, FL 7/30/85 - $105	Third National Corporation Nashville, TN 11/26/86 - $5,062	Peoples Banc-shares Inc. Lebanon, TN 2/10/87 - $68 SWG Financial Enterprises Morristown, TN 3/19/87 - $123 London County Bankshares, Inc. Lenoir City, TN 4/14/87 - $70	Commercial Bank in Panama City Panama City, FL 4/12/88 - $238	
Synovus Financial Corp Columbus, GA ($2,410)					Northwest Florida Banking Corporation Quincy, FL 1/11/88 - $81 Fort Rucker National Bank Fort Rucker, AL 7/20/88 - $40	Farmers and Merchants Bank of Russell County Phenix City, AL 3/15/89 - $39 Vanguard Banks, Inc. Valparaiso, FL 8/8/89 - $143 Bank of Pensacola Pensacola, FL 11/9/89 - $55

Table 3 (continued)

Acquirer	1984	1985	1986	1987	1988	1989
			Louisiana			
Hibernia Corporation New Orleans, LA ($6,697)						First State Bank Pflugerville, TX 8/24/89 - $34
						Thousand Oaks National Bank San Antonio, TX 9/7/89 - $36
						Humble Savings and Loan Associa- tion Humble, TX 9/15/89 - $52
						Trinity Valley Sav- ings and Loan Association Cleveland, TX 10/6/89 - $92
						United National Bank of Plano Plano, TX 11/9/89 - $44
						Executive National Bank San Antonio, TX 11/16/89 - $11
						Love Field National Bank Dallas, TX 11/16/89 - $44

Acquirer	1984	1985	1986	1987	1988	1989
Louisiana (continued)						
						Greater Texas Bank North, N.A. Austin, TX 11/30/89 - $28
						First National Bank in Frisco Frisco, TX 12/7/89 - $11
						Westheimer Memorial Bank, N.A. Houston, TX 12/8/89 - $48
Mississippi						
Deposit Guaranty Corporation Jackson, MS ($3,745)						Commercial National Corporation Shreveport, LA 11/27/89 - $1,041
Tennessee						
Union Planters Corporation Memphis, TN ($4,003)						United Southern Corporation Clarksdale, MS 1/13/89 - $346
						National Commerce Corporation New Albany, MS 9/22/89 - $151

Table 3 (continued)

Acquirer	1984	1985	1986	1987	1988	1989
Tennessee (continued)						
						Steiner Bank Birmingham, AL 10/6/89 - $18
District of Columbia						
First American Bankshares, Inc. Washington, D.C. ($11,528)				NBG Financial Corporation Atlanta, GA 6/26/87 - $1,600		Bank of Escambia, N.A. Pensacola, FL 2/16/89 - $36
New York[1]						
Citicorp New York, NY ($230,643)	Biscayne Federal Savings and Loan Association Miami, FL 1/20/84 - $1,945				Caribank Dania, FL 12/9/88 - $629	
North Carolina						
NCNB Corporation Charlotte, NC ($66,494)	Ellis Banking Corporation Bradenton, FL 2/15/84 - $1,695	Pan American Banks Miami, FL 11/27/85 - $1,563	National Bank of Florida Miami, FL 2/14/86 - $93	The County Bank Palmetto, FL 2/13/87 - $60	USBancorp Inc. St. Petersburg, FL 5/31/88 - $49	Southern Florida-Banc Federal Savings Bank Boca Raton, FL 10/6/89 - $143

[1] *New York is not included in Florida's (or other southeastern states') regional, reciprocal interstate banking laws. The federal banking laws that allow nationwide acquisitions of commercial banks and thrifts, usually troubled institutions, are the Garn-St Germain Act of 1982, Competitive Equality in Banking Act of 1987, and Financial Institutions Reform, Recovery, and Enforcement Act of 1989.*

Acquirer	1984	1985	1986	1987	1988	1989
North Carolina (continued)						
		Southern National Bancshares Atlanta, GA 11/27/85 - $79				Freedom Federal Savings and Loan Association Tampa, FL 10/13/89 - $1,402
First Union Corporation Charlotte, NC ($32,131)		Atlantic Bancorporation Jacksonville, FL 10/16/85 - $3,752 Central Florida Bank Corporation Dade City, FL 10/24/85 - $178	Citizens DeKalb Bank Clarkston, GA 2/14/86 - $42 First Bankers Corporation of Florida Pompano Beach, FL 4/17/86 - $1,292 Georgia State Bancshares, Inc. Atlanta, GA 8/8/86 - $153 Bank of Waynesboro Waynesboro, GA 9/12/86 - $43 First Railroad and Banking Company of Georgia Augusta, GA 9/22/86 - $3,377 Collier Bank Naples, FL 11/12/86 - $38	Roswell Bank Roswell, GA 2/27/87 - $143 Commerce National Bank Naples, FL 4/6/87 - $43 First North Port Bancorp North Port, FL 4/10/87 - $71 First Sarasota Bancorporation Sarasota, FL 4/22/87 - $51 Sarasota Bank and Trust Company Sarasota, FL 5/19/87 - $24 First State Bank of Pensacola Pensacola, FL 9/1/87 - $111	Florida Commercial Banks, Inc. Miami, FL 1/7/88 - $984	Florida National Banks of Florida, Inc. Jacksonville, FL 12/22/89 - $7,828

Table 3 (continued)

Acquirer	1984	1985	1986	1987	1988	1989
			North Carolina (continued)			
First Wachovia Corporation Winston-Salem, NC ($24,041)		Wachovia Corporation & First Atlanta Corporation Winston-Salem, NC & Atlanta, GA 11/4/85 - $14,468 (Holding Company Formation)	Security National Bank Fort Meyers, FL 11/12/86 - $51 Forsyth County Bank Cumming, GA 12/15/86 - $92	Community Bank of Manatee Bradenton, FL 10/28/87 - $60 Bank of Bellevue Nashville, TN 11/20/87 - $9 FA Bankshares, Inc. Monroe, GA 4/23/87 - $75 North Georgia Bankshares, Inc. Canton, GA 12/23/87 - $47		First Bank and Trust Company Fayetteville, GA 6/6/89 - $35
			Virginia			
Dominion Bankshares Corporation Roanoke, VA ($10,119)			Nashville City Bank & Trust Company Nashville, TN 10/20/86 - $589	First Dickson Corporation Dickson, TN 1/2/87 - $96 First National Financial Corporation Clarksville, TN 4/1/87 - $209 Ashland City Bank and Trust Company Ashland City, TN 6/17/87 - $53	Merchants & Planters Corporation Newport, TN 2/18/88 - $120 Citizens Union Corporation Rogersville, TN 2/22/88 - $170	

Acquirer	1984	1985	1986	1987	1988	1989
			Virginia (continued)			
				First Springfield National Corporation Springfield, TN 8/20/87 - $76		
				First National Bank of Sparta Sparta, TN 11/16/87 - $94		
				Franklin First National Corporation Decherd, TN 11/27/87 - $55		
				The Peoples National Bancorp, Inc. Shelbyville, TN 11/27/87 - $90		
				UNB Corporation Fayetteville, TN 11/27/87 - $97		
Sovran Financial Corporation Norfolk, VA ($25,442)				Commerce Union Corporation Nashville, TN 10/1/87 - $3,930[1]	First Bank of Marion County South Pittsburgh, TN 12/19/88 - $97	
					First National Bank of Collierville Collierville, TN 12/19/88 - $47	

[1] *This transaction provided Sovran Financial with entry into the state of Kentucky. Commerce Union controlled eleven subsidiary banks when it was acquired in October 1987. Ten subsidiaries were Tennessee-based, and one—Planters Bank and Trust Company of Hopkinsville, with total assets of $182 million—was Kentucky-based.*

Interstate Banking Legislation by State: Southeast Compact

State	Description (The years in parentheses are the dates the laws were passed.)
Alabama	Reciprocal, twelve states and DC (AR, FL, GA, KY, LA, MD, MS, NC, SC, TN, VA, WV). The in-state institution to be acquired must have been in existence at least five years. The law became effective July 1, 1987 (1986).
Arkansas	Reciprocal, sixteen states and DC (AL, FL, GA, KS, LA, MD, MS, MO, NE, NC, OK, SC, TN, TX, VA, WV). The effective date of the law is January 1, 1989 (1988).
District of Columbia	Reciprocal, eleven states (AL, FL, GA, LA, MD, MS, NC, SC, TN, VA, WV) (1985). Permits bank holding companies outside the preceding listed states to acquire existing DC banking organizations, provided certain commitments are made regarding investments and operations in DC (1986).
Florida	Reciprocal, eleven states and DC (AL, AR, GA, LA, MD, MS, NC, SC, TN, VA, WV). The in-state bank to be acquired must have been in existence at least two years (1984). Under a 1972 law, NCNB and Northern Trust Corporation are grandfathered and can make further in-state acquisitions.
Georgia	Reciprocal, ten states and DC (AL, FL, KY, LA, MD, MS, NC, SC, TN, VA). The in-state bank to be acquired must have been in existence at least five years (1984). MD and DC were added to the region effective March 13, 1987 (1987).
Kentucky	National, reciprocal. The in-state bank to be acquired must have been in existence at least five years. Initially, there was a contiguous state requirement, but that was dropped July 15, 1986 (1984).
Louisiana	National, reciprocal. The statute originally permitted bank holding company entry on a reciprocal basis, but only by holding companies located in AL, AR, FL, GA, KY, MD, MS, NC, OK, SC, TN, TX, VA, WV, and DC. The law became effective July 1, 1987 (1986). Regional restriction was dropped on January 1, 1989. Parallel legislation covers S&Ls.
Maryland	Reciprocal, fourteen states and DC (AL, AR, DE, FL, GA, KY, LA, MS, NC, PA, SC, TN, VA, WV) (1985). For DE, VA, WV, and DC, the law became effective July 1, 1985. The law became effective July 1, 1987, for the remaining states.
Mississippi	Reciprocal, thirteen states (AL, AR, FL, GA, KY, LA, MO, NC, SC, TN, TX, VA, WV). The law became effective for the contiguous states of AL, AR, LA, and TN on July 1, 1988 (1986). The same statute specified that the region would be expanded to include the remaining southeastern states effective July 1, 1990.
North Carolina	Reciprocal, thirteen states and DC (AL, AR, FL, GA, KY, LA, MD, MS, SC, TN, TX, VA, WV) (1984).
South Carolina	Reciprocal, twelve states and DC (AL, AR, FL, GA, KY, LA, MD, MS, NC, TN, VA, WV). The law became effective January 1, 1986. The in-state bank to be acquired must have been in existence at least five years (1984).
Tennessee	Reciprocal, fourteen states (AL, AR, FL, GA, IN, KY, LA, MD, MS, MO, NC, SC, VA, WV). The banking organization to be acquired must have been in existence at least five years (1985). National, reciprocal effective January 1, 1991.
Virginia	Reciprocal, twelve states and DC (AL, AR, FL, GA, KY, LA, MD, MS, NC, SC, TN, WV). The in-state bank to be acquired must have been in existence at least two years (1985).
West Virginia	National, reciprocal. The law became effective January 1, 1988, and applies to both bank holding companies and savings and loan holding companies. The in-state institution to be acquired must have been in existence for two years (1986).

Sources: *Banking Expansion Reporter* (1990, 7-10); King, Tschinkel, and Whitehead (1989, 35-36).

Notes

[1] Maine enacted an interstate banking law in 1975, but the law had a reciprocity restriction. Out-of-state banks could not acquire Maine banks unless Maine banks could acquire banks located in the acquirer's headquarters state. The reciprocity restriction effectively prevented any interstate acquisition of Maine banks until another interstate banking law was adopted. Such a law was adopted by Massachusetts legislators in 1982. The laws governing interstate banking in the Southeast are discussed in detail in the box on page 52.

[2] Commercial banking transactions in the Southeast are the focus of this article. The study includes transactions in which the acquiring organization or principal merger partner is a commercial bank holding company, bank holding company subsidiary, or independent bank. Commercial bank consolidations with thrift institutions are included as well. Transactions involving consolidations solely among thrift organizations are excluded. Also excluded are transactions involving commercial bank purchases of branch offices, either bank or thrift branches, and consolidations involving trust companies. The Southeast as defined here includes Alabama, Florida, Georgia, Louisiana, Mississippi, and Tennessee.

[3] The State Superintendent of Banking authorized statewide branching privileges to Alabama-chartered banks in a policy letter issued on May 31, 1990. The superintendent cited a May 22, 1990, decision by the Comptroller of the Currency to approve the application by The First National Bank of Florence, Lauderdale County, Alabama (FNB Florence), to establish a branch bank in Muscle Shoals, Colbert County, Alabama. Commercial banks in Alabama were restricted previously to branching within their home counties by local laws and general laws of local application, and the limitation applied to Lauderdale and Colbert counties. The superintendent's decision for this Alabama case was based on the fact that national banks and state-chartered savings and loan associations in Alabama have the authority to establish branch offices throughout the state. That is, state-chartered savings and loan associations in Alabama are authorized by Alabama law to do whatever federal savings and loan associations domiciled in the state may do. If a federal savings and loan association in Alabama may branch statewide, an Alabama-chartered savings and loan association may do likewise. The Comptroller of the Currency had previously decided that national banks could branch statewide because federal savings and loan associations, which held essentially the same powers of a national banks, could do so.

Specifically, under Alabama's Equal Powers Act, Section 5-5A-18.1, Code of Alabama 1975, a state-chartered bank in Alabama is authorized to exercise any powers that a national bank domiciled in Alabama may exercise. This includes the power to establish branch offices. The Alabama Superintendent of Banks, therefore, advised state-chartered banks that they may branch statewide because national banks, as well as state-chartered savings and loan associations, are permitted to do so. Previous decisions by the Comptroller of the Currency involved national banks' branch applications in Mississippi, Tennessee, Texas, Louisiana, Florida, Missouri, Wisconsin, Kansas, and Indiana.

Florida banking statutes allowed statewide branching privileges for its commercial banks in 1989. Louisiana banking law permitted statewide branching for state-chartered banks in June 1988. Lawmakers provided Mississippi-chartered banks statewide branching rights on July 1, 1989, subject to minimum population requirements for municipalities. In 1990 Tennessee banking law allowed state-chartered banks to establish branch offices at any location in any county of the state.

[4] During most of the six years under study, branching laws as specified below applied to each of the states in the Southeast. (See note 3 for an explanation of statewide branching rights granted to southeastern states except Georgia during 1988, 1989, or 1990.)

Alabama. Prior to 1990 state-chartered commercial banks could establish branch offices within the parent bank's home county. Local laws and general laws of local application applied because state banking statutes did not explicitly address branching rights.

Florida. Before 1989 a Florida-chartered bank could establish branches within the borders of the parent bank's home county.

Georgia. Commercial banks are still restricted to branching within county borders. An exception has been made for banks in densely populated Fulton and DeKalb counties. State-chartered commercial banks in one of these contiguous counties may establish branches in the other county.

Louisiana. Before 1988 state-chartered banks (with capital of $100,000 or more) could open branch offices in parishes in which there were no state-chartered banks. Also, not more than one branch (in addition to branches in the parent bank's home parish) could be opened in any one parish other than the home parish. All banks in the parishes of Allen, Calcasieu, Cameron, or Jefferson Davis could establish branch offices in any one or more of the other named parishes.

Mississippi. Commercial banks could establish branch banks within the following applicable radius from the parent bank: (a) 100 miles from July 1, 1986, through June 30, 1987; (b) 150 miles from July 1, 1987, through June 30, 1988; (c) 200 miles from July 1, 1988, through June 30, 1989; and (d) the geographical boundaries of the State of Mississippi on and after July 1, 1989.

Tennessee. Before 1990 a Tennessee-chartered commercial bank could not establish branches other than in the county where its principal office was located.

All six states permitted their state-chartered banks to establish new branches in the case of emergency or failure of existing banks. See Amel and Keane (1986).

[5] For a more detailed discussion of legislation pertaining to interstate banking see King, Tschinkel, and Whitehead (1989).

[6] These grandfather provisions of the Douglas Amendment account for a few large regional organizations that operated from 1956 to 1985, when the nation's Supreme Court ruled that regional interstate banking compacts did not violate antidiscrimination provisions of the Con-

stitution. Examples of such cross-state organizations are Interstate Bancorp in the West and First Bank System and Norwest in the upper Midwest.

[7]Purchased funds consist of certificates of deposits with balances of roughly $100,000 or more. The legal maximum per named account covered by the Federal Deposit Insurance Corporation is $100,000. These deposits tend to be withdrawn quickly when a banking organization's financial difficulties, perceived or actual, become known to the public. Purchased funds may also be withdrawn when interest rates paid on deposits by the offering bank drop significantly relative to the rates paid by competitors. Core deposits, on the other hand, are nonbrokered FDIC-insured accounts with balances under $100,000.

[8]For a fuller discussion of these motivations for bank acquisitions and mergers, see Hunter and Wall (1989).

[9]Three troubled depository institutions in Florida were acquired by New York bank holding companies under provisions of the Garn-St Germain Act of 1982 or the Competitive Equality in Banking Act of 1987. Citicorp acquired a failed thrift and a troubled commercial bank in 1984 and 1988, respectively, and Chase Manhattan Corporation purchased a troubled commercial bank in 1986.

First American Bankshares, Inc., a District of Columbia holding company, received approval to acquire Bank of Escambia, N.A., Pensacola, Florida, in February 1989. This acquisition was not mentioned in the text because Bank of Escambia possessed only

$35.5 million in total assets and was acquired through a debt previously contracted situation.

[10]Two interstate transactions were publicly announced as mergers of equals, with individual and presumably equal headquarters for the new organization's two subsidiary banks. In 1985 SunTrust Banks, Inc., was formed as a bank holding company with Trust Company of Georgia, Atlanta, Georgia, and SunBanks, Inc., of Orlando, Florida, as bank subsidiaries. Also in 1985 First Wachovia Corporation was formed as a bank holding company with two subsidiary banks. They were Wachovia Corporation of Winston-Salem, North Carolina, and First Atlanta Corporation of Atlanta, Georgia. For both formations, total assets held by the two subsidiary banks were roughly equal, but over time SunTrust's Georgia operations and First Wachovia's North Carolina operations have become dominant. This study, therefore, views the SunTrust and First Wachovia transactions as acquisitions of Florida and Georgia banks by Georgia and North Carolina bank holding companies, respectively.

[11]A merger between Citizens and Southern Corporation in Georgia and Sovran Financial Corporation in Virginia was approved in 1990, after the end of our sample period. This merger was billed as a merger of equals. Although one organization almost always dominates transactions billed as mergers of equals, it is not yet possible to discern which organization will prevail or whether this is in fact a merger where both organizations will be equal.

References

Amel, Dean F., and Daniel G. Keane. "State Laws Affecting Commercial Bank Branching, Multibank Holding Company Expansion, and Interstate Banking." *Issues in Bank Regulation* (Autumn 1986): 30-40.

Banking Expansion Reporter 9, no. 16, August 20, 1990.

Code of Alabama, 1975. Vol. 4, title 5. 1987-1990 Cumulative Supplement. Charlottesville, Va.: The Michie Company, 1990.

Florida Statutes Annotated. Vol. 19, section 658. 1989 Revision, 1989 Cumulative Pocket Part. Norcross, Ga.: The Harrison Company, 1989.

Goudreau, Robert E., and B. Frank King. "Recovering Bank Profitability: Spoiled Again by Large Banks' Loan Problems." Federal Reserve Bank of Atlanta *Economic Review* 75 (May/June 1990): 30-43.

Hunter, William C., and Larry D. Wall. "Bank Merger Motivations: A Review of the Evidence and an Examination of Key Target Bank Characteristics." Federal Reserve Bank of Atlanta *Economic Review* 74 (September/October 1989): 2-19.

King, B. Frank, Sheila L. Tschinkel, and David D. Whitehead. "Interstate Banking Developments in the 1980s." Federal Reserve Bank of Atlanta *Economic Review* 74 (May/June 1989): 32-51.

Louisiana Revised Statutes. Vol. 2, title 6. 1986-1990 Cumulative Annual Pocket Part. St. Paul, Minn.: West Publishing Company, 1990.

Mississippi Code 1972 Annotated. Vol. 18, title 81. Revised 1989-1990 Cumulative Supplement. Rochester, N.Y.: The Lawyers Cooperative Publishing Company, 1990.

Official Code of Georgia, Annotated. Vol. 5, title 7. 1989-1990 Supplement. Charlottesville, Va.: The Michie Company, 1990.

Tennessee Code Annotated. Vol. 8, title 45. 1987 Replacement, 1987-1990 Supplement. Charlottesville, Va.: The Michie Company, 1990.

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.

90-1   Thomas J. Cunningham
Some Real Evidence on the Real Bills Doctrine Versus the Quantity Theory

90-2   Joseph A. Whitt, Jr.
The Long-Run Behavior of the Real Exchange Rate: A Reconsideration

90-3   Jon R. Moen and Ellis W. Tallman
The Bank Panic of 1907: The Role of Trust Companies

90-4   Lucy F. Ackert and William C. Hunter
A Sequential Test Methodology for Detecting Futures Market Disruptions with Applications to Futures Margin Management

90-5   Peter A. Abken
Valuing Default-Risky Interest-Rate Caps: A Monte Carlo Approach

90-6   Peter A. Zadrozny
Estimating a Multivariate ARMA Model with Mixed-Frequency Data: An Application to Forecasting U.S. GNP at Monthly Intervals

90-7   Charles H. Whiteman and William Roberds
Monetary Aggregates as Monetary Targets: A Statistical Investigation

90-8   William C. Hunter and Stephen G. Timme
Quasi-Fixed Inputs in the Estimation of Bank Scale Economies

90-9   Ellis W. Tallman and Ping Wang
Human Capital and Endogenous Growth: Evidence from Taiwan

90-10   Thomas J. Cunningham and Rosemary Thomas Cunningham
The Behavior of Real Rates of Interest in a Small, Opening Economy

90-11   George J. Benston, William C. Hunter, and Larry D. Wall
Potential Diversification and Bank Acquisition Prices

90-12   Vikram Kumar
Effects of Devaluation with a Partial Wage Floor

Previous working papers are also available. For more information contact the Public Information Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., 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:

❑90-1   ❑90-2   ❑90-3   ❑90-4   ❑90-5   ❑90-6   ❑90-7   ❑90-8   ❑90-9   ❑90-10   ❑90-11   ❑90-12

Name_____

Address_____

City _____ State _____ ZIP_____

Return to the Public Information Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., Atlanta, Georgia 30303-2713.

Book Review

The Energy Crisis and the American Political Economy: Politics and Markets in the Management of Natural Resources

by Franklin Tugwell.
Stanford, Calif.: Stanford University Press, 1988.
294 pages. $32.50.

Even before Iraq's invasion of Kuwait on August 2, 1990, most Americans could easily recall the trauma of the energy crisis during the 1970s and early 1980s. Recent events have no doubt reminded many people of the skyrocketing oil prices and long lines at gasoline pumps that accompanied the scarcity of oil supplies during that period. However, the events that led to the 1970s' crisis and the circumstances surrounding it may not be remembered in detail or clearly understood even by those it most seriously affected.

In *The Energy Crisis and the American Political Economy*, Franklin Tugwell has provided an excellent review of the events and circumstances leading to the energy crisis that began in the winter of 1973-74, as well as the nation's reactions to it, especially the government's largely unsuccessful attempts to alleviate its impacts. The author, who is Professor of Government at Pomona College and Claremont Graduate School, also draws some worrisome conclusions about responses if American oil supplies should again be restricted. In short, he believes that the country's dealings with the first crisis left it little better prepared to deal with a second one.

In part one of his four-part book, Tugwell traces the evolution of both public and private efforts to deal with competition within the major energy regimes: coal, oil, natural gas, electricity, and nuclear power. The second section contains a chronology of the development of the energy crisis and the country's response to it. In the third part Tugwell examines changes in energy's political economy during the crisis and public policy's role in shaping those developments. Finally, the author looks at current shortcomings associated with the U.S. system of economic governance—specifically, the failure to achieve a lasting solution to the energy crisis—and suggests other arrangements he believes would prove more effective in dealing with such problems.

Until the 1970s' crisis the major energy forms used in the United States bore little relationship to each other in either private ownership or public policy. The author looks in detail at the development of the primary energy regimes and the ways in which each one has responded to its particular production and marketing problems. These responses have involved both private arrangements and

public intervention designed to moderate market forces. Tugwell presents evidence to show that each fuel type developed its own political subsystem in which industry representatives teamed up with specialized congressional and bureaucratic groups to shape public policies. Given this assemblage of distinct policy-making groups, often labeled "iron triangles" because of the rigidly closed system within which policies were concocted, it is not surprising that, prior to the oil embargo, there was no attempt to develop a coherent public policy for the energy system as a whole.

Government intervention in the coal industry never worked well for long periods because of the geographical dispersion and independence of producers, who were unable to agree among themselves and with mine workers on the objectives of coal policy. As a result, supply could not be controlled sufficiently to manage markets.

By contrast, the oil industry was amenable to government regulations partly because of its early monopolistic control, which allowed industry leaders to exercise great influence in defining problems and designing regulatory solutions. Supply control, with the objective of supporting prices, was always a clearly understood goal of oil industry leaders.

Granted monopoly status because of its cumbersome, costly delivery system, the natural gas industry experienced strong consumer representation in policy-making from the outset. Consumers' predominant influence forged regulations that kept prices low and stymied investment and development, which give low returns to producers.

The electricity industry was set up early on as a system of legal monopolies in local service districts. Producers' relationships with regulatory authorities were cooperative rather than conflictual, and authorities adopted a strong promotional attitude toward electric power. This promotion encouraged growth that provided increasing scale economies, benefiting investors, managers, and consumers alike. Although initially state government regulation dominated electrical supply, federal government intervention expanded when supply networks reached beyond state boundaries, particularly when atomic energy

began to be adopted as a power source. However, the structure of the electricity regime made it unusually vulnerable to the energy crisis in 1973, when the relative tranquility of proprietary regulation was upset by emerging adversarial relationships with environmentalists, antinuclear activists, and disgruntled consumers.

By the early 1970s each energy regime had developed a somewhat stable set of procedures for governing its particular piece of the industry. When the energy crisis struck, each energy type was, for different reasons, vulnerable to the severe market adjustments caused by rapidly rising oil prices. Consequently, the nation's response to the crisis was heavily influenced by the lack of uniformity among regulatory procedures for the various fuel types.

As Tugwell reminds readers, the oil embargo engineered by the Arab members of the Organization of Petroleum Exporting Countries (OPEC) followed the outbreak of the Yom Kippur War in 1973. The oil ministers of the Arab oil-exporting countries chose to draw attention to the Arab cause by imposing production cutbacks in late 1973. Within the next few months, the supply of oil reaching international markets fell by 15 percent and the price of world oil jumped from $3 to $12 per barrel. The impact on the world economy was estimated to be equivalent to a $100 billion tax, claimed by some to be the largest reallocation of income in history.

The embargo struck a particularly telling blow to the U.S. economy, Tugwell contends, partly because it occurred in the midst of other problems. During the years before the embargo, the United States had suffered from shortages of natural gas and some petroleum products because of price ceilings that limited producers' incentive to supply those products. Escalating inflation accompanying the government's large expenditures for the Vietnam War and the Great Society's domestic programs had led President Richard Nixon to institute an economywide wage and price freeze in 1971, two years before the oil embargo. Tugwell claims that the freeze aggravated the shortages, especially in energy and petroleum products, but that the oil embargo and the public's resentment of its effects

made it politically impossible to remove energy price controls even when general price controls were scrapped. In fact, he argues, the Nixon administration's price control programs set the stage for the subsequent elaborate regulatory system that involved the government ever more deeply in energy matters.

Although the government had intervened in varying degrees in each energy regime prior to the Nixon era, coal and oil were distributed largely under a market system until price controls were imposed in 1971. Tugwell contends that the measures adopted to manage shortages—including price controls, supply allocations, entitlements, and vast expansions in the government's regulatory functions—created a broad set of stakeholders who, feeling threatened by a return to the market allocation system, became additional political forces to be reckoned with in attempting to formulate energy policy.

All of these factors contributed to the set of conditions that Tugwell labels the energy crisis. He enumerates several major elements of the crisis: the abrupt restrictions of imported oil supplies, on which the nation had become dependent; rapidly dwindling domestic oil reserves, reflecting a century of growing usage; a lack of readily available alternative energy sources; the vast transfers of wealth from consumers to foreign oil producers because of the rising cost of imported oil; and the United States' virtual helplessness in thwarting embargo organizers' aims.

Although several energy policies emerged from the turmoil, Tugwell argues that these typically moved further and further from the actions needed to alleviate the energy crisis, basically because the political process was deadlocked by multiple special interest groups. Moreover, from the mid-1970s onward this stalemate has tended to stymie most political efforts to devise energy policy. President Carter promoted a host of new policies designed to reduce imports, increase energy use efficiency, and distribute more equitably the burden of high energy prices. Though Tugwell endorses these as goals, he argues that Carter essentially could not overcome the stalemate and in the end merely accelerated the process of deregulation.

Tugwell attributes to the Reagan administration the final dismantling of petroleum price and allocation controls. Despite that administration's vigorous efforts to abolish the Department of Energy and move to freer energy markets, limited headway was made in further reducing the government's role in the energy regimes, Tugwell reports, because legislators proved reluctant to abandon programs so recently brought into existence through hard-won compromises. Such measures as environmental regulations impinging on energy development and programs to help poor people pay for increased costs of energy continued.

The author acknowledges a number of accomplishments during America's decade-long struggle with the energy crisis: the improved

"Although several energy policies emerged from the . . . [1970s' crisis], Tugwell argues that these typically moved further and further from the actions needed to alleviate the energy crisis, basically because the political process was deadlocked by multiple special interest groups."

fuel efficiency of autos and trucks, reductions in home heating fuel use resulting from numerous conservation measures, reduced energy intensity in the industrial sector, and considerable substitution of other energy sources for oil. He points out, however, that many of these achievements occurred in spite of the energy policies enacted, which often intended different results. For example, one major energy policy objective dating from the early days of the second Nixon administration in 1974 was to reduce the country's dependence on imported oil and restore energy self-sufficiency. The Emergency Petroleum Allocation Act, which Nixon signed into law in late 1973, was intended to protect consumers from rising oil prices. Price controls partially succeeded in holding crude oil prices down but at the same time discouraged domestic production, leading to an increase in oil im-

ports relative to what they otherwise would have been. In fact, from 1973 to 1977 oil imports had risen 31 percent, and the same Arab countries that had embargoed the United States were providing a large share of the total.

Largely because of sharp price increases in 1973 and again in 1979-80, energy use patterns changed and efficiency increased, eventually cutting into the growth in energy consumption. Had prices risen faster, domestic production might have increased more quickly and by a greater amount; consumption undoubtedly would have grown less rapidly. These events would have hastened the restoration of the supply-demand balance that eventually brought prices under control.

In Tugwell's opinion, one of the most obvious impacts of national energy policy was a

"According to Tugwell, a major reason for the relative ineffectiveness of the energy regulatory structure was the political rift between energy producers and energy consumers."

quite unintended transfer of wealth across national borders as well as from consuming to producing areas within the United States. As U.S. oil imports increased, Tugwell states, OPEC countries received an additional estimated $1 billion to $6 billion annually between 1975 and 1979 from U.S. consumers. He points out that low-income people were particularly hard hit by these price shocks, as industries laid off low-wage earners or cut their pay in an effort to restore industry competitiveness and pave the way for economic recovery. In addition, in an attempt to deal with the energy crisis, the federal government established a large regulatory apparatus that imposed a costly administrative burden on the private sector. Regulations intended to protect consumers from paying high prices for oil and to prevent owners of domestic petroleum resources from reaping unearned returns from

the price jump extended in part from 1972 through 1981. The author argues that the accomplishments of these regulations were mixed at best.

According to Tugwell, a major reason for the relative ineffectiveness of the energy regulatory structure was the political rift between energy producers and energy consumers. This division was compounded by the disagreement between the executive and legislative branches of government concerning the procedures and policies needed. In effect, each group was powerful enough to thwart the other's initiatives. It is perhaps because energy matters affect the great cross section of the country in such an immediate and conspicuous manner that energy constituent groups became so active and so entrenched in their positions that compromise ultimately proved impossible. An impasse developed on the key issues of petroleum and natural gas price regulation. Despite the Reagan administration's strong efforts to rid energy of government controls, the stalemate was eventually resolved only by the passage of time, which brought reduced consumption of petroleum, increased supplies, and the de facto emergence of world prices in U.S. energy markets. Tugwell states: "The decision to terminate regulation was the result neither of 'principled' choices by decision makers nor of political consensus, but of a persistent stalemate in the body politic that made this the only possible outcome. It was, if you will, deregulation by default."

Looking ahead, the author contends that another energy crisis is almost a foregone conclusion. He sees little hope that the U.S. political system will do more than repeat "the cycle of failure" that characterized the nation's handling of the 1973 energy crisis. As before, he expects that the same powerful interest groups will become intent on protecting their constituencies' interests at the expense of the common good.

Tugwell's bleak outlook for the ability of the American democratic process to handle crises leads him to examine alternative systems for the management of economic affairs. He asserts that a government that is weak, divided, and easily influenced by special interest groups, as America's increasingly appears

to be, can do little to remedy the inefficient distortions of economic activity and waste of resources that result from wild market fluctuations. Such a government, he states, can intervene in markets only on an ad hoc, temporary basis, following no consistent doctrine or objective except to provide immediate compensation to the most influential injured claimants. This situation, Tugwell contends, seriously threatens continued economic progress in an advanced industrial society.

The alternative he favors is "a system of planned markets," which he describes as "an arrangement in which the government accepts responsibility for the operation of the economic system and the values it promotes, but wherever possible eschews detailed regulation or micromanagement, instead relying on competitive, or market-like, processes. It does so by consciously and continuously setting the terms for competition . . . and enforcing them, while at the same time carefully attending to the needs of those who are hurt as a result." This proposed system is needed, Tugwell asserts, because the founders of America's democratic system failed to foresee that corporate organization would become dominant in the political process and that the government would be swamped by multiple demands for access to the policy-making process and for security and at the same time be held responsible for managing the details of economic activity.

In the author's view, this planned market system would enable the government to make decisions and direct resource allocations in a manner that would generate the maximum benefit to society as a whole while avoiding the paralysis that tends to occur in policy-making when some interest groups are powerful enough to block actions that benefit other groups. In the case of the 1970s' energy crisis, such a system presumably would have allowed fuel prices to rise; then taxes or assessments on producers would have been levied to remove windfall profits and those revenues used to alleviate the unusual hardships low-income consumers suffered as a result of the higher prices for gasoline, heating oil, and other necessary energy-based items. Tugwell gives short shrift to the increased hardships that unhampered price increases

would have imposed on limited-resource consumers and businesses, who would have had to pay even higher prices in the short run and suffer the longer-term consequences of a more severe economic contraction.

The author ignores the argument that the policy of holding down domestic oil prices, even though only partially successful, produced some major stabilizing benefits in the short term. He also fails to address the destabilizing influences on industry production and investment that typically result from wild price swings, such as those that occurred for oil during the 1970s and 1980s. The rush to expand production that accompanied the jump in domestic crude oil prices from $9 to $32 per barrel from 1978 to 1981 sharply escalated costs of drilling equipment and resources as producers, many inexperienced, scrambled to take advantage of the boom. The drop in crude oil's price to $12.50 per barrel in 1986, when expanded oil supplies met declining demand, caused massive financial losses from which oil-producing regions have yet to recover. If policy had kept a firmer lid on oil prices, some financial losses from such high-cost expansion might have been avoided.

Tugwell fails to bolster his case by citing examples where planned markets are working or have worked to accomplish the objectives he claims have not been obtained through the U.S. political process. He also does not make it clear what mechanism would ensure that a planning group would, in fact, achieve a planned market system.

There are reasons to believe that Iraq's invasion of Kuwait will not bring about the replay of the mistakes of the 1970s that Tugwell forecasts. The experience of having survived the 1973-74 embargo appears to be alleviating the public panic that was widespread during the previous crisis, when energy supplies were severely restricted and sometimes could not be obtained. This time there is a clear villain manifest in the country of Iraq and its leader Saddam Hussein. During the 1970s' crisis it was commonly suspected that the major oil companies were in league with the OPEC cartel, precipitating the crisis in order to increase their profits. That mind-set produced immense public resentment of fuel price increases and generated measures intended to

prohibit domestic fuel costs from rising. So far in the Iraq-Kuwait conflict, the public has neither demanded policies to restrict price increases nor asked for special supply management programs, even though domestic energy producers and suppliers stand to reap one-time windfall gains.

In addition to a changed public attitude, other factors may cause the 1990s' energy policy to differ from that resulting from the 1973-74 experience. A large government-held strategic oil reserve, not available in the 1970s, would provide some cushion should oil imports be sharply curtailed. With that reserve, oil consumers now have additional time to make adjustments. Moreover, prices had already been rising before Iraq's invasion, setting the forces in motion to begin to reduce demand. Furthermore, substitutes (primarily natural gas and coal) are now more readily available, and some users are already equipped to switch to these alternate fuels. In addition, thanks to the earlier experience with high energy prices, effective conservation measures have been developed that can help counter possible oil market shortfalls.

Alternate fuels and substitute energy sources are less remote in 1990 than they were during the 1970s. The removal of price controls on natural gas has encouraged that industry to expand so that it can now assume a significantly larger share of the energy market. The potentially higher returns to gas producers since price decontrol have stimulated renewed drilling of gas wells as well as increased capture and use of the natural gas that is a joint product from oil wells. As oil prices rise, the distribution of gas, a more costly procedure than that for oil, could become increasingly profitable. Abundant nearby gas supplies in the United States, the Gulf of Mexico, and Canada now help bolster energy security over the long run.

Other energy sources—such as wind, solar heat, and biomass—are closer to a dependable availability for common usage in 1990 than they were in the 1970s. Thanks to crash research programs instituted during the 1970s, progress in harnessing these energy forms could now proceed more rapidly—after having essentially ceased during the 1980s when oil prices dropped—in the event that higher energy prices seem likely to be sustained.

Overall, Tugwell's book provides a worthwhile look at the history of the energy crises of the 1970s. Although the book may provoke readers to think more critically about U.S. energy policy, readers have reason to be skeptical about the public policy prescriptions Tugwell espouses.

Gene D. Sullivan

_____

Index for 1990

The Federal Reserve Bank of Atlanta
will sponsor a banking conference

THE BANKING INDUSTRY:

THE NEXT DECADE IS HERE

Rules, Risk, Reform

The conference program will include interactive
panel discussions as well as several featured speakers:

Former Senator William Proxmire
Governor Wayne Angell, Federal Reserve Board
David Lascelles, Banking Editor of *Financial Times*
William Taylor, Director, Banking Supervision and Regulation,
Federal Reserve Board

March 19-20, 1991

Westin Peachtree Plaza
Atlanta, Georgia

For information and registration,
call Ms. Jess Palazzolo at 404/521-8747

Economic Review

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

Address Correction Requested