A Slow Recovery in the Third District: Evidence From New Time-Series Models

*Theodore M. Crone**

A state budget director wants to know how much income tax revenue will be collected in the coming fiscal year. A department store manager wants to know whether sales will increase next quarter. A plumbing contractor wants to know how many new houses will be built in the spring and summer. Each of these persons will rely on some implicit or explicit forecast to make decisions about spending, inventory, or employment levels. Sometimes the forecast will be based on private information or just "gut feelings"; at other times it will be based on a more formal statistical model. In either case, the value of the forecast will depend on how accurate and how relevant it is. For example, if personal income is likely to vary significantly from the forecast or if the forecast is only for income at the national level, the budget director may not be able to accurately predict income tax revenue. And she may not know whether to recommend a reduction in spending in order to balance the state's budget.

The need for accurate forecasts of regional economic conditions has spurred the development of forecasting models for individual states. Recently, separate models have been developed for each of the states in the Third Federal

*Ted Crone is Assistant Vice President in charge of the Regional Economics section in the Philadelphia Fed's Research Department.

Reserve District—Pennsylvania, New Jersey, and Delaware. These models are based on time-series techniques and are currently forecasting a slow and gradual recovery from the recent recession.

TIME-SERIES MODELS DIFFER FROM EARLIER STRUCTURAL MODELS

The earliest regional forecasting models were developed in the 1960s and early 1970s and were fashioned after the national models of the day. They are often referred to as "structural models" because they are designed to capture in a set of equations the basic economic relationships or the "structure" of the regional economy.[1] Economic theory plays a critical role in the construction of these models. Theory dictates which variables are to be explained by the model (the endogenous variables) and which variables are inputs into the model but are not explained by the model (the exogenous variables). Theory also determines how the variables in the model interact. For example, as a family's income increases it will tend to spend more money, so the level of personal income should affect the level of retail sales in a region. Most of the items we buy, however, are produced and sold by many firms in nationwide, or even worldwide, markets on a competitive basis. So this increased demand for goods in a relatively small region will not necessarily put pressure on the prices of those goods and raise the general price level. On the other hand, since local manufacturers tend to use business services that are close by, an increase in manufacturing employment in a region may increase local jobs in business services. A complete structural model would consist of hundreds of equations specifying the full range of economic relationships. In practice, the small number of data series available at the state and local level severely limits the number of equations in a regional structural model.

In the 1980s several new regional forecasting models were developed applying time-series techniques (see *State Forecast Models Developed at Federal Reserve Banks*). These models differ from the more traditional structural models in their relationship to economic theory. More important for time-series models than any specific theory are the statistical regularities among economic variables. There may be a consistent pattern, for example, between personal income growth in a given quarter and the change in personal income and employment in the previous two quarters. Time-series models attempt to determine these consistent patterns from historical data and use them to forecast the future.

Even though statistical relationships form the basis of time-series models, the models are not totally divorced from economic theory. Model-builders will naturally include variables that theory suggests have some economic relationship to one another. Moreover, they sometimes impose restrictions on how certain variables may influence others. In a regional model, for example, it is often the case that past values of regional variables are not allowed to affect the current national variables. Past regional data are already contained in the national data included in most time-series models. Last quarter's employment in New Jersey, for example, is already incorporated into last quarter's national employment data, and New Jersey's employment should influence the national economy in basically the same way as employment in any other state.

Both structural and time-series models are in common use for national as well as regional forecasts. In terms of accuracy it is not clear that one has an advantage over the other. In studies that have compared time-series and structural

[1]For a brief description of some of these models, see Norman J. Glickman, *Econometric Analysis of Regional Systems* (Academic Press, 1977); and Roger Bolton, "Regional Econometric Models," *Journal of Regional Science*, 25 (1985), pp. 495-520.

State Forecast Models Developed at Federal Reserve Banks

Economists at several Federal Reserve Banks have developed state forecast models based on time-series techniques. The following is a list of articles on the major models developed over the past decade.

Amirizadeh, Hossain, and Richard M. Todd. "More Growth Ahead for Ninth District States," *Quarterly Review*, Federal Reserve Bank of Minneapolis (Fall, 1984).

Gruben, William C., and Donald W. Hayes. "Forecasting the Louisiana Economy," *Economic Review*, Federal Reserve Bank of Dallas (March, 1991).

Gruben, William C., and William T. Long III. "Forecasting the Texas Economy: Applications and Evaluation of a Systematic Multivariate Time-Series Model," *Economic Review*, Federal Reserve Bank of Dallas (January, 1988).

Gruben, William C., and William T. Long III. "The New Mexico Economy: Outlook for 1989," *Economic Review*, Federal Reserve Bank of Dallas (November, 1988).

Hoehn, James G., and James J. Balazsy. "The Ohio Economy: A Time-Series Analysis," *Economic Review*, Federal Reserve Bank of Cleveland (Third Quarter, 1985).

Hoehn, James G., William C. Gruben, and Thomas B. Fomby. "Time Series Forecasting Models of the Texas Economy: A Comparison," *Economic Review*, Federal Reserve Bank of Dallas (May, 1984).

Kuprianov, Anatoli, and William Lupoletti. "The Economic Outlook for Fifth District States in 1984: Forecasts from Vector Autoregression Models," *Economic Review*, Federal Reserve Bank of Richmond (February, 1984).

models of the national economy, neither type has been found to be consistently more accurate than the other.[2] Less rigorous comparisons have been made between time-series and structural models of regional economies, and the evidence again suggests that neither method is obviously superior.[3] Thus, the choice of which type of model to prefer will depend on the resources available and the ultimate use of the model.

Time-Series Models Have Some Advantages. The loose link between time-series models and economic theory can be an advan-

[2]Stephen K. McNees, "Forecasting Accuracy of Alternative Techniques: A Comparison of U.S. Macroeconomic Forecasts," *Journal of Business and Economic Statistics*, 4 (1986), pp. 5-15; Robert B. Litterman, "Forecasting with Bayesian Vector Autoregressions--Five Years of Experience," *Journal of Business and Economic Statistics*, 4 (1986), pp. 25-38; and Roy H. Webb, "Vector Autoregressions as a Tool for Forecast Evaluation," *Economic Review*, Federal Reserve Bank of Richmond (January/February, 1984).

[3]Paul A. Anderson, "Help for the Regional Economic Forecaster: Vector Autoregression," *Quarterly Review*, Federal Reserve Bank of Minneapolis (Summer, 1979).

tage. Even when two analysts disagree about how economic variables influence one another, both might accept the results of a time-series forecast because the models do not pretend to capture theoretical relationships. One can question the set of variables selected for a time-series model on theoretical grounds, but the historical pattern among the variables can be recognized without appealing to any economic theory.

Because time-series models do not have to account for the many relationships that define the structure of the economy, they require fewer variables than the traditional structural models. For this reason, they are better suited for state forecasts where data are limited. Typically, regional time-series forecasts contain only four to 10 variables for any individual state.

Time-series models can be estimated with a small number of variables and without the need to specify the theoretical relationships among them, and this reduces the amount of research time and computer resources needed to develop the models. As a result, economic analysts are able to develop and maintain time-series models more easily than traditional structural models.

Time-Series Models also Have Limitations. The greater accessibility of time-series models does not imply that they should always be the model of choice or that constructing them presents no difficulties.

Time-series models do not attempt to reproduce basic economic relationships, rendering them less useful than structural models in analyzing the effects of policy changes. In a structural model, exogenous policy variables, such as government spending or tax rates, directly affect other variables in the system. By generating forecasts using different assumptions about these policy variables, one can gauge how changes in policy would work their way through the economy.[4] In time-series models, there are no strictly exogenous variables, and policy variables are seldom represented in the models. Thus, the effects of policy changes are not so easily tracked.

This does not imply that time-series models are of no use for policy analysis. The baseline forecast from a time-series model implicitly assumes that policymakers will respond to any future shocks to the economy as they have in the past.[5] The time-series forecaster can, however, construct a "what if" scenario. A change in policy is likely to immediately affect some variable, such as a short-term interest rate, that is included in the model. For policy analysis, the forecaster must decide to what extent a policy change is likely to cause that variable to deviate from the baseline forecast. He can then use the model to estimate how the path of all the variables in the model would be altered as a result of the change in policy.

The development of time-series forecasting models often poses another problem. These models are susceptible to "overfitting." The problem arises when the number of explanatory variables in an equation is nearly as large as the number of observations we have on each variable. For example, we might try to predict this quarter's employment level using the past values of 30 other economic indicators for which

[4]This use of structural models for policy analysis is appropriate when analyzing small or marginal changes in some policy variable such as a tax rate. When there is a significant policy change, however, such as the introduction of a new tax, the various participants in the economy may react differently under the new policy than they would have under the old one. Since the parameters of the structural model are estimated under the old set of rules, they may tell us little about how the economy will respond to the new policy. See Robert E. Lucas, "Econometric Policy Evaluation: A Critique," *The Phillips Curve and Labor Markets*, Carnegie-Rochester Conference Series on Public Policy, 1 (1976), pp. 19-46.

[5]See Robert B. Litterman, "Forecasting and Policy Analysis with Bayesian Vector Autoregression Models," *Quarterly Review*, Federal Reserve Bank of Minneapolis (Fall, 1984).

we have only 40 quarters of data. In this case the estimated model may explain the historical data very well, but it may not be a good model for forecasting purposes. The estimated model may reflect not only the stable relationships among the variables but also those relationships that were peculiar to the period from which the data were drawn to estimate the model. When the model is then used to forecast, these temporary patterns will be projected into the future, diminishing the accuracy of the forecast.

There are several ways to overcome the overfitting problem. The models developed for the states in the Third District employ the so-called Bayesian Vector Autoregression (BVAR) method. Basically, this technique begins by limiting the weight that each explanatory variable can have on the forecast based on the model-builder's belief about how important that variable is likely to be. These initial restrictions are then gradually adjusted to improve the forecasting ability of the model. Each adjustment produces another specification of the model. After numerous adjustments, the specification that has the smallest forecast errors for a period not used in the estimation is chosen as the forecasting model. This method of choosing the final specification of the model makes it less likely that temporary patterns among the variables will be projected into the future.[6]

THE NEW MODELS FOR THE THIRD DISTRICT STATES

The selection of variables for a time-series model depends heavily on how the model is to be used and, of course, the available data. Since the new models for Pennsylvania, New Jersey, and Delaware are intended to forecast overall economic conditions in each state, three general state-level variables were included—establishment employment, personal income, and the unemployment rate. We would have included gross state product among the state variables, but the data are available only on an annual basis and are published with a long lag, rendering them of little use for forecasting purposes. Since we do not have a timely measure of gross state product, employment is often viewed as the most comprehensive measure of economic activity at the state level.[7] Personal income, the chief component of which is wages and salaries paid in the state, also reflects the general level of economic activity. And even though some components of personal income, such as rents, dividends, and interest, may be earned outside the state, they are likely to influence future economic activity in the state. The third variable of primary interest in these new models is the state unemployment rate, an indicator of the overall slack in the economy.

Besides the three variables reflecting general economic activity, the models for Pennsylvania

[6]For an excellent introduction to BVAR models and the overfitting problem, see Richard M. Todd, "Improving Economic Forecasting with Bayesian Vector Autoregression," *Quarterly Review*, Federal Reserve Bank of Minneapolis (Fall, 1984). For a more technical discussion see Thomas Doan, Robert Litterman, and Christopher Sims, "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3 (1984), pp. 1-100. Economists at the Dallas Fed and at the Cleveland Fed have devised a two-step method for limiting the number of explanatory variables in time-series models and thus mitigating the overfitting problem. See James G. Hoehn and James J. Balazsy, Jr., "The Ohio Economy: Using Time-Series Characteristics in Forecasting," Federal Reserve Bank of Cleveland, Working Paper 8508 (1985); and William C. Gruben and William T. Long III, *Economic Review*, Federal Reserve Bank of Dallas (January, 1988).

[7]We use nonfarm establishment employment rather than resident employment because the establishment employment series contains smaller measurement error. For smaller states like Delaware, resident employment is estimated using nonfarm establishment employment and an estimate of self-employed and agricultural workers.

and New Jersey contain two more state-level variables, and the model for Delaware contains one more (see *State-Level Variables in the Forecast Models*). Housing permits are included in the models for all three states. The residential construction industry often leads the economy over the business cycle, since the purchase of a new house generally results in the purchase of other goods, such as appliances and furniture, and relatively good data are available on housing permits. The models for Pennsylvania and New Jersey also include retail sales to reflect the strength of the consumer sector at the state level. These sales data are not available for Delaware. In each of our models any state-level variable is permitted to influence any other state-level variable.

Since the national economy plays such an important role in most state economies, several national variables are included in the models (see *National Variables in the Forecast Models*). According to common practice, the national variables are allowed to influence any of the state variables, but the state variables are not allowed to influence the national variables. The national counterparts to the five state-level variables are included in the forecast models. In addition, the models include gross domestic product and the spread between the 10-year Treasury bond yield and the federal funds rate. Gross domestic product, or the value of all goods produced in the U.S., is included because it is the most comprehensive measure of the domestic economy. The final national variable, the spread between the 10-year Treasury bond yield and the federal funds rate, reflects conditions in financial markets. Several recent studies have found that interest rate spreads contain valuable information in forecasting the national economy.[8] And the inclusion of this particular spread substantially improved the state forecasts.

Three major adjustments are made to both the national and state-level variables in the time-series models for Pennsylvania, New Jersey, and Delaware. First, all the variables except the interest rates are adjusted for seasonal variation. There is no seasonal variation in interest rates. Second, all variables expressed in dollar terms (gross domestic product, personal income, and retail sales) are in constant dollars, that is, in 1982 dollars for personal income and retail sales and in 1987 dollars for

State-Level Variables in the Forecast Models

Nonagricultural Establishment Employment
Personal Income
Unemployment Rate
Housing Permits
Retail Sales (Pennsylvania and New Jersey)

National Variables in the Forecast Models

Gross Domestic Product
Nonagricultural Establishment Employment
Personal Income
Unemployment Rate
Housing Permits
Retail Sales
Spread Between 10-Year Treasury Yield and Fed Funds Rate

[8]See Ben Bernanke, "On the Predictive Power of Interest Rates and Interest Rate Spreads," NBER Working Paper 3486 (October, 1990); and Benjamin M. Friedman and Kenneth N. Kuttner, "Why Does the Paper-Bill Spread Predict Real Economic Activity?" Federal Reserve Bank of Chicago, Working Paper Series on Macro-economic Issues 91-16 (September, 1991).

gross domestic product. Third, in estimating the models and producing forecasts, we use the log of each variable except for the unemployment rates and the interest rates.

The new models were estimated using quarterly data, and the previous four quarters of all the variables in the models were allowed to influence the forecast of each state-level variable. The final specification chosen for each model was based on how well it would have forecast the state's economy between 1981 and 1990.[9]

WHAT LIES AHEAD FOR THE THIRD DISTRICT STATES?

The overall economic condition in each state is best reflected by three variables in the new forecast models: employment, personal income, and the unemployment rate. Historical data on employment and personal income show that the recession that began on the national level in July 1990 was more severe in each of the three states in the Third District than in the nation generally. The current forecasts of employment and personal income from the new time-series models indicate that the region's economy will recover slowly from the recent recession. In fact, growth will not be rapid enough to significantly lower unemployment rates from their current levels.

The timing of the recent downturn varied from state to state in the Third District. Employment in New Jersey began to decline three years ago, in the second quarter of 1989. One year later, job levels began to fall in Pennsylvania. Delaware followed the national pattern much more closely; jobs began to decline in the third quarter of 1990. The resumption of job growth in the three states is occurring in the reverse order. Delaware seems to have entered a period of sustained job growth in the fourth quarter of 1991. Jobs in Pennsylvania also increased in the final quarter of 1991 but backtracked somewhat in the first quarter of this year. Employment was still declining in New Jersey in the first quarter of 1992 (Figure 1).

The forecasts of employment growth from the first quarter of 1992 to the first quarter of 1993 reflect this staggered timing of the recovery: Delaware's job growth over the next four quarters is projected to be greater than 2.5 percent, Pennsylvania's to be 0.7 percent, and New Jersey's only 0.6 percent.[10] These projected growth rates are well below the average first-quarter to first-quarter growth rates during the 1982 to 1990 national expansion, which ranged from 2.1 percent for Pennsylvania to 4.1 percent for Delaware. For the region as a whole, the new forecasts signal a slow, gradual recovery from the recent recession.

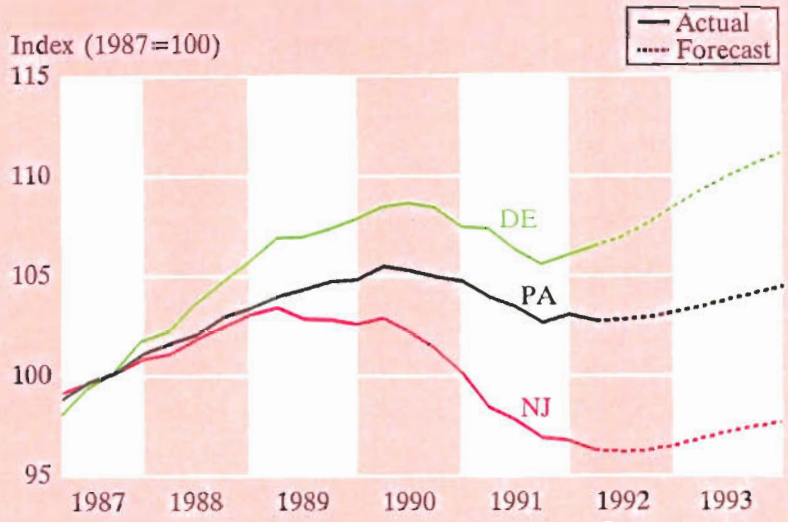
During the 1990-91 recession, the decline in real personal income in the region lasted for a much shorter period than the decline in employment. Real personal income fell for three consecutive quarters in Pennsylvania and New Jersey and for two consecutive quarters in Delaware. The percentage decline in each state, however, was greater than the decline at the national level. In all three states real personal income has already recovered somewhat from its recession low, and further improvement is forecasted through 1992 (Figure 2). Each state's real personal income should increase about 1 percent or more from the last quarter of 1991 through the last quarter of this year. The increase in Delaware is projected to be much greater than the increases in the other two

[9]The models are re-estimated as new data become available, and the estimates are based on the constraints imposed when the models were first developed.

[10]All the forecasts reported in this article are based on the data available May 15, 1992. At that time employment and unemployment rates were available through the first quarter of 1992. Personal income at the state level was available only through the fourth quarter of 1991.

Figure 1

Employment Growth in the Region 1987



Establishment Employment
Percent Change 1992:I to 1993:I

	Pennsylvania	New Jersey	Delaware
Model Forecast	+0.7%	+0.6%	+2.6%
Range (+/- mean absolute forecast error 1981:I to 1990:IV)	-0.2% to +1.6%	-0.1% to +1.3%	+1.0% to +4.2%

states. But each state's projected growth rate of personal income is lower than the average fourth-quarter to fourth-quarter growth rate during the 1980s' expansion, which ranged from 2.6 percent for Pennsylvania to 4.6 percent for Delaware.

These patterns of slower than average growth will keep the state unemployment rates above their prerecession levels through the first quarter of 1993. The forecasted rates for the first quarter of 1993 for Pennsylvania (7.1 percent) and for Delaware (5.0 percent) are virtually unchanged from the first-quarter 1992 rates. In New Jersey the unemployment rate is projected to decline to 6.2 percent, but this is still well above its prerecession level of less than 4 percent.

HOW ACCURATE ARE THESE FORECASTS LIKELY TO BE?

All forecasts, whether derived from structural or time-series models, are subject to error. Some indication of the possible error in these new state forecasts is available from the past performance of the models. In developing the models, we calculated the errors these models would have produced for forecasts one quarter ahead and four quarters ahead of the latest available data. The mean absolute forecast errors from 1981 through 1990 are presented in the table on page 12.

Two general patterns appear in these forecast errors. First, the errors for any particular variable become larger as the time horizon increases. The further into the future one looks, the more difficult it is to predict the course of the economy. Second, the errors are generally larger for the small state of Delaware than for the two larger states of Pennsylvania and New Jersey.

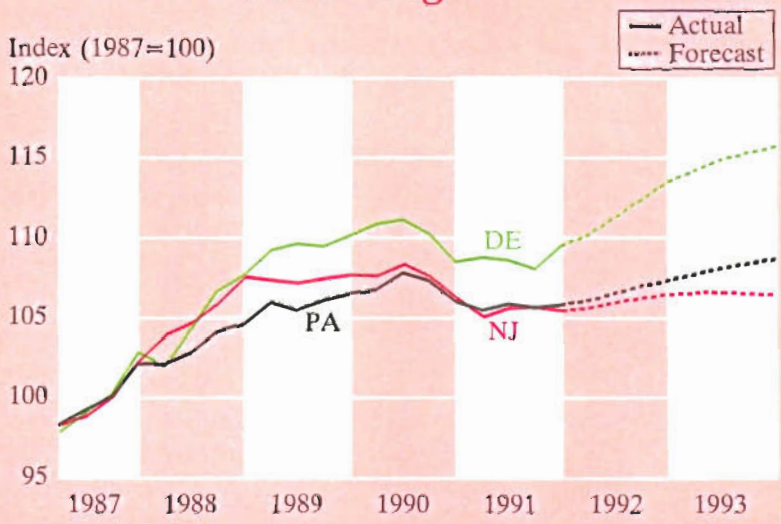
What do these historical errors imply about the current forecasts? They underscore how slow the regional recovery is likely to be. Employment in Pennsylvania and New Jersey is currently forecast to grow so slowly that the

predicted growth rates over the next four quarters are less than the average absolute forecast errors in the 1980s. In other words, given the history of the forecasts and the slow growth scenario being predicted, we cannot rule out no growth in employment in those two states for the year (see the ranges in Figure 1). This is not the most likely outcome but a possible one.

Time-Series Forecasts Should Not Be Used in Isolation. These new forecasts, like all forecasts, are not precise, and therefore they should be used with other information about the nation or the region. The models generate their own forecasts of the national variables. But it is possible to add information by substituting forecasts of national variables from other time-series or structural models. In this way several state forecasts can be derived under alternative scenarios for the national economy. Currently the new models are projecting a pattern for the national economy similar to that of the major forecasters, so the use of other national forecasts at this time would have little effect on the state forecasts.[11]

Figure 2

Real Personal Income Growth in the Region



Real Personal Income
Percent Change 1991:IV to 1992:IV

	Pennsylvania	New Jersey	Delaware
Model Forecast	+1.5%	+0.9%	+3.6%
Range (+/- mean absolute forecast error 1981:1 to 1990:IV)	+0.5% to +2.5%	-0.9% to +2.7%	+1.6% to +5.6%

Historical forecast errors also suggest that econometric forecasts should not be used to the

[11]State-level forecasts were generated from the new models using DRI's forecasts for the national variables. The employment and personal income growth rates for the three states were little changed from the forecasts reported in Figures 1 and 2. The first-quarter over first-quarter employment growth estimate was reduced by 0.1 percentage point for New Jersey and Delaware and was increased by 0.1 percentage point for Pennsylvania. The fourth-quarter over fourth-quarter personal income growth estimate was increased by 0.4 percentage point for Delaware, 0.3 percentage point for Pennsylvania, and 0.1 percentage point for New Jersey. In a procedure similar to our use of the DRI forecast, researchers at the Dallas Fed used the *Blue Chip* consensus forecast of national variables in the forecasting phase of their model of the Louisiana economy. See Gruben and Hayes, *Economic Review*, Federal Reserve Bank of Dallas (March, 1991).

exclusion of other types of information about future conditions. For example, the Philadelphia Fed conducts a monthly survey of manufacturers in the District (*Business Outlook Survey*), inquiring about current business conditions and expectations for the following six months. This survey has been found to contain reliable information about the future course of the regional economy.[12] The index of current activity from the survey has recently turned positive, and expectations are still high, indicating continued improvement in manufacturing. This is consistent with the evidence from the new forecast models.

CONCLUSION

Like all tools of economic analysis, time-series models have certain advantages and limitations. They are not as useful as structural models in analyzing changes in policy. They are particularly helpful, however, in forecasting regional economies for which data are limited. The new models for the Third District states are intended to forecast general economic conditions in the region. Current forecasts suggest that a sustained but gradual recovery will be in place throughout the region in the second half of 1992. Like all forecasts, these are subject to error and should not be used to the exclusion of other information.

[12]See John Bell and Theodore Crone, "Charting the Course of the Economy: What Can Local Manufacturers Tell Us?" this *Business Review* (July/August, 1986).

Average Absolute Errors of the Forecasts
One Quarter and Four Quarters Ahead
1981:I to 1990:IV

	PA	NJ	DE
Employment Growth			
One Quarter	0.3%	0.3%	0.5%
Four Quarters	0.9%	0.7%	1.6%
Real Personal Income Growth			
One Quarter	0.5%	0.6%	0.8%
Four Quarters	1.0%	1.8%	2.0%
Unemployment Rate			
One Quarter	0.4%	0.3%	0.4%
Four Quarters	1.0%	0.6%	0.7%