Traditional and Nontraditional Data as Indicators of Economic Activity in Rural Communities

Richard McHugh, Ronald Cooper, Sheila Martin,
Daniel Otto, and Stanley R. Johnson

Staff Report 90-SR 40

January 1990

Traditional and Nontraditional Data as Indicators of Economic Activity in Rural Communities

by Richard McHugh, Ronald Cooper,
Sheila Martin, Daniel Otto, and Stanley R. Johnson

Staff Report 90-SR 40
January 1990

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011

Richard McHugh is an associate professor in the Department of Economics, University of Missouri; Ronald Cooper is with the National Governors' Association; Sheila Martin is a graduate assistant at CARD; Daniel Otto is an associate professor with the Department of Economics, Iowa State University; and Stanley R. Johnson is a professor with the Department of Economics and the director of CARD.

This paper was prepared with funding assistance from the Aspen Institute and the National Governors' Association.

Contents

Tables

Abstract

The objective of this study was twofold: (1) to identify nontraditional sources of data that can be used to monitor economic activity in rural areas, and (2) to test the relationship between such data and trends in traditional measures, such as income and employment.

The need for alternative sources of data arises primarily out of the inconsistent quality and availability of traditional economic data at the substate level. New sources of data can help reduce this metro-rural data gap and provide a fuller picture of the diverse experience and structure of substate economies.

Two sources of nontraditional data were explored: state government administrative records, and data from local private and quasi-private companies, such as utilities and banks. In particular, banking data from the Federal Reserve Bank of Chicago, phone connection data from Northwest Bell (U.S. West), and food stamp program data were used. The data series consisted of county-level data for Iowa, Minnesota, and Wisconsin.

These data were statistically compared to employment growth levels in each county. Banking data (reflecting bank willingness and ability to lend), phone data (reflecting new household and business formations), and food stamp participation correlated significantly (though weakly) with growth. The weak correlations are expected because the nontraditional data series are generally more volatile, or erratic, than employment data series.

Composite indexes were constructed from the data and correlated with the employment data series. The results showed improved correlations and suggest that such indexes may be useful as leading indicators of economic activity. Overall, the results show that there is considerable potential in the development of economic indicators from nontraditional data.

Introduction

Over the past two decades in the United States, economic growth and decline have varied significantly by industry and region. While one group of industries might have performed well for a period of time, another group of industries simultaneously will have floundered. A few years later the fortunes of these industries will have changed. For example, during the past seven years, the construction, manufacturing, mining, and farming industries have had periods of economic weakness, although their overall level of economic activity has grown continuously. Likewise, and in part as a consequence of diversity in industrial performance, some regions of the country have experienced rapid growth and the burdens of tight resource markets while, at the same time, other regions have grappled with the consequences of little or no economic growth.

An important implication of this spatial and industrial diversity in economic conditions is that broad measures no longer can be comfortably relied upon as comprehensive indicators of economic performance. Economic growth can vary dramatically within and among states. This has been particularly evident in the past few years, during which metropolitan areas of many states have generally fared much better than the rural areas. To some extent, these discrepancies in performance have been the case even among rural areas. Declining real net farm income and the adoption of laborsaving technologies are the most often cited causes of rural economic decline.

The interindustry and interregional diversity of economic structure and performance presents a special challenge to policymakers. When economic trends were more uniform across sectors of the economy, policymakers responded to problems with general, nontargeted measures. Tax reform and abatement and regulatory reform are examples. However, these blunt policy instruments are inappropriate in circumstances of interindustry and interregional divergences in economic activity. In short, "targeting" has become an important element of effective domestic economic development policy. Along with the concept of targeting has come the requirement for more state and local government involvement in economic development policy.

Effective targeting of economic development resources and efforts requires an ability to set priorities and evaluate outcomes by closely monitoring economic performance within states, industries, and regions. In fact, the monitoring of local economic conditions is one of the most important tasks of state government. For the design of specialized economic development programs, policymakers must be more fully aware of the underlying economic conditions in areas within states. And, as noted, aggregate indicators of economic performance can give false impressions of local and industry-specific activity. Overall state economic health may mask local economic decline. Recently this has been especially the case for many rural areas.

State governments also require indicators that will do more than simply reflect the pace of past economic activity by region or industry. For effective intervention and management, state policymakers must be knowledgeable of economic problems as they emerge. Ideally, state policymakers should have information reflecting preconditions or precursors of trends in local or regional economic activity. In short, there is a need for high-quality economic data for subareas within states that can be used effectively for monitoring economic activity and identifying areas for policy action.

Many of the economic development policies undertaken in the past few years by state and local government have been highly experimental in nature. These have resulted from a combination of political pressures from areas or industries in distress and a fragmented concept of the development process. To illustrate, customized labor force training and main street initiatives are more or less untested policies that have become commonplace. And, the innovations in economic development policy show no sign of abating. An improved system for timely monitoring of economic activity, providing rapid feedback to help in fine-tuning the design of new and untested development policies, is critical to efficient economic development programs.

Unfortunately, economic performance data for rural areas are limited. And, the most widely cited sources of these data are from the federal government. The U.S. Departments of Commerce and Labor, in particular, produce nearly all of the available county-level economic data. There are major shortcomings of these traditional data for development policy: first, while timely information on economic activity may be available nationally, many of the federal data series are available at the state and substate levels only with substantial lags; and second, the geographic coverage of this data often is unbalanced, with larger areas and metropolitan areas having more and better traditional performance data than smaller rural areas. As examples, unemployment rate and labor force data, obtained from a monthly survey of households, are available only for the larger states and some metropolitan areas. Area wage level surveys are available only for metropolitan areas; cost of living and price index data are only available for multistate regions and larger metropolitan areas. Industrial production, capital investment, and construction are other instances in which data are available for larger areas but not for rural communities.

In sum, quality data adequate for accurately monitoring rural economies can support improved development policy. Yet, the availability of data for this purpose from traditional sources is limited. If policy analysts are to give serious attention to the more careful monitoring and evaluation of rural economies and rural development, new and innovative means of data collection and analysis must be initiated. The theme underlying this exploratory project is that there exists a deep and presently untapped well of information on economic activity, data of a nontraditional nature, that can be used to fill an existing and growing gap in the capacity of traditional data for meeting development policy needs for rural areas.

Objectives and Approach

The intent of this project has been to investigate the possibility of accessing and utilizing existing yet unexploited nontraditional data sources, and to evaluate the extent to which these sources may fill some of the more evident gaps in the traditional data on the economic performance for rural communities. The overriding objective of the project has been to identify unique and nontraditional sources of data that have potential use in monitoring local economic activity and supporting economic development initiatives in rural areas.

To accomplish this objective, a process involving four distinct steps was followed. First, potential types and sources of nontraditional data were identified and reviewed for their potential relative to existing sources. Second, those responsible for maintaining the databases were contacted as a basis for better understanding collection and processing of these data. Third, the availability of the data from these nontraditional sources was assessed. This required evaluating possibilities of establishing ongoing relationships with the responsible organizations, checking for the existence of an adequate time series of data that could be used to test their reliability and usefulness, and gauging the geographic coverage of the available data. Fourth, the capacity of the data for producing indicators of economic activity that supplement those available from more traditional sources was evaluated.

Data Sources

The data for monitoring economic activity in local, rural areas must have a number of distinguishing features. First, data should be reported at a detailed geographic level. Unless at a minimum county-level economic activity can be identified, we cannot categorize the economic activity as having occurred in a rural area or urban area, for example. Counties are many times also a focus of state interventions and policies for economic development.

Second, the data should be available on a timely basis. The lag between the economic impact of an event and the reporting of performance data should be short. As has been observed, a problem with much of the traditional data is the lag between the economic event and the reporting of economic performance. It does little good to identify additional nontraditional sources of data with substantial reporting lags.

Third, these data should reflect broad-based economic trends. For example, it may be of little value for analysts of current economic trends to know that attendance at movies has slackened, or that the consumption of ice cream is up, although both are in part determined by broader economic factors. On the other hand, migration trends, banking activity, energy use, and communications are examples of features of the economies that more fully reflect the overall economic environment and the level of economic activity.

Finally, since all economic data are available with some lag, it is desirable to have more specialized economic measures that do not lag (and ideally lead) general economic performance. These leading indicators can provide input for tuning and adapting economic development assistance policy in anticipation of future events, or they can point to areas in which development assistance will be needed. Given the capacities of state and local government to change and adapt, these possibilities for lead time are especially important.

Two general sources of nontraditional data were explored: state government administrative records, and data from local private and quasi-private companies with location-specific business. Administrative records contain information that must be collected by state agencies to monitor particular programs or carry out routine governmental functions. For example, a state may assemble information on sales tax collections at local levels, particularly if local governments are permitted to levy a sales tax; a state may compile information on housing activity as a part of an effort to equalize property tax assessments for local governments; or a state may collect information on food stamp distributions or unemployment insurance claims from local administrative offices, with these offices automatically supplying local detail as a standard reporting requirement.

These administrative data can be a potentially valuable source of information for augmenting the current understanding of the economic activity in rural areas and/or smaller communities. And, to use these sources, no new costs need be incurred to collect primary data. But, to encourage use, these data must be evaluated for their potential in augmenting traditional information sources and arrangements must be made for their timely release and processing.

Private and quasi-private industries in many cases have freely available data on location-specific economic activity. For example, many industries that are or have been regulated maintain well-developed reporting systems that can be accessed, given the cooperation of the appropriate authorities. Utilities usually provide regulators with information on sales and income generated from various activities by service area. Financial institutions are also subject to significant governmental oversight, requiring maintenance of location and even industry-specific information. Thus, unlike many corporate reports, which provide only highly aggregated calculations on company activity, the locational specificity of these internal data is a key to their usefulness as indicators of performance regional economies.

For nontraditional data sources from both the governmental and private sectors, the problems of access and use for improving information on rural and smaller communities is one of organization. First, the data are collected for other purposes than the development of the desired indicators. For this reason, the data may be destroyed after a short period or processed in forms that need refinement and specialized interpretation. Second, there is the question of deciding what is important in the specialized arrays of data and how they can augment and be tailored to supplement what is already known about the regions of interest.

For this study, selected possible nontraditional data sources were investigated. From the list of alternatives, three data sources appeared to have the necessary qualifications and were also currently available with a sufficient time series to allow for an evaluation of their capacities to reflect local economic conditions. Of course, exclusion of potential data sources from the list of those used should not be interpreted to indicate that they are inherently unacceptable. Rather, data sources were excluded because insufficient historical data were available (although this could be changed given the announced desire for the data), the delivery time for the data was unduly long (although experience in the provision of the data could lead to a significant shortening), or presently unavailable legal clearance for the release of the data was required.

Potential sources of nontraditional data for economic indicators identified at the outset of the project included the following:

- Indicators of business activity
  - New incorporations
  - New electricity hook-ups
  - Retail sales
  - Business telephone spending
  - Electricity usage
  - Construction activity
  - Vehicle miles traveled
  - Commercial and industrial loan activity
  - Debt and asset positions of commercial institutions
- Indicators of wealth and personal income
  - Average wage
  - Individuals receiving food stamps
  - Children approved for school lunch programs
  - Property valuation
- Indicators of population change
  - Residential building permits
  - Telephone hook-ups
  - School enrollment
  - Mortgage loans outstanding

Many of these potential sources were omitted from the analysis because of previously mentioned problems in obtaining the data or because of the spatial or temporal coverage, likely due only to present uses and practices. A detailing of the reasons that selected data sources were not included in this exploratory analysis is supplied in the Appendix.

The three data sources eventually selected were for banking, telephone connections, and food stamps. The Federal Reserve Bank of Chicago provided the banking data; each of the three states in this pilot study (Iowa, Minnesota, and Wisconsin) is located in the Chicago Federal Reserve Bank District. The data were compiled from the consolidated report of conditions of banks in the district. These reports provide data on income, assets, liabilities, and earnings of financial institutions, by geographic area.

The food stamp participation and cost data were collected from appropriate agencies of the three states: the Department of Human Services in Iowa and Minnesota, and the Department of Health and Social Services and the Economic Assistance Bureau for Wisconsin. These county-level data are used for program management but can be reorganized to meet nonprogram management uses as well.

The telephone data were obtained from Northwestern Bell (U.S. West Communications). This company serves both Iowa and Minnesota, but it conducts no business in Wisconsin. Wisconsin Bell could provide data by exchange on an annual basis only. In essence, there were no usable telephone data for Wisconsin. It is emphasized that this is not because these data do not exist someplace in the system but rather because of current company uses and protocols. Since there was not a full complement of data for Wisconsin, the assessment of these nontraditional data for tracking economic trends in Wisconsin was not performed.

The Federal Reserve System banking data can be used to measure potential for economic expansion, as indicated by the willingness and ability to lend funds that would subsequently be reflected in real economic activity. These data can also provide an early measure of a continuance of ongoing economic trends reflected in delinquency rates, charge-offs, or recoveries. In a sense, the banking data can track the same activities as several important components of the U.S. Department of Commerce's national composite leading index--change in credit outstanding, money supply, new business formations, and real estate activity. (See U.S. Department of Commerce 1984.) Lending activity at the banks would reflect these same forces. In short, the Federal Reserve Bank data appear to be a valuable source of location and industry-specific economic data.

The telephone connections data also capture impacts of variables reflecting contemporaneous economic conditions and also may indicate potential for future expansion. New residential connections reflect the in-migration and new household formation from residents already within a community. The new business connections are clearly reflective of business formation, a leading indicator of economic activity.

Finally, the food stamp data are reflective of current overall economic circumstances, since eligibility for food stamps is based upon a means test that is defined using household personal income. With new enrollment processes, there is not a significant waiting time for the means test or for request of stamps.

From these three nontraditional data sources, we selected 17 basic data series, which could be highly correlated with, or closely tied to, local economic development and performance trends.

From the Federal Reserve Bank (county-level) banking activity data was selected:

- Deposit-liability ratio
- Percentage change in construction and real estate loans
- Percentage change in commercial and industrial loans
- Percentage change in total loans
- Construction loans as share of total loans
- Commercial and industrial loans as share of total loans
- Percentage change in total charge-offs
- Percentage change in total recoveries
- Charge-offs as a share of total loans
- Recoveries as a share of total loans
- Percentage change in interest income
- Percentage change in net income
- Income-asset ratio

In most cases, a positive relationship between each of these variables and the level of economic activity would be expected. Possible exceptions to this hypothesized positive relationship are the level of charge-offs or charge-offs as a percentage of income.

From the food stamp data, two variables are expected to be correlated with the level of economic activity:

- Numbers of households receiving food stamps
- Value of food stamps received by households

In each case, a negative relationship between these variables and the level of economic activity in the county or local community was anticipated.

From the telephone data, two variables were constructed:

- Increase in number of households with phones
- Increase in number of business hook-ups

It was expected that both these variables would correlate positively with economic activity in the community or area.

The telephone and the food stamp data are reported monthly. The banking data are reported quarterly. Given the reporting frequency of the banking statistics, the evaluation of all these data series was at the quarterly level for consistency. It is observed that the banking and telephone data are in principle available on a real-time basis, given current methods of collections. Longer-term cooperative efforts with the companies or regulators could result in more current availability of these data.

Results

Initially, the degree to which the nontraditional indicators were associated with a more traditional indicator of economic activity was tested. The traditional indicator of economic activity utilizes the quarterly level of employment by each county. Since the interest was to use these variables as traditional indicators of economic growth, the series of quarterly employment levels were differentiated and transformed into percentages.

For analysis of the correlation between the nontraditional indicators and the level of employment growth, the sample of counties was partitioned into three categories: metropolitan areas (defined as those counties located within a standard metropolitan statistical area); farm counties (defined as nonmetropolitan counties within which at least 20 percent of income is from farm sources), and nonmetropolitan/nonfarm counties (the residual).

Simple Indicators

Simple coefficients of correlation between the indicator variables and the level of economic growth reflected by the employment data were calculated and are reported in Tables 1 and 2. For several of the variables, correlations were computed both for the variables levels and for the difference in the variables from one period to the next. The second, fourth, and sixth columns of these two tables (noted change) contain the estimated correlations between the transformed variables.

Several generalities emerge from the results in Tables 1 and 2. First, the direction of the correlation between the nontraditional indicator variables and the growth in employment was generally of the expected sign. For the banking data, most of the variables measuring loan activity were positively related to employment change. The ratio of charge-offs to total loans was generally negatively related to economic activity, while recoveries were positively related.

The food stamp series, particularly the changes in the numbers of households and the value of food stamps, showed a consistent negative correlation with changes in employment. While there were occasionally variables not correlated in the expected direction, these coefficients typically were not statistically significant. There were a few cases in which the nontraditional indicators had both correct signs and highly significant correlations between themselves and the growth in employment. Although disappointing, this lack of statistically significant relationships should not be unexpected because there was a very high degree of variation in some of the nontraditional indicators, while

Table 1. Coefficients of correlation between nontraditional indicators and total employment: Iowa

Data source	Farm Level	Farm Change	Metropolitan Level	Metropolitan Change	Nonmetro/Nonfarm Level	Nonmetro/Nonfarm Change
Banking						
Deposit/liability	.0373[a]	-.0071	.1080[c]	.0196	.0808[c]	-.0253
	(.1216)	(.7707)	(.0428)	(.7183)	(.0077)	(.4120)
Change in real estate loans	.0140	----	.0542	----	.0486	----
	(.5902)	----	(.3188)	----	(.1229)	----
Change in C & I loans	.0050	----	.0042	----	-.0112	----
	(.8366)	----	(.4402)	----	(.7162)	----
Change in total loans	.0682[d]	----	.0790	----	.0574[b]	----
	(.0053)	----	(.1454)	----	(.0620)	----
Real estate/total loans	-.0042	.0128	.0135	.0484	-.0215	.0473
	(.8613	(.6236)	(.8015)	(.3728)	(.4794)	(.1338)
C & I total loans	-.0148	-.0297	-.0627	.0053	.0017	-.0490
	(.5380)	(.2244)	(.2404)	(.9220)	(.9567)	(.1122)
Change (%) in charge-offs	.0147	----	.0249	----	.0368	----
	(.6781)	----	(.7505)	----	(.4072)	----
Change (%) in recoveries	-.0246	----	-.1474[c]	----	-.0030	----
	(.4199)	----	(.0190)	----	(.9386)	----
Charge-offs/total loans	-.0569[b]	-.0232	-.0946	-.0581	-.0670	.0546
	(.1059)	(.5262)	(.2298)	(.4743)	(.1309)	(.2348)
Recoveries/total loans	.0949[d]	.0256	.0401	-.1531[c]	.0636	-.0512
	(.0004)	(.4016)	(.4856)	(.0148)	(.0589)	(.9757)
Change in interest income	.0276	----	-.0761	----	-.0058	----
	(.4323)	----	(.3311)	----	(.8959)	----
Change in net income	.0023	----	.0310	----	.0077	----
	(.9483)	----	(.6924)	----	(.8617)	----
Income/asset	-.0095	.0029	-.0594	-.0579	-.0362	.0139
	(.7226)	(.9230)	(.3018)	(.3599)	(.2820)	(.7180)
Food stamp						
Houses on food stamps	.0230	-.1846[d]	-.0089	-.2606	-.0434	-.1598[b]
	(.3399)	(.0001)	(.8680)	(.0001)	(.1519)	(.0001)
Value of food stamps	.0416[c]	-.0788[d]	.0128	-.0938	-.0156	-.1089
	(.0837)	(.0012)	(.8123)	(.0001)	(.6070)	(.0004)
Telephone						
Business gain	.1145[c]	----	.1453[b]	----	.1238[c]	----
	(.0227)	----	(.0966)	----	(.0399)	----
Residential gain	.0892[b]	----	.0909	----	.0747	----
	(.0763)	----	(.2999)	----	(.2157)	----

[a]Figures on top are the coefficients of correlation. Those in parentheses below are levels of statistical significance.

[b]Significant at 10 percent level or better.

[c]Significant at 5 percent level or better.

[d]Significant at 1 percent level or better.

Table 2. Coefficients of correlation between nontraditional indicators and total employment: Minnesota

Data source	Farm Level	Farm Change	Metropolitan Level	Metropolitan Change	Nonmetro/Nonfarm Level	Nonmetro/Nonfarm Change
Banking						
Deposit/liability	.016[a]	.046	-.006	-.013	-.040	.008
	(.706)	(.303)	(.895)	(.776)	(.239)	(.803)
Change in real estate loans	.041	----	-.011	----	-.023)	----
	(.374)	----	(.823)	----	(.520)	----
Change in C & I loans	.018	----	.043	----	-.051	----
	(.686)	----	(.360)	----	(.129)	----
Change in total loans	-.009	----	.030	----	.038	----
	(.834)	----	(.522)	----	(.265)	----
Real estate/total loans	-.004	.041	-.006	-.014	.057[b]	-.023
	(.930)	(.367)	(.890)	(.769)	(.088)	(.507)
C & I total loans	.008	-.015	.035	.937	.029	-.072[b]
	(.861)	(.735)	(.450)	(.432)	(.392)	(.040)
Change (%) in charge-offs	.092	----	.009	----	-.022	----
	(.157)	----	(.899)	----	(.663)	----
Change (%) in recoveries	.026	----	.006	----	-.028	----
	(.632)	----	(.921)	----	(.517)	----
Charge-offs/total loans	-.019	.018	-.061	.003	-.063	.031
	(.775)	(.795)	(.376)	(.965)	(.208)	(.550)
Recoveries/total loans	.035	.025	-.014	.006	-.004	-.032
	(.466)	(.651)	(.786)	(.922)	(.923)	(.461)
Change in interest income	.101	----	.054	----	-.020	----
	(.126)	----	(.437)	----	(.693)	----
Change in net income	.084	----	.019	----	.066	----
	(.196)	----	(.786)	----	(.188)	----
Income/asset	-.081	.053	-.048	.037	-.008	.048
	(.094)	(.347)	(.348)	(.526)	(.898)	(.266)
Food stamp						
Houses on food stamps	.202	-.082	-.201	-.040	-.069	.045
	(.856)	(.496)	(.039)	(.710)	(.328)	(.558)
Value of food stamps	.017	.020	-.207	-.028	-.085	.063
	(.881)	(.865)	(.034)	(.795)	(.228)	(.410)
Telephone						
Business gain	.035	----	-.001	----	.040	----
	(.681)	----	(.983)	----	(.525)	----
Residential gain	.012	----	-.076	----	.040	----
	(.890)	----	(.344)	----	(.531)	----

[a]Figures on top are the coefficients of correlation. Those in parentheses below are levels of statistical significance.

[b]Significant at 10 percent level or better.

employment tends to grow at relatively stable rates. Finally, there appears to be a high degree of conformity in the signs and significance of these coefficients across the three classes of counties.

Among the indicator variables that correlated well with economic activity, as measured by employment growth, were the ratio of charge-offs to total loans, the percentage increase in total loans, the number of households receiving food stamps (negatively related to employment growth), and the percentage increase in the value of food stamps. The telephone connections series had the anticipated signs when correlated with employment change but was not statistically significant.

Composite Indicators

The series investigated in the previous section can be extremely volatile, particularly given the small geographic areas represented. And, some of the volatility in these series may be idiosyncratic in the sense that their reporting in a particular month may have less to do with the economic conditions than, for example, with administrative decisions as to when reporting occurs (such as the reported profits for a time period), or because of backlogs or delays in recording. Also, for smaller areas in particular, a relatively minor change in a variable may translate into a statistically large effect because of a low typical value. In statisticians' parlance, the signal-to-noise ratio may be low for the specific series. That is one reason why many of the potential indicators examined had low statistical correlations with economic activity as measured by employment, although conceptually they were inextricably related to local economic activity.

One way to limit the impact of this perhaps spurious variability in the series is to create composite indices using a number of indicator variables. In the composites, the statistical impacts of spurious variability in a component series in any one period may be reduced. The addition of a new variable to the composite index reduces the weight attached to each. In effect, the noise component in the variation of the variables in the series is averaged. This approach is taken by the U.S. Department of Commerce in construction of the composite indices that are reported. For the present discussion, the method of combining series is termed the "stacked" composite approach.

An alternative for forming composites is to use principal components analysis. Principal components analysis is a statistical technique that can be used to create weighted sums of series or indices. These indices or composites, called principal components, are linear combinations of the component data series. The first principal component is the linear combination of the series of several variables, which explains or accounts for the largest share of the total variation of the combined series.

An advantage of the principal components approach for the creation of the composite index to the stacked composites, such as those prepared by the U.S. Department of Commerce, is that the latter do not explicitly take into account interactions among the individual series. For the principal components indices, the weights applied are statistically determined.

The disadvantage of the principal components approach is that it creates an index with weights or loadings on each of the component series determined by the degree to which they are related to each other rather than the extent to which the series are related to the variable to be explained. For example, the first principal component of a series of indicator variables would be exactly the same, regardless of whether it were intended to predict employment growth, employment decline or the weather.

The controlling element in the principal components approach is the selection of the component series for the index. By carefully selecting the component series, the probability of a good fit, with a plausible relationship to the economic variable of interest, can be maximized while avoiding irrelevant variations.

For this project, stacked composite indices of the type used by the U.S. Department of Commerce and principal components indices were created and compared to determine which of the two approaches generated indices that best explained economic growth in counties. The stacked indices have been used in the construction of a number of other submotional indices (see Glennon and Adams 1985; Kozlowski 1977; Loeb 1983; McHugh 1987; and Rufolo 1977). Use of the principal components technique in the creation of composites has not occurred prior to this study.

The method by which the U.S. Department of Commerce combines series to form a composite index is to take a weighted average of the change in each of the series, where the weights reflect the reliability of the series in explaining the economic activity of interest. Prior to summing these series for the composite, each is standardized. This standardization of the individual series is to insure that no component series will dominate on the composite. For example, say one particular series is inherently very volatile, while the others in the composite are stable. The variation in the stacked composite index will be dominated by the most volatile series, if the series are standardized so that the variation is equal for all series.

This basic procedure was used to create the stacked composite index for nontraditional data. The primary difference between the U.S. Department of Commerce technique and the one for this project is that differential weights were not applied to the component series. The procedure used by the Department of Commerce to determine the weights is more complex, and somewhat arbitrary, in the sense that the procedure used to give the component series scores for certain characteristics is subjective. Moreover, in the end, the relative weights appear to make little difference in the overall composite, possibly since the components move together. Also, the weights used for the U.S. Department of Commerce composite indices vary from 0.9 to 1.1. For this project it was assumed that the benefits from the differential weighting of the component series did not justify the cost in terms of specializing the weights. Equal weights were applied to each of the series in forming composite indices.

For both Minnesota and Iowa, a number of composite series were constructed. The series used for these composites were selected based upon their independent correlations with the change in economic activity measured by employment (Tables 1 and 2). Only the results for selected composite indices are reported. In selecting the indices to report, an attempt has been made to present the results from composites that include series for more than one of the nontraditional data sources (banking, telephone, and food stamps).

The composite indices and their component series:

Iowa

CL1 - Percentage Change in Bank Loans (BK_LOAN);
Increases in Business Phone Connections (PH_BUS);
Change in Households Receiving Food Stamps (FS_HOUSE).

CL2 - Increases in Business Phone Connections (PH_BUS);
Change in Households Receiving Food Stamps (FS_HOUSE).

CL3 - Deposit-Liability Ratio of Banks (BK_DEPLIA);
Percentage Change in Bank Loans (BK_LOAN);
Increases in Business Phone Connections (PH_BUS);
Change in Households Receiving Food Stamps (FS_HOUSE).

CL4 - Deposit-Liability Ratios of Banks (BK_DEPLIA);
Percentage Change in Bank Loans (BK_LOAN);
Real Estate/Total Loan Ratio (BK_REALES);
Increases in Business Phone Connections (PH_BUS);
Change in Households Receiving Food Stamps (FS_HOUSE).

CL5 - Deposit-Liability Ratios of Banks (BK_DEPLIA);
Percentage Change in Bank Loans (BK_LOAN);
Charge-offs/Liability Ratio (BK_CHARGE);
Increases in Business Phone Connections (PH_BUS);
Change in Households Receiving Food Stamps (FS_HOUSE).

CL6 - Deposit-Liability Ratios of Banks (BK_DEPLIA);
Percentage Change in Bank Loans (BK_LOAN);
Recoveries/Liability Ratio (BK_RECOV);
Increases in Business Phone Connections (PH_BUS);
Change in Households Receiving Food Stamps (FS_HOUSE).

CL7 - Deposit-Liability Ratios of Banks (BK_DEPLIA);
Percentage Change in Bank Loans (BK_LOAN);
Income-Asset Ratios of Banks (BK_INCASS);
Increases in Business Phone Connections (PH_BUS);
Change in Households Receiving Food Stamps (FS_HOUSE).

CL8 - Deposit-Liability Ratios of Banks (BK_DEPLIA);
Percentage Change in Bank Loans (BK_LOAN);
Income-Asset Ratios of Banks (BK_INCASS);
Percentage Change in Real Estate Loans (BK_REALES);
Percentage Change in Charge-off (BK_CHARGE);
Percentage Change in Recoveries (BK_RECOV);
Change in Households Receiving Food Stamps (FS_HOUSE);
Increases in Business Phone Connections (PH_BUS).

Minnesota

C11 - Recoveries/Liability Ratio (BK_RECOV);
Percentage Change in Interest Income (BK_INT).

C12 - Recoveries/Liability Ratio (BK_RECOV);
Percentage Change in Net Income (BK_INC).

C13 - Recoveries/Liability Ratio (BK_RECOV);
Income-Asset Ratios of Banks (BK_INCASS).

C14 - Recoveries/Liability Ratio (BK_RECOV);
Percentage Change in Net Income (BK_INC);
Percentage Change in Construction and Real Estate Loans BK_REALES);
Percentage Change in Households Receiving Food Stamps (FS_HOUSE)

C15 - Recoveries/Liability Ratio (BK_RECOV);
Percentage Change in Net Income (BK_INC);
Percentage Change in Commercial and Industrial Loans (BK_COMIND);
Percentage Change in Households Receiving Food Stamps (FS_HOUSE).

C16 - Recoveries/Liability Ratio (BK_RECOV) Percent Change in Net Income (BK_INC);
Income-Asset Ratios of banks (BK_INCASS);
Percentage Change in Households Receiving Food Stamps (FS_HOUSE).

C17 - Recoveries/Liability Ratio (BK_RECOV);
Percentage Change in Net Income (BK_INC);
Income-Asset Ratios of Banks (BK_INCASS);
Increases in Residential Phone Connections (PH_RES).

Principal components were estimated using these sets of series, and the associated indicators were developed based on the first principal component.

Evaluating the Stacked Composite Indices

The coefficients of correlation between the composite indices and the percentage change in employment for Iowa and Minnesota are reported in Tables 3 and 4. Because the focus of the project is on the value of nontraditional data for rural areas, the coefficients are reported only for farm counties and nonmetropolitan/nonfarm counties. In general the results were qualitatively the same for metropolitan counties.

Tables 3 and 4 also include report results for assessing these composite indices as leading indicators of economic activity. These coefficients of correlation between the composite indices and selected measures of change in economic activity one period later are reported under the column heading "Leading." For Iowa, the composite indices correlated well with concurrent changes in economic activity, particularly for the farm communities. For each of the eight composites, the series were positively correlated and at a level higher than 1 percent level of statistical significance.

For the nonfarm, nonmetropolitan areas, the results were not as robust, but they still indicated significant correlations between the indicator variables and series selected to measure changes in economic activity. In all instances, the composites were significantly correlated with changes in economic activity showing at least a 10 percent level of statistical confidence. In five of the cases, the level of confidence for the estimated correlation coefficients exceeded 5 percent.

The composite indices also appeared to be good indicators of future changes in economic activity, shown by their correlation with the rate of

Table 3. Correlation between the composite indices and percentage change in employment: Iowa

Stacked composite	Farm Current	Farm Leading	Nonmetro/Nonfarm Current	Nonmetro/Nonfarm Leading
CL1: BK_LOANS, PH_BUS, FS_HOUSE	.104[a] (.001)	.046 (.060)	.067 (.028)	.099 (.001)
CL2: PH-BUS, FS_HOUSE	.194 (.001)	.041 (.107)	.141 (.001)	.001 (.980)
CL3: BK_DEPLIA, BK_LOAN, PH_BUS, FS_HOUSE	.104 (.001)	.041 (.091)	.065 (.031)	.093 (.003)
CL4: BK_DEPLIA, BK_LOAN, BK_REALES, PH_BUS, FS_HOUSE	.094 (.001)	.026 (.282)	.073 (.017)	.093 (.003)
CL5: BK_DEPLIA, BK_LOAN, BK_CHARGE, PH_BUS, FS_HOUSE	.111 (.001)	.044 (.074)	.074 (.015)	.092 (.003)
CL6: BK_DEPLIA, BK_LOAN, BK_RECOV, PH_BUS, FS_HOUSE	.110 (.001)	.037 (.128)	.059 (.053)	.085 (.006)
CL7: BK_DEPLIA, BK_LOAN, BK_INCASS, PH_BUS, FS_HOUSE	.106 (.001)	.033 (.173)	.056 (.067)	.082 (.008)
CL8: BK_DEPLIA, BK_LOAN, BK_INCASS, BK_REALES, BK_CHARGE, BK_RECOV, PH_BUS, FS_HOUSE	.104 (.001)	.017 (.498)	.059 (.052)	.071 (.022)

[a]Figures on top are the coefficients of correlation. Those in parentheses below are levels of statistical significance.

Table 4. Correlation between the composite indices and percentage change in employment: Minnesota

Stacked composite	Farm		Nonmetro/Nonfarm	
	Concurrent	Leading	Concurrent	Leading
C11: BK_RECOV, BK_INT	.073[a] (.131)	-.009 (.852)	.004 (.919)	.002 (.959)
C12: BK_RECOV, BK_INC	.060 (.216)	.023 (.644)	.038 (.299)	.022 (.566)
C13: BK_RECOV, BK_INCASS	.067 (.171)	-.013 (.789)	.001 (.998)	.091 (.017)
C14: BK_RECOV, BK_INC, BK_REALES, FS_HOUSE	.077 (.083)	-.042 (.351)	-.019 (.566)	.067 (.054)
C15: BK_RECOV, BK_INC, BK_COMIND, FS_HOUSE	.060 (.171)	.008 (.860)	.021 (.328)	.017 (.620)
C16: BK_RECOV, BK_INC, BK_INCASS, FS_HOUSE	.094 (.054)	.001 (.977)	.039 (.297)	.035 (.357)
C17: BK_RECOV, BK_INC, BK_INCASS, PH_RES	.023 (.639)	-.053 (.285)	.020 (.593)	.0303 (.434)

[a]Figures on top are the coefficients of correlation. Those in parentheses below are levels of statistical significance.

growth in employment one quarter in the future. The estimated correlations for leading indices in the farm communities were weaker than the concurrent or contemporaneous correlations. However, the correlations were all positive and, in three of the eight cases, statistically significant at the 10 percent confidence level. Had the composites been constructed with the intention of making these correlations better, these results could have been improved.

For the nonmetropolitan, nonfarm areas, the coefficients of correlation between the composites and the future level of economic activity were more pronounced than for the farm counties. In six of the eight cases, the estimated coefficient of correlation was significant at the 1 percent level of confidence. Clearly, the nontraditional indicator series, when combined using this variant of the stacked method, gives consistent signals on likely future trends in economic activity.

The stacked composite indices for Minnesota did not track economic growth measured by percentage total employment change as well as those for Iowa. For the farm counties, two of the composites were correlated with employment growth at confidence levels exceeding 10 percent. For the nonmetropolitan, nonfarm areas, the results were weaker. The "Leading" columns in Tables 3 and 4 show the results of the composites as leading indicators of economic growth. The results mirrored, to an extent, those for Iowa. In the case of the farm counties, the correlations between the composites and economic growth one quarter into the future were weaker than the contemporaneous correlations. However, as in Iowa, the correlations between the stacked composites and the growth one quarter in the future for the nonmetropolitan, nonfarm counties were better than the contemporaneous correlations.

Evaluating the Principal Components Indices

A similar analysis to that for the stacked indices was performed for the principal components indices. Correlations between the first principal components and the employment proxy for the rate of economic growth are reported on in Tables 5 and 6.

Initially, principal components were created from all of the data series considered. For Iowa, the correlations were robust, always significant at the 5 percent level or above. In Minnesota, the correlations were not as strong, particularly for the farm counties. However, in the nonmetropolitan and nonfarm areas, the principal components indices correlated at a better than 10 percent degree of confidence with concurrent growth. And, as with the stacked composite indices, the correlations were even stronger between the employment growth and the first lag of the principal components indices.

In the Iowa counties, the results are very robust. Using the same subsets of series as for the construction of the stacked composites, the correlations between the principal components indices and the growth in employment are generally similar to those for the stacked composite indices. Out of the associated eight principal components, five were significant as coincident indicators at the 5 percent confidence level or better (seven at the 10 percent level or better) for the farm counties. For nonmetropolitan and nonfarm counties, the nontraditional principal

Table 5. Correlation between the principal components and percentage change in employment: Iowa

Principal components[a]	Farm		Nonmetro/Nonfarm	
	Concurrent	Leading	Concurrent	Leading
PC1 All items	.043[b] (.072)	.052 (.030)	.061 (.044)	.064 (.033)
PC2 BK_LOANS, PH_BUS, FS_HOUSE	.082 (.001)	.041 (.088)	.084 (.006)	.081 (.008)
PC3 PH_BUS, FS_HOUSE	.163 (.001)	.056 (.020)	.130 (.001)	.000 (.993
PC4 BK_DEPLIA, PH_BUS, FS_HOUSE	.185 (.001)	.012 (.626)	.161 (.001)	.007 (.822)
PC5 BK_DEPLIA, BK_LOAN, FS_HOUSE, PH_BUS	.095 (.001)	.039 (.100)	.105 (.001)	.077 (.012)
PC6 BK_DEPLIA, BK_LOAN, BK_REALES, FS_HOUSE, PH_BUS	.052 (.032)	.007 (.774)	.106 (.001)	.072 (.012)
PC7 BK_DEPLIA, BK_LOAN, BK_RECOV, FS_HOUSE, PH_BUS	.059 (.015)	.058 (.016)	.070 (.021)	.059 (.051)
PC8 BK_DEPLIA, BK_LOAN, BK_INCASS, FS_HOUSE, PH_BUS	.029 (.237)	.019 (.431)	.035 (.249)	.038 (.215)
PC9 BK_DEPLIA, BK_LOAN, BK_INCASS, BK_REALES, BK_RECOV, FS_HOUSE, PH_BUS	.047 (.051)	.061 (.012)	.069 (.023)	.073 (.016)

[a]Variables included are noted; see text and discussion of stacked indices for definitions.

[b]Figures on top are the coefficients of correlation. Those in parentheses below are levels of statistical significance.

Table 6. Correlation between the principal components and percentage change in employment: Minnesota

Principal components[a]	Farm Concurrent	Farm Leading	Nonmetro/Nonfarm Concurrent	Nonmetro/Nonfarm Leading
PC1 All items	.019[b] (.667)	.032 (.457)	.059 (.082)	.081 (.015)
PC2 BK_RECOV, BK_INT	.045 (.304)	.030 (.485)	.032 (.336)	.047 (.158)
PC3 BK_RECOV, BK_INC	-.028 (.525)	.059 (.177)	.034 (.302)	.101 (.002)
PC4 BK_RECOV, BK_INCASS	-.042 (.337)	.120 (.006)	.040 (.236)	.070 (.037)
PC5 BK_RECOV, BK_INC, BK_REALES, FS_HOUSE	.034 (.439)	.054 (.212)	.018 (.581)	.108 (.001)
PC6 BK_RECOV, BK_INC, BK_COMIND, FS_HOUSE	.023 (.600)	.063 (.148)	.040 (.234)	.112 (.001)
PC7 BK_RECOV, BK_INC, BK_INCASS, FS_HOUSE	-.035 (.429)	.116 (.007)	.040 (.226)	.073 (.030)
PC8 BK_RECOV, BK_INC, BK_INCASS, PH_RES	-.032 (.458)	.118 (.007)	.043 (.199)	.073 (.029)

[a]Variables included are noted; see text and discussion of stacked indices for definitions.

[b]Figures on top are the coefficients of correlation. Those in parentheses below are levels of statistical significance.

components indicators were significant at the 5 percent level or higher in seven out of eight cases. As leading indicators, these principal components indices did not perform as well, but only marginally less well in farm and nonmetropolitan/nonfarm counties.

As for the stacked composite indices, the principal components-based indicator series performed less well in Minnesota than in Iowa. One surprising result was that in the nonmetropolitan and nonfarm areas, the principal components indices performed better as leading indicators than as coincident indicators. Comparing the stacked composite indicators to the principal components indicators, there was in most cases little difference, except for the dominance of the principal components indices leading indicators for nonmetropolitan and nonfarm counties in Minnesota.

Conclusion

In the analysis of the dynamics of economic growth, economists and other public policy analysts typically rely on readily available data, collected and reported by the federal government. A potential problem with strict reliance upon these federally provided data is that their geographic coverage may be less than complete. This can be a particularly limiting aspect of the federal data for studying of rural areas and smaller communities. The federal data often are not available or are available with considerable lag for these areas. Given these lags in the official (traditional) federal data, they often are not useful in providing indications of economic trouble spots in states, and especially in providing early warnings of problems.

A wealth of untapped data exists that can be utilized to address some of the inadequacies and buttress the traditional federal data sources. These nontraditional sources may give a more accurate, more complete, and more timely picture of conditions in the local economy. The project has demonstrated the potential for nontraditional data--these data exist, can be obtained, and have economic content. The results of this initial attempt to acquire and use nontraditional data to track economic conditions at the county level show they can be a useful addition to the data sources currently used in economic development planning and programming.

The statistical significance of the relationship between the indicator variables and the traditional measures of economic growth is adequate, in light of the fact that the stacked composite and the principal components indices used were all constructed without reference to national aggregate economic factors, which in the end drive much of the performance of regional economies. The nontraditional indicator variables were county-specific, industry-specific series. A composite including these national economic variables as well as the local series would have had much higher explanatory power for the percentage change in county employment. If accurate forecasts had been the only intent of this project, higher composite correlates utilizing aggregate national data could have been constructed.

One reason for the relatively low degree of statistical correlation between individual nontraditional data series and the level of economic activity is that the data series were not developed--nor are they strictly maintained at this point--for use as indicators of local economic activity. Once monitoring economic activity as an objective has been identified, it is possible that the quality of the nontraditional data for this use may be improved. Even in rough form, the nontraditional data have value for monitoring economic activity in rural areas; they exhibit a potential for use in providing rapid feedback on economic impacts of development assistance programs.

Many of the data sources that could prove useful in this exercise were excluded because of administrative difficulties. These problems were mainly transitional, for example, obtaining first-time permission to use. The lags in the delivery of these and other nontraditional data can surely be shortened as experience with their use grows. As additional data become available for inclusion as nontraditional indicators, the breadth of the coverage of sectors will grow beyond the current experiment. Retail sales from sales tax data, housing permits, and unemployment insurance claims are examples of series used in some composite indices, although not included in these experimental indices. The potential of nontraditional databases is evident and should be ultimately exploited. As additional nontraditional data sources are identified and experience in the collection of those data leads to more timely delivery, nontraditional data series are likely to become an increasingly important element of the information base for analysis of economic development in rural areas.

Finally, the indicators developed were not intended as substitutes for the traditional federal data. Rather, the nontraditional data can be viewed as supplemental information on local economic development. Each piece of information, traditional or nontraditional, carries a unique signal with its own leading, coincident, and lagging relationship to overall economic health of a community. Composites consisting of the traditional as well as the nontraditional data are likely to provide a much broader signal or set of signals for use in designing and monitoring economic development policy.

Appendix
Excluded Variables

A detailing of the reasons for the exclusion of some of the potential economic indicator variables is presented below. These specifics are provided to suggest the problems of acquiring series that can be used as nontraditional indicators.

New Incorporations. There are questions concerning the possibility of releasing this data at the level of geographic detail needed for this project. Moreover, there are conceptual difficulties. An incorporation does not necessarily imply that a business is in operation. It simply indicates a potential to begin business. The relationship between the place of incorporation and place of operation is not necessarily coincident. Finally, decisions to incorporate rather than choose other forms of business organization are often influenced by legal and national tax factors.

New Rural Electric Connections. At this point, the data are compiled annually by a state agencies, for example, the Iowa Commerce Commission. Monthly data are made available at that time. In the absence of a more regular compilation, the data are not timely enough.

Retail Sales. Departments of revenue and finance can provide these data quarterly at the county level of detail. However, the time lag is ten months for Iowa.

Rural Electricity Usage. Like the rural electrical connections data, these data are available on a monthly basis, but once per year.

Vehicle Miles of Travel. The time lag for the delivery of these data on a county basis can be as high as six months in Iowa.

Average Hourly Earnings. These data are not available at the county level of detail.

Property Valuation. These data are available only on a fiscal year basis.

Residential Building Permits. These data are available only for reporting towns and are thus not sufficiently comprehensive. The time lag for delivery is about four to six weeks.