United States Department of Agriculture

ERS

Technical Bulletin Number 1850

An Economic Research Service Report

Factors Associated with Rural Economic Growth

Lessons from the 1980's

Lorin D. Kusmin
John M. Redman
David W. Sears

It's Easy To Order Another Copy!

Just dial 1-800-999-6779. Toll free in the United States and Canada. Other areas, please call 1-703-834-0125.

Ask for ***Factors Associated with Rural Economic Growth: Lessons from the 1980's*** (TB-1850).

The cost is $15.00 per copy. For non-U.S. addresses (including Canada), add 25 percent. Charge your purchase to your VISA or MasterCard. Or send a check (made payable to ERS-NASS) to:

ERS-NASS
341 Victory Drive
Herndon, VA 20170-5217

For additional information about ERS publications, databases, and other products, both paper and electronic, visit the **ERS Home Page** on the Internet at http://www.econ.ag.gov/

Factors Associated with Rural Economic Growth: Lessons from the 1980's. By Lorin D. Kusmin, John M. Redman, and David W. Sears. Rural Economy Division, Economic Research Service, U.S. Department of Agriculture. Technical Bulletin No. 1850.

Abstract

This report identifies characteristics of rural areas conducive to economic growth. Its analysis explains growth in total real earnings in nonmetro U.S. counties from 1979 to 1989 using multiple regression analysis. Factors associated with improved county earnings were attractiveness to retirees, right-to-work laws, high levels of high school completion rates, public education expenditures, and access to transportation networks. Factors associated with poor earnings growth include higher wage levels and concentrations of transfer-payment recipiency. Counties with higher concentrations of African-Americans also experienced slowed earnings growth, although the reasons for that association cannot be identified from this analysis. The mix of industries active in a county and the rate of earnings growth in nearby counties were also strongly associated with county earnings.

Keywords: Counties, rural, nonmetropolitan, regional, economic growth, earnings, economic development, commuting zones, fixed effects, regression analysis

Washington, DC 20005-4788 September 1996

Acknowledgments

The comments and assistance of many people in and outside of the Economic Research Service (ERS) have contributed to this project. Among those whom the authors would like to acknowledge are:

- Molly Sizer, formerly of ERS, who played a central role in the early phases of this report, and who is one of the developers of the commuting zone classification that plays a major part in this report;
- David Hopkins of the Information Services Division at ERS, who compiled much of the data used here in a source file, and who provided additional programming assistance over the course of the project;
- Calvin Beale, Peggy Cook, Robert Hoppe, Richard Reeder, Peter Stenberg, and Paul Swaim of ERS, as well as Thomas Johnson of the Virginia Polytechnic Institute, Glenn Nelson of the University of Missouri, and Sheena McConnell of Mathematica for their comments on the manuscript;
- Linda Atkinson of the ERS Information Services Division for her assistance with statistical and programming problems during the course of the project, comments on technical portions of the manuscript, and assistance with graphics;
- Sharon Lee and Tom McDonald of the ERS Information Services Division for their editorial assistance in preparation of the manuscript;
- Charlie Hallahan and Pete Kostik of the ERS Information Services Division for additional assistance with statistical and programming problems;
- David McGranahan of ERS, and Jim McGlone and Felice Marlor, formerly of ERS, who provided us with unpublished data developed during research of their own;
- Cathy Kassab, formerly of the General Accounting Office, who provided assistance and citations on the subjects of robust regression and bootstrapping;
- Bill Amt, formerly of ERS, for assistance with preparation of tables during an earlier phase of the project;
- Several attendees at the 1991 meetings of the Regional Science Association in New Orleans, who made useful comments on a closely related report; and
- Jerry Hausman of the Massachusetts Institute of Technology, for guidance on the assessment of the random effects estimator.

Contents

Highlights

This report uncovers evidence on characteristics of rural areas conducive to economic growth. It explains growth in total real earnings by place of work in nonmetro U.S. counties from 1979 to 1989 as a function of local and regional characteristics at the beginning of the period. Variables selected for analysis include demographic characteristics, labor market characteristics, education levels and activity, transportation access, business and banking structure, amenities, relationship to metro areas, and economic base.

- Earnings in retirement counties grew 4.5 percentage points more than earnings in other counties over the 1979-89 period.

- The percentage of the county population that was African-American was inversely associated with earnings growth. A 10-percentage-point difference in the African-American share of the county population was associated with a 1.9-percentage-point difference in cumulative 1979-89 earnings growth.

- Wage levels had a significant and substantial negative effect on earnings growth rates. A 10-percent difference in earnings per job was associated with a 2.35-percent difference in total earnings growth over 10 years.

- State right-to-work laws had a significant positive effect on earnings growth. The estimated effect of these laws was a 5.2-percent difference in earnings growth during the 1980's.

- Education levels were positively related to economic growth: A difference of 10 percentage points in the high school completion rate among adults was associated with a difference of 3.3 percent in total 1979-89 earnings growth.

- Public education expenditures are conducive to higher earnings growth. A difference of $1,000 in annual per-pupil expenditures was associated with an additional 3.8 percentage points in growth between 1979 and 1989.

- Counties that had an airport with scheduled passenger service within 50 miles experienced 3.4 percentage points in additional earnings growth during the decade.

- Access to interstate highway interchanges contributed to earnings growth in rural areas during the 1980's. Each such interchange within a county was associated with approximately 0.42 percentage point in additional growth during the period.

- The proportion of goods-producing businesses that are small and independent had a negative relationship with earnings growth during the 1980's. A county could expect a reduction of 1.1 percentage points in earnings growth over the decade if 80 percent rather than 70 percent of all county goods-producing businesses were small independent businesses.

- An additional $100 in transfer payments per capita was associated with a 1.6-percentage-point reduction in cumulative earnings growth.

- Industry structure was an important determinant of county earnings growth. Among the industrial sectors with positive and significant associations with earnings growth were

transport services, real estate, hotels, miscellaneous business services, education services, and State and local governments. Among those with significant negative effects were forestry, metal mining, oil and gas extraction, coal mining, heavy construction, lumber and wood products, primary metal manufacturing, electrical machinery manufacturing, and railroads.

- These results reflect application of a multivariate regression model of earnings growth with corrections for heteroskedasticity, non-normal errors, and fixed commuting-zone effects, which we refer to as our "preferred method." Results are also presented that reflect omission of some or all of these corrections.

- Some variables yielded little or no evidence of a significant relationship with earnings growth. These variables include urban population, population aged 25 to 64, labor force participation, college completion rate, high school dropout rate, local tax level, presence of an airport or an intersection of two major highways within the county, liberal branch banking laws, topography, and population of nearby metro areas.

- Past growth rates had a very modest effect on 1979-89 growth rates.

Factors Associated with Rural Economic Growth: Lessons from the 1980's

Lorin D. Kusmin
John M. Redman
David W. Sears

Introduction

This report provides policymakers, practitioners, evaluators, and researchers in the field of subnational economic growth with evidence of which characteristics of rural areas are conducive to economic growth.[1] We considered a wide range of factors that may be associated with growth. While the patterns found here apply to economic growth in the 1980's, many of them may persist into the 1990's and beyond.

Policymakers and practitioners will be particularly interested in our findings concerning the effect of variables that may be determined or heavily influenced by government action, such as educational spending levels or access to transportation infrastructure. Researchers and evaluators will also be interested in our findings concerning other local characteristics that have strong associations with rural economic growth. Even if these cannot be readily changed by public action, it is important to take into account the effects of these background or control variables when one is assessing the economic effect of particular policies or programs.[2]

Basic Issues of Modeling Local Economic Growth

There is an extensive econometric literature on factors that may be associated with subnational economic growth (see Kusmin, 1994). Many of these studies focus on one or a few potential influences on economic growth (taxes, for example, or education) while disregarding others. Most seek to explain differences in economic activity across States or metropolitan areas, and a few focus on rural areas.

The model developed in this report explains growth in earnings in rural (nonmetro) U.S. counties from 1979 to 1989 as a function of local and regional characteristics at the beginning of the

1. Strictly speaking, we refer to "relative economic growth." More than half of the rural counties covered in this report experienced declines in total real earnings during the study period. Thus, some factors that we identify as facilitating growth may have operated primarily by retarding or arresting the decline of county economies, and different factors might be more relevant in a period of widespread rural economic growth.

2. In this report, "regional" is used to refer to States, combinations of States, and multicounty areas. "Local" is used to refer to single counties and to individual communities; it is also used in reference to local government as conventionally defined. "Subnational" is used to refer to geographic areas within the United States that may be compared with one another, thus "subnational economies" may be either "regional" or "local."

period.[3] We used multivariate linear regression to estimate the relationship between growth and each of a variety of local and regional characteristics included in the model. Both 1979 and 1989 represent peaks in the national business cycle. Hence, differences among counties in earnings growth rates calculated over this period should represent long-term trends and not intercounty differences in cyclical sensitivity. To assess the sensitivity of our results to the period of growth considered, we also estimated our econometric model for 1978-88 and for 1980-90. (These results are available from the authors.)

While many past studies of subnational economic growth have attempted to explain employment growth rather than earnings growth, policymakers in recent years have increasingly emphasized the importance of attracting higher wage jobs, rather than just any jobs. At the same time, we were interested in the economic well-being of rural communities as a whole; hence, we do not focus on earnings per job or earnings per capita, indicators that may rise even when a community as a whole is in decline. Therefore, in this report we focus on total county earnings growth as the single measure best reflecting the overall experience of the local economy. (References to "earnings growth" in this report refer to real earnings growth, after adjustment for inflation. The measurement of real earnings growth is discussed on pages 6-7.)

One of the key dilemmas debated in the rural development literature is whether development should be aimed at improving the condition of rural people or rural places. Our choice of earnings growth as our key indicator reflects the consensus view that both people and places are important.

The Structure of a Model of Local Economic Growth

What kinds of variables influence changes in the level of local economic activity? That is, what causes economic growth or decline? In simple economic models where the economy is assumed to be always in equilibrium, the current characteristics of a local area (and the region surrounding it) are already reflected in the area's present level of economic activity. Hence, those characteristics should not be useful in explaining the direction or magnitude of future economic growth. In such a world, changes in economic activity levels can be explained only by changes in other characteristics.

However, as noted in some earlier studies (Plaut and Pluta, 1983; Newman and Sullivan, 1988), a model that links the initial levels of various local and regional characteristics to subsequent growth may be justified if initial levels of output are not equilibrium levels, and if areas with particular characteristics are more likely to be below/above their equilibrium level of output initially. But, as also pointed out by Newman and Sullivan, equilibrium levels of output will depend on both supply and demand factors, so that appropriate implementation of this approach requires independent measurement of both.

Even if local economies were initially at or near economic equilibrium, a model that relates the initial levels of local and regional characteristics to subsequent change in economic activity may still be justified if national or global economic changes (including technological changes) altered the

---

3. Nonmetro counties are those that are not part of any Metropolitan Statistical Area (MSA), Consolidated Metropolitan Statistical Area (CMSA), or New England County Metropolitan Area (NECMA), while metro counties are those that do belong to an MSA, CMSA, or NECMA. Metropolitan areas (MSA's, CMSA's, and NECMA's) are defined by the Office of Management and Budget using census data. These areas normally include a population nucleus (a city or urbanized area) of 50,000 or more, the county or counties containing that nucleus, and adjacent counties that meet specified requirements for commuting to the central counties and for metropolitan character. In New England, MSA's and CMSA's are composed of towns and cities rather than counties. Thus, county-level analysis of metro and nonmetro areas in New England must use the NECMA's instead (U.S. Department of Commerce, 1991).

demand for particular products, types of labor, or other local area characteristics during the period, and local economic activity subsequently rose or fell toward the new equilibrium.[4] Few would deny that extensive global realignment occurred during the 1980's. We believe that this realignment is sufficient by itself to justify use of a model in which the level of local characteristics is used to explain economic growth and decline.

It is also likely that changes in local and regional characteristics during the 1979-89 period affected equilibrium economic activity levels and therefore growth during that period. However, growth during the period may also have caused many changes in local characteristics. To reduce the risk of wrongly identifying the results of growth as the determinants of growth, contemporaneous change variables are excluded as regressors in the model. (However, in one alternative model, we do include the contemporaneous earnings growth rate in other nearby counties as a control variable.)

Even with this design, there is still some risk that we might wrongly identify characteristics that were actually results of growth (or decline) as having been causes of growth (or decline). If some characteristics of counties as measured between 1975 and 1980 were consequences of earlier growth, and if the factors giving rise to that earlier growth persisted into the 1980's, then these variables may appear to have been causes of growth in the 1980's, when they were really consequences of the earlier growth. However, the low correlation of 1970's nonmetro earnings growth rates and 1980's rates, documented in *Growth and Stability of Rural Economies in the 1980's: Differences Among Counties* (Sears and others, December 1992), suggests that such persistent growth factors have played at most a modest role in nonmetro county earnings growth. Further, tests of models that include lagged county growth rates (see pp.25 and B-11) indicate that earnings growth in the 1970's had little if any influence on growth in the 1980's. Hence, any association between our independent variables and growth in the 1970's is unlikely to bias our estimates of the effects that these variables had on growth in the 1980's.

The model presented in this report is a cross-sectional model. That is, for each case (county), we observed the earnings growth rate over a single period, and associated differences in the growth rate across counties with differences in county characteristics.

An alternative approach to identifying the determinants of earnings growth employs panel data. Panel data are data collected on a group of individuals or jurisdictions at a number of different points in time, so that there are multiple observations in the data set for each individual or jurisdiction. A panel data approach would control for all differences across counties in characteristics that were constant over time, and would test whether past changes in local or regional characteristics have been associated with changes in local area growth rates from one period to the next. The panel data approach permits estimation of the effect of county characteristics only if: (a) those characteristics have changed over time, and (b) data on those characteristics are available for each period.

---

4. One should note that if local economic growth or decline during the period reflected adjustment to such a change in national or global demand, and if that adjustment was complete by the end of the period, then there is no reason to expect that future growth will be associated with the same characteristics. However, if levels of economic activity had only partially adjusted to changes in demand by the end of the study period, then growth trends in the 1980's may continue into the 1990's, as the adjustments to those demand changes continue. Further, if trends during the 1980's in the demand for particular factors, outputs, or local/regional characteristics continue into the 1990's, then the associations between local or regional characteristics and economic growth that were a consequence of those trends will also continue. Therefore, the extent to which any of the relationships that are found in this report persist through the 1990's will depend on: (a) the speed of local adjustment to demand changes, and (b) the persistence of 1980's demand trends into the 1990's.

This approach is used in some studies of the determinants of State economic growth (see, for example, some of the studies reviewed in Kusmin, 1994), but panel data for many of the variables of interest to us were not available at the county level. In addition, proper implementation of a panel data approach would require careful consideration of any probable lags in the effect of changes in local or regional characteristics on local growth rates. Determination of the appropriate lags would be difficult in many instances.

A few of the studies covered in our review of the literature (Kusmin, 1994) employ a simultaneous equations approach to the modeling of subnational economic growth (see, for example, Fox, 1981, and Carlino and Mills, 1987). These studies treat population growth and migration flows as factors that exert their own influence on economic growth, as well as being influenced in turn by economic growth trends. They model the determination of both population and economic trends, as well as the influence of each upon the other. However, the more common approach in the literature is to ignore this simultaneity and concentrate on a single regression equation explaining economic outcomes. This is the approach taken in the present report. This may be viewed as a reduced-form approach, where the resulting estimate of the effect of any county characteristic on economic growth will reflect both: (a) the characteristic's direct effect on economic activity, and (b) its indirect effect on economic activity through its effect on population.

A Linear Model

In this report, we assume a linear relationship between each local or regional characteristic and our index of growth in most instances. That is, each increment of one unit in the level of that characteristic has the same effect on the growth index, without regard to the initial level of that characteristic or other characteristics, or the initial level of the growth index itself. This assumption is made to simplify the analysis in a model with many independent variables.

Note, however, that the dependent variable in our model is a change in the logarithm of earnings (see p. 6). Hence, the model might instead be described as a semi-logarithmic one. However, the assumed relationships between the dependent variable, as defined, and each of the dependent variables as they have been defined are linear relationships. This permits use of conventional linear regression techniques for estimation. (It should also be noted that, because the dependent variable is logarithmic, while most of the independent variables are not defined in logarithmic terms, use of this model to generate expected growth rates can yield improbably large expected rates for combinations of values of the independent variables that are very different from those actually observed.)

In a few instances, we also use an alternative form for the regressor. For example, the average wage level is entered in logarithmic form. Local college enrollment levels are converted into a dichotomous (dummy) variable (see Appendix B, p. 6); and data on activity at local or nearby airports is converted into dummy variables indicating the presence of such airports. (Other variables that represent the conversion of an underlying continuous variable into dichotomous form are the dummy variable for retirement attraction counties and the dummy variable for the presence of the intersection of two interstate highways.)

Geographical and Industrial Coverage

Most of the past empirical analyses of subnational economic growth reviewed for this report focused on a particular sector of the economy (for example, manufacturing), or on a particular geographic area or collection of areas (for example, one State or selected metropolitan areas). A focus on a particular sector permits more attention to local or regional characteristics presumed to have a specific effect on that sector. Similarly, a focus on a single area or a limited number of

geographic units (for example, a sample of MSA's) can permit development or use of data on variables for which published or readily available sources with nationwide coverage do not exist.

However, such narrow coverage also limits the possibility of generalization; further, when a small number of cases is included in the study, the number of local or regional characteristics that can be included in the model is limited as well. For these reasons, the coverage of this report is inclusive both geographically and sectorally, seeking to explain the 1979-89 growth in total earnings across all rural counties throughout the United States.

Units of Analysis

This report uses county-level data, in part because counties are the smallest units for which a wide variety of economic and geographic data are available nationwide.

It has been suggested that individual counties do not constitute distinct substate economies, and that a larger unit of analysis is more appropriate for the study of local economic growth (Killian and Parker, 1991). However, the results of our analysis confirm that earnings growth in an individual county within a multicounty labor market area can often differ sharply from the performance of that area as a whole, and that such differences can be explained, at least in part, as a function of the characteristics of that particular county. This indicates that while there may be important interactions among counties within such a labor market area, those areas cannot be viewed as homogeneous economic units.

Further, use of the individual county as a unit of analysis allows us to use regional (multicounty) groupings to control for unmeasured regional effects (see pp. 24-25). This would not be feasible if such groups of counties were treated as the basic units of analysis.

Finally, performing the analysis at the county level, rather than at a more aggregated level, permits a report that covers the rural universe neatly by covering essentially all nonmetro areas while excluding all metro areas.[5] In contrast, the exclusively nonmetro commuting zones studied by Killian and Parker (1991) cover only 1,572 of 2,360 nonmetro counties, and only 57 percent of the 1980 nonmetro population. The remaining 788 nonmetro counties are in commuting zones that also include metropolitan counties (in most cases, the population of these zones was predominantly metropolitan.)

5. In New England, where the boundaries of Metropolitan Statistical Areas (MSA's) follow town rather than county boundaries, we cover only those counties that do not fall within the boundaries of any New England County Metropolitan Area (NECMA). Thus, our analysis excludes some nonmetro towns in New England.

Variables That May Be Associated with Local Economic Growth

This report is based on an analysis of data on 2,346 rural counties in 47 States.[6] Data sources are listed in Appendix A.

Measure of Growth

Local economic growth is represented in this report by the change in county earned income by place of work during the 1980's. The data are taken from the Bureau of Economic Analysis county income data file. As noted on page 2 and in our previous report (Sears and others, 1992), this is a broader measure of change in economic well-being than employment growth, because it reflects changes in compensation per job as well as changes in the number of jobs.[7] More precisely, the measure of county economic growth used is:

$$100 * [\log(1989\ RLPI) - \log(1979\ RLPI)],$$

where RLPI is real labor and proprietor income by place of work.

For small changes, this index is approximately equal to the percentage change in real income over the 1979-89 period.[8] For larger changes, the two statistics diverge to some extent. A 25-percent decrease in earnings corresponds to a change of 28.77 units in the index, and a 50-percent decrease corresponds to a change of 69.31 units. On the other hand, a 25-percent increase in earnings corresponds to a change in the index of just 22.31 units; a 50-percent increase corresponds to a change of just 40.55 units. Of the 2,346 counties covered in this report, 100 experienced earnings growth of 50 percent or more in 1979-89, while 18 experienced decreases of 50 percent or more. For these counties, the differences between the earnings growth index and the percentage change in earnings are substantial.

There are several reasons for using the change of the logarithm of earnings ($\Delta\log$(total earnings)) rather than the percentage change:

1) If we assume that $\Delta\log$(earnings per job) and $\Delta\log$(number of jobs) are each linear functions of our independent variables, then $\Delta\log$(total earnings) will also be a linear function of these variables (because $\Delta\log(A*B) = \Delta\log(A) + \Delta\log(B)$). An analogous relationship does not hold for $\%\Delta$(total earnings), $\%\Delta$(earnings per job), and $\%\Delta$(number of jobs).

2) If the influence of distinct county characteristics on the level of economic activity is multiplicative rather than additive, and if we wish to estimate these effects using linear regression, then use of the change in logarithm may be preferable. For instance, if earnings (Y) are determined by independent variables X and Z, and if

$$Y_{1989}/Y_{1979} = e^{a} * e^{bX} * e^{cZ} * e^{r},$$

---

6. Alaska and Hawaii are not included because of data limitations; New Jersey has no nonmetropolitan counties; and 11 counties in the remaining States were excluded due to missing data for some key variables.

7. While there is nothing wrong with a place where people live well while earning a good living in another county, we chose to look at how well places did in providing earnings to their workers.

8. The factor of 100 in the definition of the dependent variable has no substantive meaning, but permits us to interpret the results in terms of effects on percentage growth.

where a is a constant and r is an error term, then

$$\log(Y_{1989}) - \log(Y_{1979}) = a + bX + cZ + r,$$

and the coefficients of the latter equation can be estimated using linear regression.

3) Use of Δlog (total earnings) somewhat reduces the influence of the small number of counties that experienced large unexplained growth in earnings.

Values of this growth index vary widely, ranging from -147.8 to 197.0, with a mean of -1.42 and a standard deviation of 25.85.[9]

Note that the above -1.42 mean value was not the growth index for the U.S. nonmetro economy as a whole. Each county has been given equal weight in computing the statistic shown here. In contrast, the growth index for the nonmetro United States as a whole gives greater weight to the experience of those counties with larger economies. The value of this aggregate growth index for the 1979-89 period was 3.09 for the 2,346 counties included in this report. Total real earnings in these counties rose by a cumulative total of 3.14 percent over this 10-year period.

For simplicity, changes in our index of earnings growth will be referred to as "percentage" changes in earnings in the remainder of this report.

Potential Determinants of Growth

The independent variables in our model of county earnings growth represent a variety of county and regional characteristics as measured at or near the beginning of the 1979-89 period. To the extent that data were available, variables in the model include most of those factors that past researchers have identified as likely influences on regional or local economic growth.

Demographic Variables

Urban Population

Among the independent variables in the model is the total number of people in the county living in areas defined as urban.[10] This reflects the example of past studies, which have suggested both population and percentage urban as potential determinants of subnational economic growth. (See Kusmin (1994) for a more detailed discussion of the classes of variables that have been considered as possible determinants of subnational economic growth in past reports. That report covers most classes of variables discussed in this report, and includes citations of individual studies.)

To the extent that a report's units of analysis correspond to natural economic units, population may be viewed as a measure of potential agglomeration economies, which would imply an expected positive association between population and economic growth rates. Population may also be included to capture some scale effects not represented by other variables. Several studies have included measures of total or percentage urban population among potential predictors of subnational economic activity. A variety of reasons have been offered for thinking that more urban

9. These are values for the 2,346 counties included in the regression analysis.

10. Generally, these are incorporated or unincorporated places with a population of 2,500 or more.

areas might experience either slower or faster economic growth, ranging from more highly developed transportation systems to differences in tastes between urban and rural consumers.

In this report, we focus on total urban population because: (a) any agglomeration economies are likely to be associated with urban populations rather than rural ones; and (b) most of the likely effects of urban populations seem more likely to be associated with the size of the urban population, rather than with the percentage of the population that was urban, which may reflect the density of settlement in rural parts of the county as much as the degree of urban settlement.

Racial/Ethnic Mix

Some authors have suggested that the percentage of racial minorities in an area might have an influence on changes in business activity, perhaps because of differences in their education, training, or wages. Both positive and negative effects on economic growth have been hypothesized, and some evidence of each has been reported (Kusmin, 1994, pp. 56-57). One possibility is that the relative size of the African-American population might have reduced growth if geographically mobile employers avoided locating businesses in areas with concentrations of racial minorities. Alternatively, growth might have been positively associated with African-American population concentrations if the intensity of this employer behavior fell over the 1979-89 period. Growth might also have been lower in areas with large African-American populations, because African-Americans have historically been less likely to start new businesses, and such new businesses were a major source of new jobs during the 1980's. Similar hypotheses could be constructed for other minority groups, such as Hispanics. Because of these suggestions, the percentage of county residents who are African-American and the percentage who are Hispanic have also been included in our model.[11]

Retirement Attraction

Also included in the model is a retirement county dummy variable. Retirement counties experienced substantial net in-migration of people at or near retirement age (Bender and others, 1985). Specifically, retirement counties are counties where the 1980 population aged 60 and over was at least 15 percent higher than the level expected from survival of the local 1970 population aged 50 and over. If these counties continued to attract retirees during the 1980's, retirement counties are likely to have experienced faster earnings growth, associated with demand for goods and services generated by in-migration. In addition, retirement counties are likely to be those with amenities that may also attract other relocating businesses and individuals. One report, which included a similar variable among the potential determinants of manufacturing employment growth, found a highly significant positive relationship.

Other Demographic Variables

Based on suggestions made by past researchers who included measures of an area's population age distribution, we include the percentage of the total population in the 25 to 64 age range as one measure of the availability of an experienced labor force. We hypothesize that higher levels of this variable may be positively associated with economic growth.

---

11. Because our model controls for educational attainment and wage levels (see pp. 9-11), lower wage rates and school completion rates associated with African-American or Hispanic populations should not influence our estimate of these ethnicity effects.

Labor Market Characteristics

Wages

If economic growth is primarily driven by business location and expansion/contraction decisions, and if those decisions are motivated by the desire to hold down production costs, then higher wage levels (for labor of a given quality) will result in relative decreases in business activity.

However, this is not the only possible effect of wage levels on economic growth. To the extent that changes in business activity are driven by individual migration decisions, and those decisions are motivated by the desire for higher wages, then higher wages could be associated with an increase in business activity.

Further, even if changes in business activity are driven by the search for the lowest production costs, empirical studies that do not adjust labor cost measures for the quality of labor may fail to find the expected negative relationship between wages and changes in business activity. Without such adjustment, high wages may act as a proxy for high labor quality, which is expected to have a positive effect on business activity.

Wages may be an important influence on subnational economic growth. A number of earlier studies included some wage measure as a predictor of business activity. Some of these studies found a significant negative relationship between wages and business activity; however, others found a significant positive relationship, at least for some sectors of the economy or some measures of activity. Still others found no clear or consistent relationship (Kusmin, 1994, pp. 36-39).

The county wage level measured in this report is the 1976-78 average value of earnings per job.[12] Differences in the average level of earnings per job reflect differences in job and worker characteristics as well as variation in the wage scale. However, the inclusion of separate controls for many of these job and worker characteristics in the model implies that any estimated effect of earnings per job in our results should primarily reflect the effect of true wage differences.

Union Power

It is widely suggested that many employers who can choose their location will seek to avoid areas where unions are powerful, reducing expected economic activity and growth in those areas (see articles reviewed in Bartik, 1991, and in Kusmin, 1994, pp. 39-42). In addition to the higher wages that unions command, it is thought that business will seek to avoid both the limitations on management discretion imposed by union agreements and the prolabor State laws and regulatory policies that may be associated with union political strength. Further, if union agreements and prolabor public policies do in fact make businesses less competitive, then businesses in nationally or globally competitive markets will grow more slowly in areas of greater union power, even when

12. The 1979 data are not used to compute the measure of initial county wage level, because they have been used to compute the dependent variable (1979-89 earnings growth). Use of the same data to compute values for an independent variable would lead to a negative bias in the estimated relationship between the two variables. This follows because any perturbations in 1979 county earnings values that might result from either measurement errors or random factors would have opposite effects on 1979 earnings-per-job and on the 1979-89 growth index. For instance, in areas where 1979 income was under-reported, the data would understate true 1979 earnings per job, while overstating 1979-89 earnings growth. This would lead to an apparent negative association between initial earnings per job and subsequent earnings growth, even in the absence of any true relationship.

such power does not lead individual businesses to move. However, some researchers also propose that unions may play a role in enhancing productivity (Economic Policy Institute, 1992).[13]

The presence of State right-to-work laws is included in this report as one measure of union power. (Right-to-work laws prohibit labor agreements that require union membership as a condition of continued employment of a person who was not a member when hired.) Unionization levels are not included, however, because data were available only at the State level, and State averages may have only limited relevance to nonmetro labor markets. In contrast, because right-to-work laws are applicable statewide, they would have affected union power in all counties.

Labor Force Participation

The model also includes a measure of labor force participation, as suggested by some other researchers. The variable included is the ratio of the labor force to the working age (25 to 64) population. This is a measure of the size of the potential labor force that was available to new or expanding employers. A high ratio means that most residents of working age were already in the labor force, so that there was less opportunity for employers to attract more people into the labor force without in-migration. Therefore, this variable is hypothesized to have a negative relationship to earnings growth.

Education

Education Levels

There is widespread agreement that the level of skill and education demanded in the labor marketplace is increasing, although there is uncertainty about the magnitude of this increase and about whether the process is accelerating (Teixeira and Swaim, 1991). Under these conditions, it is plausible that areas with a greater supply of more educated labor might experience more economic growth. At the same time, the demand for educated labor varies across industries. For some industries, an area where educational attainment is lower may be attractive.

Many previous studies attempted to assess the effect of educational attainment levels on business location decisions or business activity. Of those assessed in our literature review (Kusmin, 1994, pp. 44-47), most included only a single measure of educational attainment, either median years of schooling completed or percentage of adults who have completed high school. A few recent studies include variables that are intended to assess the effect of college-educated populations. However, only one of those reviewed directly compared the effect of high school completion and that of college completion on growth or business location. One study covered by our review tested for and found a significant positive relationship between the high school dropout rate among young adults and economic growth in metro areas. That is, metro areas with more dropouts experienced more rapid growth (Killian and Parker, 1991). Overall, the studies reviewed provide limited and sometimes contradictory evidence concerning the effects of educational attainment on economic growth.

The model here includes several distinct measures of the educational status of the local labor force. Adult educational attainment measures include both the percentage of the adult population (those

13. If union power influenced business decisions only through its effect on initial wage levels, we would not expect any union-related variable to have an effect on employment or earnings in any model that includes a separate control for such initial wage levels. However, if union power had an effect on rates of wage growth or decline (relative to productivity) during the 1980's, this will be reflected in our results, together with any effect of such power on changes in employment levels.

aged 25 or more) who had completed college, and the percentage of adults who had completed high school but not college.[14] In addition, the supply of young, lesser skilled labor is represented by the percentage of young adults (aged 16 to 19) who have dropped out of high school.

The model of growth used in this report controls for differences in the initial occupational and industrial mix of county employment (see pp. 16-18). Therefore, the estimated effects of 1980 educational attainment on subsequent earnings growth do not reflect any role that educational attainment levels may have played prior to 1980 in attracting or retaining industries that would contribute to county earnings growth during the decade of the 1980's. However, if educational attainment levels contributed directly to the attraction or retention of growing industries during the 1979-89 period, these effects will be reflected in our results.

Higher Education Activity

Several authors studying the determination of economic growth or industrial location decisions at the substate level have suggested that proximity to institutions of higher education might have a positive effect. For instance, one study suggested that access to such institutions may be "important as a means of reducing personnel cost through manpower development and in-service training, and . . .[may] make the community a more acceptable residential area for company personnel" (Smith, Deaton, and Kelch, 1978, p. 25). These authors tested several possible measures of access to such institutions, but found no consistent evidence of the proposed relationship.

The presence of significant levels of access to higher education is measured in this report by the ratio of college enrollment to the current labor force at the county level. This ratio has been converted to a dichotomous (dummy) variable that is equal to 1 in counties where the ratio is greater than 0.10, and equal to zero in all other counties. This variable serves as a measure of the relative importance of institutions of higher education in the local economy. Our working hypothesis is that colleges and universities may foster local economic growth by providing a source of specialized consultants and other highly skilled labor, and by offering cultural amenities that may be attractive to some types of businesses and individuals. This variable may also indicate the availability of more highly skilled young adults in the community. (While most full-time students will not be simultaneously employed in highly skilled positions, communities with institutions of higher education may benefit from the skills of those who complete their education and wish to remain in the community, and those who are employed in the community while attending college on a part-time basis.)

A dummy variable is used because these effects seem likely to hinge on the presence of significant higher education activity in the community, rather than being linear functions of college enrollments. In fact, the relationship of institutions of higher education to the local economy may be substantially different in the small number of counties where the ratio of college enrollments is far above the 0.10 threshold.

Local Revenue and Expenditure Variables

Many earlier studies of subnational economic growth assessed the effects of taxation. The conventional hypothesis is that higher levels of taxation have an adverse effect on economic

---

14. The percentage who had completed college is approximated by those who reported completion of 16 or more years of schooling, while the percentage who had completed high school but not college is approximated by those reporting 12 to 15 years of schooling. Data on actual high school or college completion are not available from the 1980 census.

growth through their influence on business and individual decisions. This hypothesis has been supported by many but not all studies (see reviews in Kusmin, 1994; Bartik, 1991; Wasylenko, 1991). However, it has also been suggested that the estimated effects of taxation may be sensitive to the treatment of public expenditures in the analysis. If taxes are used to pay for public goods and services that are sufficiently valuable to businesses and mobile individuals, the net effects of higher taxation and spending on economic growth may be positive, and estimated tax level effects might be positive if such expenditures are not controlled for separately (Bartik, 1991).

Local Taxes

The model considered in this report includes a measure of local tax levels (revenue effort) for each county. Because a comprehensive measure of the local revenue base was not available at the county level, local revenue effort is defined here as the ratio of locally raised general revenue to local personal income.[15] Local revenue effort is calculated for 1977, the date of the last Census of Governments prior to the beginning of our analysis period. (Data for 1977 are used because of the availability of Census of Governments data for all counties; data for intercensal years are not available for most smaller jurisdictions.)

Local Education Expenditures

Local spending on primary and secondary education may also affect local growth. In the short run, a higher level of local education spending might augment growth by attracting or assisting in the retention of adults concerned about the quality of education for their children. However, it might also repel childless adults as well as many businesses, as they might benefit from local expenditures on road maintenance, public safety, and other services, but are likely to see little immediate benefit from taxes paid for local education.[16] In the long run, local education spending may also attract business and augment growth if it leads to a higher quality labor force.

Several past studies of subnational economic growth have found significant effects associated with education spending (Kusmin, 1994, pp. 21-36). Our model includes an estimate of per-pupil local spending on primary and secondary education for each county in 1977.[17] Our estimates of spending per pupil reflect estimated student population sizes that are computed from 1977 populations and 1980 population age distributions for each county (Dubin, 1989).

Transportation

Access to markets for locally produced goods and services and to inputs needed for local production are factors thought to be key in determining the level of economic activity, and likely to influence changes in that level. Transportation systems provide access to both input and output

---

15. This ratio is expressed in percentage terms. Although we treat local taxes and local revenue as synonymous in the text, local own-source general revenue includes tax revenue (about 73 percent of the total in 1977), current charges including hospital and sewer charges, and miscellaneous general revenue including interest earnings (Census of Governments, 1977).

16. Because we have incorporated a separate variable to capture the effects of the level of local taxation on earnings growth, the estimated effects of per-pupil expenditures associated with this variable will not reflect any higher local taxes that may be required to pay for higher local education spending. However, for given levels of local revenue, and assuming constant intergovernmental assistance, higher levels of education spending imply lower levels of spending on other local services that may be more valued by childless adults and some businesses.

17. Data for 1977 are used because of the availability of Census of Governments data for all counties; data for intercensal years are not available for most smaller jurisdictions.

markets and, therefore, are expected to augment economic activity and growth. Hence, measures of access to transportation are frequently included in models of growth determination.

A number of earlier studies included a measure of access to interstate or other major highways, and several found a significant positive relationship between such access and economic growth or business location decisions (see citations in Kusmin, 1994, pp. 48-49). Several also included measures of access to air service, but these yielded little evidence that such access is important.

Four variables are included in the current model to capture possible effects of access to long-distance transportation systems. Access to major highways is measured by the number of major highway interchanges (exits) in the county, and by a dummy variable showing the presence of an intersection of two interstate highways. (The major highways included interstate highways and other divided, limited-access highways. Where a county had a limited-access bypass, but had no through limited-access highways attached to the interstate highway system, any interchanges associated with the bypass were not counted.)

Access to air travel is measured by the presence of an airport with scheduled passenger service, either: (a) within the county, or (b) within 50 miles of any part of the county. (General aviation airports with no scheduled service are not reflected in our analysis.)

Data on the distribution of highway interchanges and intersections and on county proximity to airports are based on unpublished tabulations done at the U.S. Department of Agriculture's Economic Research Service. Beginning-of-period data were not readily available for these variables, so more recent data were used.[18]

Business and Banking Structure

Our literature review included one earlier study that suggested a significant relationship between the structure of the business sector, in particular the size and vigor of the small independent business component of the economy, and local economic growth (Bluestone and others, 1989, cited in Kusmin, 1994, p. 67). In this report, we test separately for the importance of small independent businesses in each of three sectors of the economy: the goods-producing, producer services, and other services sectors.

The percentage of establishments that were both small and independent in 1980 is computed separately within each of the three sectors. An establishment may be an independent firm, or a branch or the corporate headquarters of a multi-establishment firm. Establishments with fewer than 20 employees are classified as small. The analysis focused on the percentage of establishments that were small and independent, rather than on the percentage of all employment in such establishments, because we believe that the proportion of establishments is a more sensitive measure of the vitality of the small business sector. These variables do not reflect the relative importance of each of these sectors as a whole in the local economy.

In light of some evidence in one study included in our literature review that a more open banking system fosters small business start-ups (Bartik, 1989, cited in Kusmin, 1994, p. 68), a variable to

---

18. To the extent that transportation access was expanded or reduced in response to county economic growth or decline, use of more recent data could bias upward our estimates of the effects of transportation access. This bias is unlikely to be substantial in the case of our estimates of highway access effects, as there were only modest additions to the highway system in nonmetro areas between 1979 and 1989, and most of these likely reflected plans already made in 1979. However, such a bias is more likely with regard to air service, since deregulation of airlines and substantial reductions in subsidized service to smaller airports occurred during the 1980's.

capture the effects of State branch banking laws on local growth is also included in the model. Based on a tabulation of changes in branch banking law reported in Milkove and Sullivan (1990), States are coded in terms of the liberality of their branch banking laws at the beginning of the study period.

Amenities

While cultural and natural amenities (such as access to musical or theatrical performances, beauty of the landscape, or resources for outdoor recreation such as boating, fishing, or camping) do play some role in the location of economic growth, they are not readily quantified and, therefore, have not appeared frequently in econometric studies of growth. In this report, three variables are included that reflect the presence of some natural amenities that might foster faster earnings growth. These three measures of natural amenities were developed by the Human Resources Branch of the Agriculture and Rural Economy Division at ERS (see McGranahan, 1993).

Climate

The index of climate attractiveness is defined as:

JANTEMZ + JANSUNZ - JULHUMZ - JULTEMRZ,

where

JANTEMZ = Z-score (average January temperature),[19]

JANSUNZ = Z-score (average number of days with sun in January),

JULHUMZ = Z-score (average July humidity), and

JULTEMRZ = Z-score (residual when average July temperature is regressed against average January temperature).

This index reflects presumed preferences of individuals and businesses for mild, sunny winters and for mild, dry summers. Data for this index are taken from the Area Resource File (ARF) of the Department of Health and Human Services (Office of Health Professions Analysis and Research) (Appendix A).

Topography

The index of landscape desirability or topographical index is defined as:

TOPOGZ - ELEVZ,

where

19. For observation (county) I, Z-score $(x_i) = (x_i - x^*)/SD_x$, where x^* is the mean value of x over all observations, and SDx is the standard deviation of x over all observations. That is, for a specified variable and set of observations, the Z-score for the ith observation indicates how many standard deviations it is from the average value.

TOPOGZ = Z-score (index of mountainous or rugged terrain),[20] and
ELEVZ = Z-score (average elevation).[21]

This index reflects presumed individual preferences for areas with rugged terrain and for low elevation. Rugged terrain is presumably desirable for its visual attractiveness and role in some recreational activities; while high elevation, by itself, is likely to be unattractive, because reduced oxygen levels will impede comfort and enjoyment of recreational activities for some people.

Bodies of Water

Water coverage is measured by:

ZLWATD = Z-score (percentage of land area covered by water).

Higher levels of this variable indicate the presence of lakes or rivers that may be suitable for recreational activity.[22] (The data are from the Census Bureau water coverage data file. The ratio of land to water area is truncated at 250 percent to preclude assigning exceptionally high values to certain counties that border the Great Lakes whose nominal land area extends far into the lakes.)

Other Amenities

In addition, as noted earlier (see p. 11), counties with a large concentration of college students are more likely to offer some cultural amenities, and counties identified as retirement counties are frequently among those with attractive natural or cultural features. Thus, any estimated association between growth and status as a retirement county or a college county may reflect, in part, the effect of such amenities.

Relationship to Metro Areas

Also included in our model is a measure of access to nearby metropolitan markets. The total populations of nearby or proximate Metropolitan Statistical Areas (MSA's) for each county were computed at ERS based on 1983 definitions of MSA's and 1980 MSA populations. An MSA is deemed proximate to a county if the central city of the MSA was within 50 miles of any portion of the county.[23] (The total proximate metro population for each county reflects, in some cases, the population of more than one nearby metro area.) Measures of access to metropolitan markets have been included in several past studies of growth in economic activity in nonmetro communities (see citations in Kusmin, 1994, pp. 49-50). The conventional hypothesis is that communities with

---

20. Coded at ERS based on topographic map.

21. Data from the ARF file.

22. Such bodies of water will also be useful in some cases for industrial purposes (for example, as a source for water to be used as a coolant in industrial processes) or transportation (navigable rivers or lakes connected to them, seacoasts with usable harbors), although this measure is not designed specifically to capture such uses.

23. This differs from the definition of MSA adjacency used to classify counties in the ERS system of urban-rural continuum codes (Beale codes). The latter codes reflect the strength of commuting ties with metro areas as well as physical proximity (see Butler, 1990). Because rural economic growth may be influenced by access to metropolitan markets even in the absence of substantial commuting, our analysis defines metro proximity in terms of distances alone. For example, rural industries may provide goods for nearby metropolitan areas, or residents of such areas may provide a clientele for any local tourism industry.

access to such markets will experience greater economic activity and, perhaps, faster economic growth than communities that lack such access.

Industrial Structure

The general economic success of an area over any period of time generally reflects the success of the industries that have concentrated there. A surge in demand for airplanes is likely to mean a growth spurt in Washington State, while growth in the domestic automotive industry is still good news for Michigan. Especially in the short run, growth in demand for goods and services already produced in an area is one of the most important factors influencing the rate of local economic growth. Therefore, it is perhaps surprising that so few of the studies covered in our review of the literature attempted to take these effects into account in explaining differences in subnational economic growth. Two of the most recent, however, do so. These studies include, among the factors predicting subnational employment growth, a weighted average of industry growth rates, with the weights reflecting the mix of industries in each area. Both find such measures to be significant predictors of growth (Kusmin, 1994, pp. 64-65).

In this report, variables are included in the model corresponding to the initial (1979) share in county employment of each of 75 industrial sectors to control for the influence of local industrial structure on county economic growth. These sectors correspond in most instances to two-digit industries in the Standard Industrial Classification (SIC). Employment shares include wage and salary employment, plus an estimate of the number of proprietors in each sector. See Appendix F for a list of the individual sectors, with the mean and standard deviation of each sectoral employment share variable.

The specification used here is considerably more flexible than one that summarizes the effect of an area's industrial structure with a single expected growth rate. The single-variable approach assumes that the importance of an industry for local or regional economic growth is simply proportional to that industry's employment or earnings share and the industry growth rate. However, this need not be so. A concentration of employment in an export-base industry that is growing or declining nationally may have a disproportionate effect on the local economy, as changing employment and earnings of local workers in that industry affects consumer demand in other sectors. Also, industrial sectors that have a small share of local employment may have a disproportionate effect on the local economy, if their presence or absence influences the migration decisions of businesses or individuals. For example, this might be true for some producer service industries. On the other hand, changes in economic activity in many locally oriented trade and service industries are especially likely to reflect the prosperity of the local or regional economy as a whole, rather than driving that prosperity. The specification used in this report allows for all of these possibilities. Because it does not place constraints on the estimated values of industry-effect parameters, it allows the data to indicate whether the employment share of a particular industry had a large or small effect on local earnings growth. This approach would be impractical for a small data set, but the use in this report of data for 2,346 counties reduces the importance of conserving degrees of freedom.[24]

Estimated effects for some industries may be subject to a substantial level of error, because of collinearity among the employment share variables for these industries (see Appendix B, p. B-8,

---

24. The degrees of freedom for a linear regression are equal to the number of observations minus the number of parameters estimated (coefficients of explanatory variables, where explanatory variables may include an intercept term and dummy variables that capture fixed regional effects) (Myers, 1990, p. 19). If the number of variables exceeds the number of observations, the effects of each variable cannot be estimated. To estimate these effects with any precision, the number of observations must be considerably larger than the number of variables.

and Appendix G). However, our primary interest here is to control for the effects of industrial mix and to assess the overall importance of such effects, not to assess the effects on growth of concentrations of employment in individual industries. Therefore, uncertainty about the effects associated with particular industries is not of crucial importance.

Diversification

Industrial diversification is generally viewed as a determinant of economic stability. Diversification *per se*, as opposed to diversification into fast-growing industries, is not necessarily expected to influence the rate of subnational economic growth. None of the empirical studies of economic growth included in the literature review for this report treat diversification as a potential growth predictor.

There are some theoretical reasons, however, to believe that the degree of diversification could influence growth. Theories of innovation and the diffusion of innovation that emphasize communication among those making similar products suggest that excessive diversification might retard growth. A small nonmetro might also require the benefits of specialization before any industry can operate at a large enough scale to be competitive. On the other hand, other theories of innovation emphasize communication among those making different products, and these suggest that a more diversified economy might be more progressive, and lead to more growth (Glaeser and others, 1991).

Beyond these theoretical bases for including diversification in a growth model, if industrial diversification is identified as a significant positive influence on stability, it is important to know of any evidence that such diversification also has either positive or (especially) adverse effects on economic growth.

Our measure of diversification in county i is:

$$DIVERS_i = -\sum_j |p_{ij} - \bar{p}_j|,$$

where p_{ij} is the percentage of employment in county i that is in industrial sector j and $\bar{p}_j$ is the percentage of all nonmetropolitan employment that is in industrial sector j. This is a variant of the national average diversification measure. A national average measure of diversification is one that defines an economy as perfectly diversified if the area's industry employment shares are all equal to the corresponding national industry employment shares. Specialization may be measured, as is done in this report, by the sum of absolute deviations of actual industry employment shares from this perfectly diversified economy, or by the sum of squared deviations or some other sum of transformed deviations.[25]

---

25. An earlier version of the model was also tested using an entropy measure of diversification. An entropy measure defines diversification by:

$$E_i = c(-\sum_j P_{ij} \log P_{ij}),$$

where P_{ij} is the proportion of employment in each sector j in county i and c is an arbitrary constant. This formula implicitly defines perfect diversification as the condition where the employment share of each industry is equal. The national average measure of diversification and the corresponding entropy measure had about equal explanatory power with respect to earnings growth. However, when both measures were included in the same model of earnings stability (see Maddala, 1988, p. 445), only the national average measure was significant. Hence, the latter has been used for consistency in our modeling of both earnings growth and earnings stability.

Because we control separately for the employment shares of industrial sectors, any effect estimated for this variable may be viewed as a pure diversification effect, rather than a reflection of any correlation between our measure of diversification and the concentration of fast-growing industries in the county.[26]

Other Economic Base Characteristics

Transfer Payments

Transfer payments per capita in 1979 are also included as a variable in the model. Transfer payments consist of cash or goods, received by individuals, that are not returns to investments or to current employment. Most transfer payments are associated with Government programs, such as Social Security, Medicare, food stamps, unemployment compensation, and veterans' benefits. Retirement programs account for the largest share of transfer-payment income in both metro and nonmetro areas, although unemployment insurance, income maintenance programs, and veterans' benefits also account for substantial payments (Bentley, 1988).

A few past studies have included some measure of State expenditures on transfer payments as a potential predictor of State economic growth, but none of the studies that we reviewed specifically addressed the effects of transfer-payment recipiency. While some types of transfer-payment recipients (for example, welfare recipients) may be viewed negatively by potential in-migrants or relocating businesses, we make no *a priori* assumptions about how transfer-payment recipiency as a whole might affect earnings growth.

Occupational Structure

While infrequently represented in the econometric analyses of subnational economic growth that were reviewed for this report, the occupational mix of the local labor force may have a significant influence on growth (see, for example, Thompson and Thompson, 1987). For example, new employers, particularly those in higher wage industries, may be more likely to locate in areas where more of the labor force has technical or professional experience.

The model used in this report takes into account the percentage of the local employed population in each of seven broad occupational groups: professional, managerial, technical, sales-and-clerical, skilled crafts, operators and laborers, and service occupations. However, these are included in the model primarily to control for any occupational mix effects. We do not focus at length on the estimated effects of the individual variables.

Regional Dummy Variables

Regional dummy variables have frequently been found to have significant influences on growth in past studies of subnational economic growth. Such variables are included in the basic model to

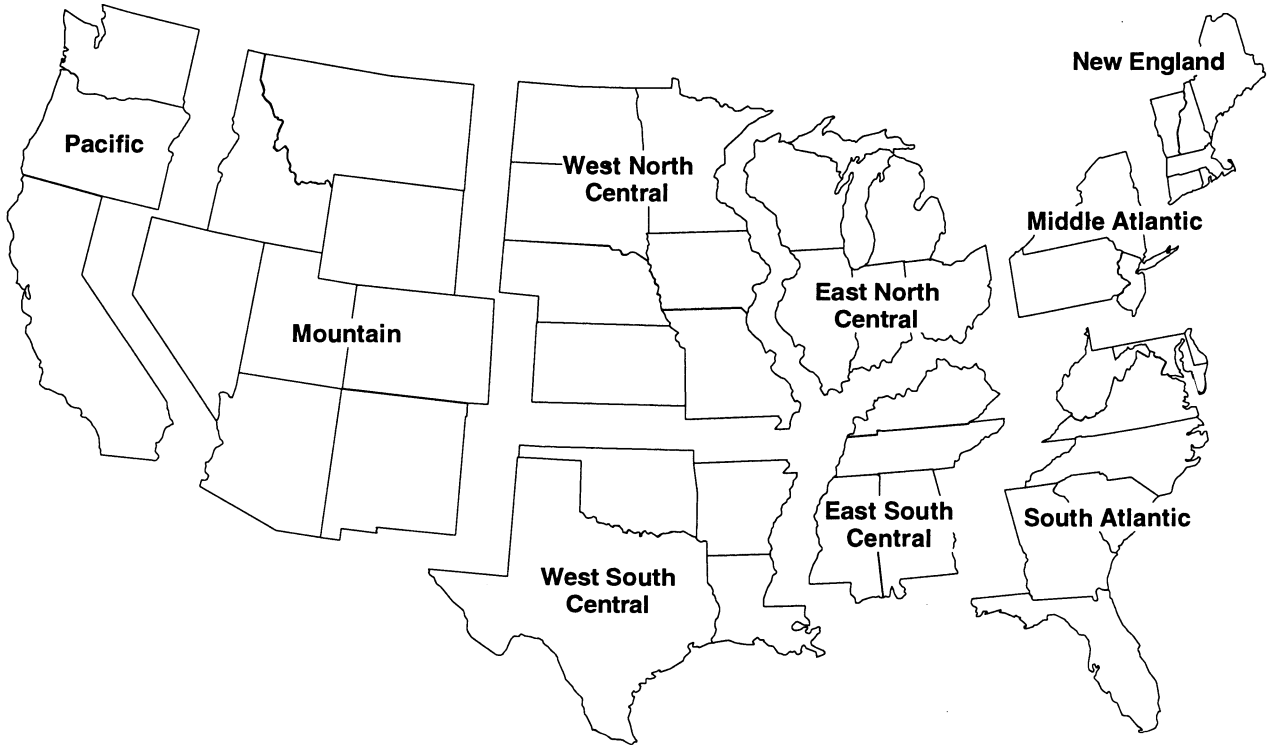
---

26. Such a correlation could arise, for example, if employment in declining or slow-growing industries occurred primarily in counties where employment was dominated by a single industry (for example, farming or coal mining), while employment in fast-growing industries (for example, a variety of service industries) tended to be found in counties that were less specialized. Empirically, our data show that 1979 employment shares for a number of manufacturing industries, as well as most trade and service industries, tended to be higher in more diversified counties, while diversification is negatively correlated with employment shares in farming, mining, construction, textile and wood products manufacturing, hotels, amusement services, and Federal employment.

capture any effect of regional differences not otherwise accounted for in the model.[27] Such differences could include differences in variables that are potentially measurable, but were not specifically included in the model; examples include geographic access to domestic or foreign markets. Regional dummies may also capture differences that are harder to quantify, such as regional differences in political climate or cultural receptivity to innovation.

Dummy variables are included for each of the nine census divisions (see figure 1). Treatment of dummy variables and controls for regional differences are discussed below (p. 26).

Figure 1.
U.S. Census Divisions



Source: U.S. Statistical Abstract, 1994-95 (114th Edition), Table No. 600.

Table 1 lists variables used in this report, with a brief description of each, and the mean and standard deviation for each variable. Corresponding data for variables capturing the percentages of employment in each industry and occupation are reported in appendix tables F-1 and G-1. Data sources are listed in Appendix A. Reported means and standard deviations are based on actual

27. When a sample or a population falls into a set of mutually exclusive categories, such as census division, a series of dummy variables (defined in Appendix B, p. 6) can be defined to capture the effect of membership in each of these categories.

Table 1--Variables included in the model

Variable	Description	Mean	Standard deviation
DEPENDENT VARIABLE:			
Growth index	100 * change in logarithm of total real labor and proprietor income, 1979-89	-1.4	25.8
INDEPENDENT VARIABLES:			
Demographic characteristics--			
Total urban population (1,000)	Urban population of county in thousands, 1980	8.6	12.0
Percent African-American	African-American population as percentage of county population, 1980	8.3	15.2
Percent Hispanic	Hispanic population as percentage of county population, 1980	3.7	10.6
Retirement county	1 if retirement county, 0 otherwise	0.205	0.403
Percent of population aged 25-64	Population aged 25-64 as percentage of county population, 1980	44.70	2.87
Labor market characteristics--			
Mean annual earnings (log)	Logarithm of mean 1976-78 real earnings per job (in $)	9.108	0.193
Right-to-work law	1 if right-to-work State, 0 otherwise	0.551	0.497
Labor force participation rate	Total labor force as percentage of population aged 16-64, 1980	69.86	6.41
Education levels and activity--			
Percent high school graduates	Percentage of adults aged 25+ who have completed at least 12 but less than 16 years of schooling, 1980	47.06	9.90
Percent college graduates	Percentage of adults aged 25+ who have completed 16 or more years of schooling, 1980	10.26	4.32
Percent dropouts (age 16-19)	Persons aged 16-19 who are not in school and have not completed 12 years of schooling as a percentage of all persons aged 16-19, 1980	14.95	7.65
Local college enrollment	1 if ratio of enrolled college students resident in county to county labor force was greater than 0.10 in 1980, 0 otherwise	0.122	0.327
Local taxes and expenditures--			
Local tax level	Local own-source general revenue as a percentage of local personal income, 1977	5.86	3.12
Education spending per pupil (K-12)	Current expenditures on primary and secondary education per student, 1977, in thousands of (1977) dollars	1.415	0.439
Transportation access--			
Highway interchanges	Number of interchanges (exits) of interstate and other limited-access highways	1.75	3.02
Highway intersections	1 if county contains an intersection of two interstate highways, 0 otherwise	0.020	0.140
Airport in county	1 if there is an airport with regularly scheduled passenger service in county, 0 otherwise (1985)	0.071	0.257
Airport within 50 miles	1 if there is an airport with regularly scheduled passenger service within 50 miles, 0 otherwise (1985)	0.884	0.321

See notes at end of table.

continued--

Table 1--Variables included in the model--continued

Variable	Description	Mean	Standard deviation
Business and banking structure--			
Small independent businesses as percentage of all businesses in--	Percentage of all business establishments in a sector that were independent establishments with fewer than 20 employees, 1980		
Goods-producing sector		77.6	10.6
Producer services sector		80.6	6.0
Other services sector		71.2	16.5
Branch banking law	1 if State laws permitted open statewide branch banking in 1979, 0 otherwise	0.138	0.345
Amenities--			
Climate quality index	Index reflecting temperature and humidity in July and in January	0.033	1.889
Topography index	Index reflecting mountainousness and elevation	-0.082	1.185
Water coverage index	Index reflecting percentage of county covered by water	-0.140	0.988
Relationship to metro areas--			
Population of metro areas within 50 miles (in millions)	Total population of Metropolitan Statistical Areas within 50 miles, in millions	0.527	0.857
Region--			
New England	1 if county is in New England, 0 otherwise	0.017	0.128
Middle Atlantic	1 if county is in Middle Atlantic, 0 otherwise	0.026	0.159
East North Central	1 if county is in East North Central, 0 otherwise	0.129	0.335
West North Central	1 if county is in West North Central, 0 otherwise	0.236	0.425
South Atlantic	1 if county is in South Atlantic, 0 otherwise	0.166	0.372
East South Central	1 if county is in East South Central, 0 otherwise	0.125	0.331
West South Central	1 if county is in West South Central, 0 otherwise	0.159	0.366
Mountain	1 if county is in Mountain, 0 otherwise	0.107	0.309
Pacific	1 if county is in Pacific, 0 otherwise	0.035	0.185
Economic base--			
Transfer payments	Receipt of transfer payment income per capita ($1,000), 1979	0.964	0.196
Industrial diversification	Measure of diversification: (-1) x sum of absolute percentage deviations of county industrial mix variables from nonmetro average (possible values are from -200 (perfect specialization) to 0 (perfect diversification)	-71.2	16.7

Values are for 2,346 counties included in regression analyses.
Means and standard deviations for the percentage of county employment in each occupation and industry are reported in appendix table F-1.

Source: U.S. Department of Agriculture, Economic Research Service.

values of the variables, and do not reflect the transformation of selected variables for computational purposes that is described on page 26.

Constructing an Econometric Model of Local Economic Growth[28]

This section discusses the techniques used in this report to estimate the effects of local and regional characteristics on earnings growth over the 1980's. Readers familiar with linear regression methods and basic econometrics will be able to skip much of this section. Also, readers who are interested in our results but not in our methodology may wish to skip to the section that discusses these analyses (beginning on p. 29).

Linear Regression Model

The basic technique that we use to examine the factors that might be associated with economic growth is multivariate linear regression. This is a standard and widely used statistical tool. It is used to describe the relationship between an outcome of interest, often called the "dependent variable," and a number of other variables that are thought to influence the outcome, often called the "independent variables" or the "regressors."[29] The estimated effect of each independent variable on the outcome is measured by a regression coefficient for that variable, sometimes called simply a "coefficient," and also sometimes described as an estimated parameter.

One important feature of multivariate regression is that it estimates the effect of each of the independent variables after taking into account the estimated effect of each other variable. For instance, simple two-way correlations might show that both location in the South and the percentage of the population that is African-American were positively associated with rural county growth during the 1980's. However, multivariate regression might tell us that only one of these characteristics has a distinct effect on growth, while the other has no effect after the estimated effect of the first has been taken into account, and is correlated with growth only because most rural African-Americans live in the South.

There is a standard error associated with each of the estimated coefficients of the econometric model; this is a reflection of how much uncertainty there is about the value of that coefficient. If the model used to estimate the coefficients has been correctly designed, and if the number of observations in the data is sufficiently large relative to the number of coefficients being estimated (which is true for this report), then there is a 95-percent chance that the absolute difference between the estimated value of a particular coefficient and the (unknown) true value of that coefficient is less than 1.960 times the coefficient's standard error. Accordingly, the 95-percent confidence interval for the coefficient is defined as the range (estimated coefficient -1.960 and +1.960 standard errors); there is a 95-percent chance that the true value falls within that range.[30]

---

28. The term "econometric model" is used in this report to refer to the model of growth that specifies a linear relationship between earnings growth and a set of local and regional characteristics, in combination with the particular set of econometric techniques used to estimate the coefficients of the model and the associated standard errors.

29. The list of regressors may also include proxy variables thought to be useful in capturing the effects of other (unmeasurable) variables that influence the outcome.

30. To be precise, given an estimated value and an unknown true value, the true value either does or does not fall within the confidence interval. It is therefore more accurate to say that, if one estimates confidence intervals for coefficients in a large number of independent studies, the confidence interval will contain the true value 95 percent of the time.

While a 90-percent, 99-percent, or other confidence interval can be analogously defined, the 95-percent confidence interval is the one most commonly used in social science research.

The ratio of a coefficient estimate to the standard error of that estimate is known as the coefficient's "t-statistic." If the true value of a regression coefficient is zero, there is a 5-percent chance that the absolute value of the estimated t-statistic will be greater than 1.960, and a 1-percent chance that it will be greater than 2.576. Therefore, an estimated coefficient with a t-statistic of greater than 1.960 or less than -1.960 is described as "(statistically) significant at the 5-percent level," meaning that an observed value this large in magnitude would be relatively unlikely if the true value of the coefficient were zero, or were of the opposite sign from the estimated coefficient. Similarly, if the absolute value of the t-statistic is greater than 2.576, the coefficient is described as "(statistically) significant at the 1-percent level."[31] In this report, if a finding is described as "significant" or "statistically significant" without further qualification, it means that it is statistically significant at the 5-percent level.

The simplest and most widely used multivariate regression technique is ordinary least-squares (OLS). A fuller description of multivariate linear regression and OLS can be found in Appendix B.

Econometric Issues

Heteroskedasticity

Use of OLS regression is simple and straightforward. However, OLS may not yield the most accurate coefficient estimates if the assumptions that underlie it are not met. For instance, OLS assumes that the model will fit all kinds of observations equally well. However, it may be that the model fits the data better for some kinds of counties than for others. If the variance of actual outcomes around their expected values is not the same for all observations, we have the problem of heteroskedasticity.

One way to address the problem of heteroskedasticity is by using weighted least-squares (WLS). This approach gives less weight to observations that are less predictable. There are several techniques by which weights can be assigned to observations to correct for heteroskedasticity. Among the more flexible, used in this report, is a technique that estimates weights based on the assumption of multiplicative heteroskedasticity.

Heteroskedasticity and the weighting technique used in this report are discussed in more detail in Appendix B.

Robust Regression

For OLS to be efficient, the differences between actual and expected outcomes must follow a normal statistical distribution. This also holds for WLS after the data are weighted. If these differences are not distributed normally, the resulting coefficient estimates will be inefficient, and the estimated t-statistics and confidence intervals may be misleading. In particular, if the tails of the residual distribution, or the distribution of weighted residuals in the case of WLS, are "fatter" than the tails of a normal distribution, then OLS or WLS will give too much weight to those cases with large residuals.

---

31. Note, however, that these statements are true only if the major assumptions of the linear regression model hold true. For instance, they assume that no significant independent variables have been left out of the model, and the differences between actual and fitted values of the dependent variable fit a normal statistical distribution with mean zero.

Transformation of Selected Variables

As noted earlier (footnote 27, p. 19), when categorical variables are included in a regression model, they are represented by dummy variables. When a categorical variable has several possible values, one category must be omitted to identify the model. Often, the choice of the omitted category is arbitrary, but the interpretation of the results may be influenced by that choice.[34] Similarly, where a group of continuous variables is related in such a fashion that the sum of their values is fixed (for example, the percentage of total employment in each industry must sum to 100 across all industries), then ordinarily one of the variables would be omitted from the regression specification to make estimation possible.

In this report, an alternative approach to these cases has been used, adapted from a suggestion of Suits (1984). The alternative approach involves a transformation of the data so that the regression coefficients for a group of related variables can be interpreted as deviations from the average effect of those variables, rather than as deviations from an arbitrarily selected category. This transformation is described in more detail in Appendix B. This approach has been applied to the divisional dummy variables included in the model for some methods, to the percentage of employment in each industry, and to the percentage of employment in each occupation.

Econometric Methods[35]

In the following section, we discuss results for four econometric models of local earnings growth. (Results for some other econometric models are addressed in some of the appendixes, and briefly in the text and footnotes.) Three of these models reflect application of distinct econometric methods to the linear regression model that was described earlier (pp. 1-22). A fourth model reflects inclusion of an additional variable in the regression model to capture possible inter-county spillover effects on economic growth.

Preferred Method

In this discussion, we will emphasize the results yielded by what we consider our "preferred method." This might also be described as the "fixed-effects multiplicative-heteroskedasticity-corrected bounded-influence method." This econometric method incorporates the corrections for fixed spatial effects, heteroskedasticity, and non-normal errors that are discussed above (pp. 23-25) and in Appendix B (pp. 3-6 and 9-10). Statistical evidence supports each of these corrections, and so we view an analysis incorporating all three as the most reliable.

Other Econometric Methods

We believe that each of the techniques embodied in our preferred method is appropriate, and that the resulting estimates of how local and regional characteristics affect earnings growth are better than estimates that do not reflect these techniques. However, it may still be useful to assess how sensitive our results are to the particular estimation techniques used. For example, suppose that the estimated effect of a particular variable is similar whether or not we correct for

---

34. For example, suppose our model includes dummy variables for all but one region. If the omitted variable represents the slowest growing region, then the estimated coefficients for each of the other dummy variables will be positive, and it may be difficult to distinguish the fastest growing regions from those experiencing average growth. On the other hand, if the omitted variable represents the fastest growing region, then all estimated regional effects will be negative.

35. The term "econometric method" is used in this report to describe the combination of techniques that have been used to estimate coefficients of the regression model and their standard errors.

heteroskedasticity. We then have some reason to expect that another econometric method that reflected a different approach to heteroskedasticity would yield a similar estimate of that effect. On the other hand, if the estimated effect is extremely sensitive to the application of the multiplicative heteroskedasticity technique, then use of a different technique to correct for heteroskedasticity might yield substantially different results.

Accordingly, we also estimate the regression model using alternative econometric methods that omit certain corrections that are part of the preferred method. Besides the preferred method results described above (p. 26), we also report:

1) OLS results for each effect. These reflect none of the corrective techniques discussed on pages 23-25.

2) Results for an alternative econometric method that incorporates corrections for heteroskedasticity and non-normal errors, but that does not correct for fixed spatial effects. This may be described as the multiplicative-heteroskedasticity-corrected bounded-influence method, or more simply as the robust-weights method.

In addition, we consider an alternative approach to modeling the effects of commuting-zone characteristics on county earnings growth. While our preferred method assumes that there is a commuting-zone-specific fixed effect for all counties in a commuting zone that is not further explained by the model, this alternative approach assumes instead that the correlation of growth rates across counties within a commuting zone can be modeled as a spillover effect. That is, earnings growth in any given county may be stimulated by increased demand for goods and services due to earnings growth in nearby counties in a common labor market. Thus, the model estimated with the robust-weights method is augmented by inclusion of the contemporaneous growth rate of earnings in the other counties of the commuting zone as an additional independent variable. This is referred to as the rest-of-CZ method.

We discuss below (pp. 42-51) the extent to which results of the preferred method are similar to and reinforced by the results of these alternative methods, and the extent to which particular results are sensitive to our corrections for econometric problems.

Table 2 exhibits the econometric problems addressed and the techniques applied for each of the four econometric models used in this report. Here and throughout this report, the four econometric models considered are listed from left to right in order of increasing control for statistical problems. The OLS model does not control for any of the problems discussed here, while the robust-weights model controls for both heteroskedasticity and non-normality. The rest-of-CZ model also includes a limited attempt to address the correlation of earnings growth rates within commuting zones, while the preferred model addresses this problem in a more powerful way. Finally, the backward elimination model may be viewed as addressing each of these problems as well as that of multicollinearity. However, it does so at the cost of raising issues of omitted variable bias and inference problems associated with a method that picks out selected independent variables based on model fit, and, hence, we do not view it as the preferred method.

Table 2--Econometric problems addressed by each econometric method

Problem	OLS	Robust-weights method (MH bounded-influence)	Rest-of-CZ method	Preferred method (fixed-effects MH bounded-influence)	Fixed-effects backward-elimination
Spatial effects	No	No	Partially	Yes	Yes
Heteroskedasticity	No	Yes	Yes	Yes	Yes
Non-normal deviations from model predictions	No	Yes	Yes	Yes	Yes
Multicollinearity	No	No	No	No	Partially
Omitted variable bias	Yes	Yes	Yes	Yes	Partially
Understatement of standard errors when estimated intermediate parameters are treated as fixed	NA	No	No	No	No

OLS = Ordinary least-squares.

MH = Corrected for multiplicative heteroskedasticity.

CZ = Commuting zone.

Yes = Econometric problem is addressed by this method.

No = Econometric problem is not addressed by this method.

NA = Not applicable (this econometric problem does not arise for this method).

Source: U.S. Department of Agriculture, Economic Research Service.

Explanatory Power of Model Using Alternative Methods

Table 3 reports the adjusted r^2 value for each of the four methods, computed with both weighted and unweighted data. (Note that the robust-weights method with unweighted data is the same as the OLS method).

Table 3--Adjusted r-squared for weighted and unweighted data

Statistic	OLS	Robust-weights	Rest-of-CZ	Preferred method
Adjusted r-squared	0.4006	0.5107	0.5486	0.6700
Adjusted r-squared for unweighted estimates	0.4006	0.4006	0.4458	0.5534

OLS=Ordinary least-squares.
CZ=Commuting zone.
Unweighted robust-weights estimates are equivalent to OLS estimates.
Source: U.S. Department of Agriculture, Economic Research Service.

Because controls for fixed effects capture much of the variance in earnings growth without explaining such variance in a substantive fashion, and because the weighting procedures used reduce the weight given to observations that are difficult to explain, the conventionally computed adjusted r^2 of 0.670 for the preferred method overstates the explanatory power of the model. If the variance explained by the commuting-zone effects is subtracted from both the explained and the total variance, and if the associated degrees of freedom are adjusted appropriately, the r^2 that results is 0.263. This value reflects the power of the remaining variables, including industry and occupational mix variables, to explain variation in growth rates among counties within commuting zones. (It is equivalent to the r^2 that would result from converting all variables into deviations from commuting-zone means, estimating the model with the resulting datasets, and adjusting for the degrees of freedom associated with computation of the commuting-zone means.) The corresponding value is slightly lower (0.242) for a model with fixed effects, but without weighting to correct for heteroskedasticity or non-normality. That is, intra-CZ variation in the independent variables in the model can account for about one-quarter of intra-CZ variation in earnings growth rates during the 1980's.

It is also useful to consider an analogous computation with respect to models with census division dummies but without fixed commuting-zone effects. This indicates that (after correcting for degrees of freedom) industry, occupation, and noncontrol variables can account for 30.9 percent (unweighted data) to 38.6 percent (weighted data) of the variation in earnings growth within census divisions. However, this does not mean that the estimates without commuting-zone fixed effects are superior. Rather, these results indicate that it is easier to account for variation in earnings growth between commuting zones than variation across counties within commuting zones.

Discussion of the Results of Regression Analyses Using the Preferred Method

We now turn to discussion of the results of our empirical analyses. The emphasis here is on the results of a regression analysis controlling for heteroskedasticity, non-normal residuals, and fixed commuting-zone effects, which we have termed the "preferred method." We discuss in the next chapter results estimated with several alternative methods, reviewing the extent to which the

results are robust across the preferred and alternative estimation procedures. (The model also includes variables that reflect the percentage of 1979 county employment in each of 75 industrial sectors, and the percentage of 1980 employment in each of eight occupational groups, and results for these variables are reported in Appendix C. However, results with respect to the effects of individual industrial sectors or occupational groups are discussed only briefly in the body of the text and are not included in the tables that appear in the body of the text.)

Tables 4 and 5 show results estimated using the preferred method. This method includes corrections for fixed commuting-zone effects, for heteroskedasticity, and for non-normality in the process that generates earnings growth rates. These corrections are discussed in more detail on pages 23-26 and in Appendix B.

Table 4 identifies county characteristics (excluding industrial sector and occupational group effects) for which estimated effects were statistically significant at the 5-percent level when the preferred method was applied to estimating the growth model, and the sign of each of those effects. Characteristics that were included in the model but did not have an effect significant at the 5-percent level are listed in part C of table 4. (Unless stated otherwise, when the term "statistically significant" is used later, it means "significantly different from zero at the 5-percent level" (See pp. 22-23).)

Table 5 provides additional detail on the results of estimating the growth model using the preferred method. This table indicates for each variable in the model its estimated coefficient, a standardized measure of that effect (see the Standardized coefficients box), the t-statistic, and the statistical significance of the coefficient.[36]

Where our preferred regression method yields little evidence of earnings growth effects for particular variables, this does not mean that those variables do not have an effect on growth; some of them may be important growth factors. However, it does mean that we are not able to confirm the existence of such effects. While the effects of these variables may be nonexistent or negligible, it is also possible that measurement error, limited variability of the characteristic, the inherent uncertainty of the regression model, or other limitations of the data or techniques used have interfered with their measurement.

It may be helpful to compare nonsignificant preferred-method results with results from other methods, as we do in pages 42-51. However, this will not be decisive. Alternative methods may avoid some weaknesses of the preferred method, but some problems will show up in the results for all methods. Further, significant results from other methods that are inconsistent with those of the preferred method may merely reflect the limitations of those methods.

Demographic Characteristics

Two demographic characteristics of rural counties were significant predictors of earnings growth in the 1980's, according to the preferred-method estimates. These were status as a retirement attraction county, associated with faster growth, and the percentage of the population that is African-American, associated with slower growth.

---

36. Because estimation of the econometric model required treating certain estimated parameters as fixed, the t-statistics and significance levels reported here may be somewhat inflated. An earlier version of the model was estimated using the robust-weights method, and a bootstrapping technique was employed to correct the estimated parameter standard errors (see Appendix B for a discussion of bootstrapping). The results indicated that only a modest inflation of t-statistics (8.3 percent on average) resulted from using the conventionally computed standard errors. Computational difficulties precluded any application of a bootstrap technique to the preferred method estimates.

Table 4--Summary of results of estimating growth model using preferred method

A. Variables with significant positive effects on growth

Retirement county ▒ ***
Right-to-work law ▒ *
Percent high school graduates **
Education spending per pupil (K-12) *
Highway interchanges **
Airport within 50 miles ▒ *

B. Variables with significant negative effects on growth

Percent African-American **
Mean annual earnings (log) ***
Transfer payments ($1,000 per capita) ***

C. Variables with no significant effects on growth

Total urban population (1,000)
Percent Hispanic
Percent of population aged 25-64
Labor force participation rate
Percent college graduates
Percent dropouts (aged 16-19)
Local college enrollment ▒
Local tax level
Highway intersections
Airport in county ▒
Small businesses as percentage of all businesses in goods-producing industries +
Small businesses as percentage of all businesses in producer service industries
Small businesses as percentage of all businesses in other service industries
Branch banking law ▒
Topography (mountainousness) index +
Water coverage index
Population of metro areas within 50 miles
Industrial diversification index

* Statistically significant at 5-percent level.
** Statistically significant at 1-percent level.
*** Statistically significant at 0.1-percent level.

+=Estimated effects of these two variables are significant at the 10-percent level and are negative.

Coefficients of variables reflecting employment by industry and occupation are not represented.

▒= 0, 1 variable.

Source: U.S. Department of Agriculture, Economic Research Service.

Table 5--Estimated parameters of county earnings growth model using preferred method, including corrections for fixed commuting-zone effects, heteroskedasticity, and non-normality

Dependent variable is 1979-89 real earnings growth.

Variable	Regression coefficient	Standardized coefficient	T-statistic
Intercept	0	NA	0
Demographic characteristics:			
Total urban population (1,000)	-0.0263	-0.0122	-0.58
Percent African-American	-0.1920	-0.1127	-3.01 **
Percent Hispanic	-0.070	-0.0290	-0.66
Retirement county ▒	4.50	0.0702	3.85 ***
Percent of population aged 25-64	0.119	0.0132	0.52
Labor market characteristics:			
Mean annual earnings (log)	-23.54	-0.1755	-5.25 ***
Right-to-work law ▒	5.25	0.1010	2.26 *
Labor force participation rate	-0.039	-0.0097	-0.33
Education levels and activity:			
Percent high school graduates	0.331	0.1571	2.61 **
Percent college graduates	-0.173	-0.0290	-0.71
Percent dropouts (aged 16-19)	-0.0266	-0.0079	-0.33
Local college enrollment ▒	-0.43	-0.0054	-0.32
Local taxes and expenditures:			
Local tax level	-0.084	-0.0101	-0.40
Education spending per pupil (K-12)	3.77	0.0640	2.39 *
Transportation access:			
Highway interchanges	0.425	0.0496	3.23 **
Highway intersections	0.27	0.0015	0.11
Airport in county ▒	0.71	0.0070	0.48
Airport within 50 miles ▒	3.36	0.0417	2.26 *
Business and banking structure:			
Small businesses as percent of all businesses in--			
Goods-producing industries	-0.1076	-0.0439	-1.94 +
Producer service industries	0.1298	0.0303	1.45
Other service industries	-0.0188	-0.0120	-0.73
Branch banking law ▒	-2.62	-0.0350	-0.98

See notes at end of table.

continued--

Table 5--Estimated parameters of county earnings growth model using preferred method, including corrections for fixed commuting-zone effects, heteroskedasticity, and non-normality

Dependent variable is 1979-89 real earnings growth.

Variable	Regression coefficient	Standardized coefficient	T-statistic
Amenities:			
Climate quality index	NA	NA	NA
Topography (mountainousness) index	-1.321	-0.0605	-1.86 +
Water coverage index	-0.600	-0.0229	-1.07
Relationship to metro areas:			
Population of metro areas within 50 miles (in millions)	-0.301	-0.0100	-0.49
Economic base:			
Transfer payments ($1,000 per capita)	-15.67	-0.1186	-4.67 ***
Industrial diversification index	0.0629	0.0406	1.48

NA = Not estimated or not applicable.
▒ = 0, 1 variable.
* Statistically significant at 5-percent level.
** Statistically significant at 1-percent level.
*** Statistically significant at 0.1-percent level.
\+ Statistically significant at 10-percent level.

Note: Coefficients of variables reflecting employment by industry and occupation are not reported here; they are shown in Appendix C.

Source: U.S. Department of Agriculture, Economic Research Service.

Standardized coefficients

In addition to the regression coefficient and t-statistic for each variable, the standardized value of each regression coefficient is also reported in several tables. Where $Y = a + bX$, σ_X is the standard deviation of X, and σ_Y is the standard deviation of Y, the standardized coefficient of $X = b * \sigma_X / \sigma_Y$, and can be interpreted as the number of standard deviations of Y by which Y will shift if the value of X changes by one standard deviation of X. These standardized coefficients would be equal to the actual coefficients if the dependent variable and all independent variables were rescaled to have a variance of 1 before the model was estimated.

Use of standardized coefficients allows us to compare the relative importance of variables in determining county growth when those variables are not measured in comparable units. For example, it would be difficult to assess the relative importance of educational attainment and educational spending for local economic growth by a comparison of the raw regression coefficients for each variable, because while educational attainment is measured by the percentage of adults who have completed high school or college, educational spending is measured in dollars per pupil.

Even if two variables are measured in comparable units, the standardized coefficient is a better guide to the relative importance of two variables when the variance of one is much greater than the variance of the other. For example, both the African-American population and the college graduate population are measured as a percentage of total county population (population 25 and over in the second case). Yet, while the estimated effect of a 1-percentage-point difference in the African-American population (a 0.192-percent difference in cumulative earnings growth) is comparable in magnitude to the effect of a 1-percentage-point difference in the college graduate population (a 0.173-percent difference), the African-American population was a much more important predictor of economic growth in the 1980's, because there was much more variation in African-American population shares. This is reflected in the relative magnitudes of the standardized coefficients for the African-American population and the college graduate population, -0.112 and -0.029, respectively.

The reader may wish to remember that these statistics reflect variance among counties, unweighted by population. In some instances, large standardized coefficients may reflect variability in the characteristic of interest in a large number of counties with relatively sparse populations.

In most instances, the discussion in the text emphasizes results for conventional coefficients. However, the standardized coefficients have been used as a guide in describing the coefficients as large or small, with standardized values greater than 0.05 in absolute value indicating a large coefficient. It may be useful to note that dummy variables with a relatively large effect on growth in individual counties may have relatively small standardized coefficients. This can occur if the characteristic of interest is found in nearly all counties, or is found in very few counties, so that the variance of the variable among counties is small.

Retirement Attraction

Those counties that were retirement attraction counties during the 1970's experienced more earnings growth during the 1980's. This presumably reflects, at least in part, continued in-migration of retirees to these counties and the resulting creation of new jobs to meet their needs. It may also reflect the in-migration of businesses and nonretired individuals drawn to these counties by some or all of the same features that attracted retirees in the 1970's. Under the assumptions of our preferred method, earnings in retirement counties grew 4.50 percentage points more than earnings in other counties over the 1979-89 period, after all other differences between the two groups were taken into account (see table 5).[37]

African-American Population

The data provide evidence that concentrations of African-American population were negatively associated with earnings growth during the 1980's, even after controlling for wage and education levels. The preferred-method results imply that a 10-percentage-point difference in the African-American share of the county population was associated with an inverse 1.92-percentage-point

37. This effect is statistically significant at the 0.1-percent level.

difference in cumulative 1979-89 earnings growth.[38] These results are suggestive of bias against heavily African-American areas by geographically mobile employers. However, other possible explanations include concerns about the tax base and level of business-oriented government services in counties with largely African-American populations, and the lower rate of entrepreneurship in the African-American community.[39]

Other Demographic Variables

The remaining demographic variables included in the model (urban population, percentage Hispanic, and percentage of population aged 25 to 64) did not have statistically significant estimated effects under the preferred method.

Labor Market Characteristics of Counties

Not surprisingly, the terms on which labor was available in a county appear to have been among the most important influences on county earnings growth during the 1980's. Labor market characteristics with a significant effect on earnings growth include wages and the presence of right-to-work laws.

Wages

If economic growth is driven by costs to employers, it is expected that higher wage levels in rural areas will be associated with slower growth, with other things, including skill levels, being equal. This association is likely to be stronger if many rural industries are labor-intensive. Results of this report are clearly consistent with that expectation.

However, while the effects of wages are substantial, they are not dramatic. The preferred-method wage effect estimate of -23.54 implies that a 10-percent difference in earnings per job was associated with an inverse 2.35-percent difference in total earnings growth over 10 years. This is equivalent to an elasticity of earnings with respect to wages of -0.23.[40,41]

However, the longrun response of total earnings to wages may be substantially greater, depending on the speed with which economic activity levels respond to wage differentials. (The "longrun" response is the response to a change or difference that will be observed when economic agents have fully adjusted to that change or difference.) If the longrun response is moderately greater than what we observe over a 10-year period, these results suggest a wage elasticity of total county earnings on the order of -0.3, toward the lower end of the range found in the literature by Bartik (1991).

---

38. This effect is statistically significant at the 1-percent level.

39. In 1987, African-American-owned firms accounted for 3.1 percent of all U.S. firms (U.S. Department of Commerce, 1991, p. 535).

40. This parameter is statistically significant at the 0.1-percent level.

41. Because both earnings and the wage measure have been entered into the model in logarithmic form, this reported elasticity of total earnings with respect to the initial wage level holds across the range of both variables. However, the implicit assumption of a constant elasticity has not been tested.

Right-to-Work Laws

If mobile employers prefer environments where organized labor is less influential, then we might expect to find lower rates of earnings growth where union power is greater. Levels of unionization are not included in the model, for reasons that are discussed on page 10; but right-to-work laws, which both influence and reflect the level of union power, are included in the model.

However, right-to-work laws are defined at the State level. Therefore, when CZ effects are controlled for, as they are when we use our preferred method, only those counties in commuting zones that cross State lines contribute to the resulting estimates.[42] Hence, it is difficult to estimate the effects of this variable with precision.

In spite of this, a significant effect does emerge from the data. The presence of right-to-work laws had a significant positive effect on earnings growth in nonmetro counties, according to the preferred method. Our results suggest that earnings in rural counties located in right-to-work States grew by 5.25 percent more than other rural counties during the 1980's, all else equal. This may mean that mobile employers had a strong preference for locations in States with right-to-work laws, or that rural employers in States that permitted closed union shops fared less well under the competitive conditions of the 1980's. The presence of right-to-work laws may also serve as a proxy for other State policies and programs that served to attract or retain employers in rural areas during the decade.[43]

Other Labor Market Characteristics

The estimated effect of variation among counties in the labor force participation rate is near zero. This suggests that, at least in the 1980's, availability of a pool of potential additions to the labor force had little effect on business location and expansion decisions. If availability of a large pool of potential labor force additions did have any positive effects on employment growth, these effects may have been offset by adverse effects on wage levels, resulting in a negligible net effect on total earnings.

One other variable that may be viewed as a labor market characteristic did have a significant effect on earnings growth: the percentage of county employment in service occupations. Counties with a high concentration of service occupation workers experienced significantly slower earnings growth during the 1980's, even after taking into account our controls for differences in industrial mix and educational levels across counties. Each additional percentage point of 1979 employment in the service occupations reduced 1979-89 earnings growth by 0.44 percentage points (see Appendix C). Other occupational mix variables included in the model did not have a significant effect on earnings growth.

---

42. More precisely, because the analysis covers only nonmetro counties, only those nonmetro counties that were part of a commuting zone that included nonmetro counties from two or more States will contribute to the coefficient estimates for any variable measured at the State level. In all, there are 464 such counties in 43 States from 112 distinct commuting zones in our study population. However, only 178 nonmetro counties in 24 States belonged to one of the 44 commuting zones that included at least one nonmetro county in a right-to-work State and at least one nonmetro county in a non-right-to-work State. Thus, it is only these 178 counties that provide information on the size of the right-to-work effect when we use our preferred method.

43. While right-to-work laws are included in our model as one indicator of State union power, this variable is not simply a proxy for State unionization levels. When the latter were entered separately in a preliminary version of the model, right-to-work laws were still significant.

Population Educational Levels

While it is known that more educated individuals receive higher earnings, there is uncertainty, as discussed earlier (p. 10), about the likely influence of educational attainment in rural populations on local aggregate earnings. Our results suggest that areas that had a higher proportion of high school graduates in their adult populations experienced greater earnings growth (or less earnings decline) during the 1980's. They also suggest that educational attainment beyond high school, as measured by the proportion of college graduates in the adult population, provided no marginal benefits for earnings growth in rural areas.

High School Graduates

Results yielded by the preferred method suggest that differences in the percentage of the adult population that had completed high school had large and significant positive effects on earnings growth. The point estimate suggests that a difference of 10 percentage points in the fraction of the population that completed high school was associated with a corresponding difference of 3.3 percent in total 1979-89 earnings growth.[44] (This result is sensitive to the period over which growth is measured. Model estimates for alternate periods are available from the authors.)

College Graduates

The estimated effect of the concentration of college graduates in the adult population on earnings growth is not significantly different from zero, and the standardized value of the estimated effect is small.[45,46]

Other Education Factors

Other indicators of educational attainment exhibit no significant relationship to local area growth. We might note that our results suggest that the effect of the initial dropout rate for young adults on earnings growth was near zero. This is consistent with Killian and Parker's results for employment in nonmetro commuting zones, but contrasts with their result for metro commuting zones. We also find no evidence in the preferred-method results that the presence of substantial local higher education activity influenced earnings growth during the 1980's.

Local Fiscal Policies

Theoretical arguments and some past research suggest that, on one hand, higher levels of local taxation may have an adverse effect on earnings growth, while, on the other hand, higher levels of expenditure on desirable programs may augment local economic growth by enhancing productivity and by attracting individuals and businesses (Kusmin, 1994, pp. 21-36). Our results provide some support for the second of these hypotheses, indicating that higher levels of expenditure on public education were associated with greater earnings growth. However, local tax levels had no statistically significant effect on rural earnings growth during the 1980's.

---

44. This effect is statistically significant at the 1-percent level.

45. Note that the high school graduate variable that is used in this analysis includes those who are also college graduates. Thus, the estimated effect of college graduates represents the marginal effect of having college graduates in the population rather than high school graduates who have not completed college.

46. Due to collinearity between the high school completion and college completion variables, the standard errors associated with each of these coefficient estimates is relatively large (see Appendix B, p. 8).

Local Expenditures on Primary and Secondary Education

Results obtained with the preferred method suggest that higher levels of public education expenditures are conducive to higher earnings growth. A difference of $1,000 in annual per-pupil expenditures was associated with a difference of 3.77 percentage points in growth over the 10-year period from 1979 to 1989. This result is sensitive to the period over which growth is measured. [47]

Local Taxation Levels

The preferred-method estimates show no significant association between 1977 local tax levels and local economic growth, when local educational spending and other variables are held constant. This result may indicate that the effect of local tax levels on earnings growth was weak, or that any effect was offset by the positive effects of spending (other than education spending) that was financed by local taxes. There may also be limitations in the data or estimation techniques that obscure any underlying relationship.

Access to Transportation

Our results tend to confirm the expectation that greater access to transportation systems was positively associated with earnings growth during the 1980's. It appears that areas with more access to long-distance transportation were proportionately better able to attract or retain economic activity (and/or improve earnings per job) during this period. Results obtained with the preferred method indicated significant positive associations between earnings growth and both the number of major highway interchanges and the presence of an airport with scheduled passenger service within 50 miles.

Highways

Counties in which there was an interstate or comparable limited-access highway with one or more interchanges (exits) experienced significantly more earnings growth during the 1980's than other nonmetro counties after other differences were accounted for. The preferred method indicated that each interchange within a county was associated with approximately 0.42 percentage point in additional county earnings growth during the period.[48]

Airports

Counties that had access to an airport with scheduled passenger service experienced significantly more earnings growth during the decade. Access to such an airport, within 50 miles of the county, accounted for 3.36 percentage points in additional earnings growth.

---

47. Model estimates for alternate periods are available from the authors.

48. This effect is significant at the 1-percent level.

Other Access Factors

The estimated effect of an airport with scheduled service within the county itself was modest and not statistically significant.[49]

The effect of having an intersection of two interstate highways within the county was also small and not statistically significant. Because there were few nonmetro counties with such intersections (only 47 of the 2,346 studied), it is difficult to estimate the latter effect with much precision, but in any case this scarcity implies that these intersections can explain little of the variance in nonmetro growth rates.

Banking and Business Structure

Some past work suggested that the vigor of small and independent firms may be an important determinant of local economic growth, and may have played a substantial role in growth in recent decades (see Bluestone and others, 1989, and additional work cited there). Accordingly, several variables measuring the relative importance of small business in major sectors of the economy are included in our model as potential influences on local earnings growth.

Small Business Presence

The preferred-method results suggest that counties with a larger concentration of small businesses in their goods-producing sector experienced modestly slower growth during the 1980's than counties where the goods-producing sector was dominated by larger firms. The results indicate that a county could expect a difference of 1.08 percentage points in earnings growth over the decade if 80 percent rather than 70 percent of all county goods-producing businesses were small businesses. The result falls just short of statistical significance at the 5-percent level. This result may reflect the greater vulnerability of small firms to the adverse economic climate that affected much of rural America through much of the 1979-89 period.[50]

The estimated effect of the small-business share within the producer services sector is positive, and is similar in magnitude to the negative effect of small goods-producing businesses described just above, but is not statistically significant at conventional levels. The estimated effect of the small-business share in the other services sector is small and not significant.

Branch Banking Laws

Some past research has suggested that variations in local financial markets might influence local economic development. As a partial test of this hypothesis, a dummy variable is included in the model to capture the extent to which branch banking was permitted in each State in 1979. (Many States liberalized their branch banking laws in the years after 1979, but in most States, the laws in effect in 1979 had been in effect for a number of years.)

---

49. Note that, since another variable captures the effect of an airport anywhere within 50 miles (which includes in-county airports), this is the marginal benefit from having such an airport within the county, rather than having one outside the county but within 50 miles.

50. Note that these estimates reflect controls for differences in industrial mix. A county with a concentration of small firms in relatively prosperous industries might well have experienced faster growth as a result of the presence of those firms, but our analysis suggests that this should be attributed to the demand for the outputs of those industries, and not to any benefits of small firm size *per se*.

The regression results yield a negative effect of liberalized banking laws on earnings growth, but the estimated effect does not approach statistical significance. As noted in the section on right-to-work laws (see p. 36), the effects of variables defined at the State level cannot be estimated with much precision when we control for fixed commuting-zone effects.

Amenities and Access to Metropolitan Areas

Amenities

Two indexes of natural amenities are included in the model.[51] The topographical amenity index, designed to give positive weight to both mountainous terrain and relatively low elevation (see pp. 14-15) has a negative association with earnings growth that is statistically significant at the 10-percent level. That is, after controlling for other differences, areas that are at higher elevations or flatter had somewhat more earnings growth than areas that are lower or more mountainous. This result suggests that natural characteristics that may have a positive effect on individual migration decisions will not necessarily have a corresponding effect on the location of economic activity.

The index of water coverage, designed to reflect access to bodies of water that may have recreational value, did not have a significant effect on earnings growth.

Proximity to Metro Populations

An indicator of proximity to concentrations of metropolitan population is also included in the model. Contrary to expectations of a positive relationship that would reflect growth of nonmetro areas on the urban fringe and the benefits of access to larger markets, this variable had no significant association with earnings growth. The estimated effect is small and negative.

This result suggests that the positive correlation between metro adjacency and growth found by McGranahan and Salsgiver (1993) and others is not a direct effect of adjacency. Instead, it may be accounted for by differences in industrial mix or other characteristics that are associated with adjacency, but are included separately in our model.

Characteristics of the Local Economic Base

County earnings growth will reflect to some extent the prosperity of industries prominent in the local economic base, as well as other characteristics of that base. Our report indicates significant associations between earnings growth and transfer payments per capita as well as the share of total employment in each of 21 industries. However, industrial diversification was not significantly associated with earnings growth in the preferred-method estimates.

Transfer Payments

We did not anticipate the large and highly significant inverse relationship found between transfer-payment recipiency and earnings growth. A difference of $100 in transfers per capita was associated with a 1.57-percentage-point difference in cumulative growth.

---

51. Because variation in climate within a commuting zone is unlikely to be great enough to have measurable economic effects in most instances, the index of climate quality described earlier (see p. 14) is not included in the model as estimated by the preferred method.

Our measure of transfer-payment recipiency may be serving as a proxy for area where poor economic conditions have led to selective out-migration of people attached to the labor force, · leaving behind concentrations of transfer recipients. Social Security and retirement benefits make up the largest share of U.S. transfer payments. Areas where Social Security and retirement benefit recipiency are high are likely to be those with high concentrations of the elderly. If such concentrations do not reflect retiree in-migration, which is largely captured by our inclusion of a distinct retirement attraction variable in the model, then they may well reflect out-migration of younger people,[52] which in turn may reflect and contribute to less economic dynamism.

Selective out-migration of persons attached to the labor force may also leave behind disproportionate numbers of welfare recipients, who would similarly contribute to high per capita transfer recipiency rates.[53] Further, a large population collecting public welfare and unemployment insurance may directly reflect weakness in economic opportunities.

To the extent that the conditions leading to out-migration (or leading to elevated need for welfare benefits or unemployment insurance) were conditions that: (a) persisted into the 1980's, and (b) are not captured by other variables in this model, their effects may be reflected in the negative coefficient on transfer-payment recipiency. Similarly, any unmeasured conditions that were selectively attractive to migrants in the labor force during both the 1970's and 1980's would tend to reduce per capita transfer payment rates while augmenting earnings growth.

Finally, the main recipients of transfer payments (the elderly and disabled, welfare recipients, and the unemployed) are all populations likely to require above-average social service expenditures. Businesses may seek to avoid locations where larger shares of their taxes must go to pay for such services.

Concentration of Employment by Industry

Variables are included in the model that correspond to the share of total county employment in each of 75 industries, as discussed elsewhere (see pp. 16 and 50). The industries are defined at approximately the two-digit SIC level. (See discussion of data on page 16.)

We do not report in detail in this report on the effect that employment concentrations in particular industries had on earnings growth; these results are given in Appendix C. But the local mixture of employment by industry was an important determinant of county earnings growth. As noted earlier (see p. 29) the independent variables included in the growth model account for 26.3 percent of intra-commuting-zone variance in earnings growth rates, after adjusting for degrees of freedom. Without sectoral employment variables, this falls to 18.5 percent. [54]

Further, 21 of the sectoral employment variables are individually statistically significant, and another 6 are significant at the 10-percent level. Sectors significantly associated with more rapid

---

52. However, differences in current and past mortality and fertility rates will also influence concentration of the elderly.

53. However, concentrations of the elderly probably play a larger role in explaining the distribution of transfer payments. In 1983, retirement-related payments made up 69.4 percent of all transfer payments in nonmetro areas; veterans' benefits, including military retirement payments, accounted for a further 8.1 percent. In contrast, income maintenance payments and unemployment insurance together accounted for only 15.9 percent of nonmetro transfer payments in that year (Bentley, 1988).

54. These results apply to estimates using the preferred method and fixed weights that are based on the full model, that is, the model that includes sectoral employment variables.

growth in total county earnings include printing and publishing, transport services, real estate, hotels and lodging, miscellaneous business services, miscellaneous repair services, educational services, miscellaneous services, and State and local government. Positive coefficients significant at the 10-percent level include those for holding companies and Federal military employment. Sectors associated with significantly slower growth include farming, forestry, coal mining, metal mining, oil and gas extraction, heavy construction, lumber and wood products manufacturing, primary metals manufacturing, electrical machinery manufacturing, railroads, trucking and warehousing, and social services. Negative coefficients significant at the 10-percent level include those for stone, clay, and glass manufacturing, nonelectrical machinery manufacturing, wholesale trade, and Federal civilian employment.[55]

These results indicate that, when all other factors are held constant, local economies concentrated in the right industrial sectors will do better than average, although the sectors associated with favorable growth trends in the future may not be the same as those associated with growth in the 1980's.

Industrial Diversification

We had no firm expectations with regard to the likely effect of industrial diversification on earnings growth, as there are theoretical arguments for either a positive or negative effect. After we controlled for the employment shares of individual industries, the estimated effect of diversification on earnings growth in the preferred model was positive, but not statistically significant. While it is sometimes suggested that smaller communities must specialize to be successful, these results do not support that conclusion.

Comparison of Results Across Estimation Procedures

As discussed earlier (see p. 25), the use of a regression method that controls for fixed commuting-zone effects has both value and limitations. Fixed commuting-zone controls may capture a great deal of variation in omitted variables that could otherwise bias coefficient estimates. Statistical tests show that inclusion of commuting-zone controls reduces unexplained variance in growth rates substantially and significantly. However, use of these controls also makes it very difficult to generate precise estimates of the effects of State-level variables.

In addition, while statistical tests support use of corrections for heteroskedasticity and for non-normal errors, corrections other than those used in this report are possible. Hence, it may be useful to consider results from a model without such corrections, to assess how much the particular corrections we have used may influence the results.

Accordingly, in this section the preferred-method results discussed in the previous section are compared with the results for three alternative estimation procedures, described earlier (see p. 26).

---

55. The large absolute (unstandardized) magnitude of several estimated coefficients, such as those for transport services, real estate, and holding companies, suggest that some sectoral employment variables may be acting as proxies for unmeasured county characteristics (real estate also has a large standardized effect). Further, the significant negative effect associated with heavy construction may also require some careful interpretation, rather than being viewed as simply a sectoral effect. Employment in heavy construction is highly cyclical, and at the level of individual nonmetro counties, it is also likely sensitive to the beginning and completion of individual construction projects. In counties where construction employment levels were high in 1979, it is likely that earnings were well above their equilibrium levels. As the projects accounting for those high employment levels were completed, total county earnings would be likely to fall. Nonetheless, this result highlights the inability of economic development efforts focused on new construction to sustain economic growth.

The OLS results reflect estimation without control for heteroskedasticity, non-normality, or fixed effects. The robust-weights results reflect an estimation procedure that does control for heteroskedasticity and non-normality, but has no control for commuting-zone effects. Finally, the rest-of-CZ model assumes that county earnings growth is influenced by growth in the other counties in the same commuting zone, but does not use fixed effects to account for all inter-CZ variation in earnings growth rates (table 6).

Demographic Characteristics

The preferred-method finding that concentrations of African-American population were associated with slower earnings growth during the 1980's is reinforced by the results for the other methods. However, the estimated magnitude of this effect is only about half as great when fixed commuting-zone effects are not taken into account.

This suggests that in commuting zones where some counties had concentrations of African-Americans, the commuting zone as a whole experienced more earnings growth than otherwise expected, even though those counties where African-Americans lived experienced less earnings growth. However, we have not explicitly tested this hypothesis.

In contrast, while the estimated retirement county dummy coefficient is large and statistically significant in all four sets of estimates, the estimated effect is much larger in the absence of fixed commuting-zone effects.[56] While the preferred-method estimate is that retirement county status accounted for 4.50 percentage points of additional 1979-89 growth, the alternative estimates range from 6.79 to 9.55 percentage points.

The preferred-method estimate might plausibly understate the size of this effect. Because this method compares only counties within the same commuting zone, the estimated effect reflects a comparison of retirement counties with other nearby counties that may, in many instances, possess similar amenities. In contrast, methods without fixed commuting-zone effects compare retirement counties with other nonmetro counties throughout their region and the Nation. As noted earlier (p. 34), these retirement county results may reflect both the demand for goods and services by retirees, and in-migration of businesses and working-age individuals who were drawn to the same things that attract retirees.

Two other demographic variables have statistically significant effects under the assumptions of one or more of the alternative methods. The preferred-method coefficient on the concentration of Hispanic population is negative, but is small and does not approach statistical significance. However, when fixed commuting-zone effects are dropped, the estimated effect associated with this variable is somewhat larger and the associated parameter standard errors are lower. This effect is statistically significant in the OLS and rest-of-CZ estimates, and the robust-weights estimate is significant at the 10-percent level. It may be that the preferred method is unable to identify this effect as significant because the extent of intra-CZ variance in Hispanic population shares is limited relative to inter-CZ variance.

The evidence is weaker for a significant effect from the prime-aged population share. Only the OLS method generates a positive effect that is large or statistically significant. The rest-of-CZ and robust-weights methods yield smaller parameter estimates that are significant only at the 10-percent level, while the preferred-method estimate is near zero. The signs of all four coefficient

56. In general, if an effect is described in this report as larger than another, this does not imply that it is large (that is, with a standardized coefficient over 0.05 in absolute value). Similarly, a smaller effect is not necessarily small.

Table 6--Comparison of parameter estimates for preferred method and selected alternative methods

Dependent variable for all methods shown is 1979-89 real earnings growth.

Variable	OLS estimates			WLS bounded-influence estimates ("robust weights")			WLS bounded-influence estimates with rest-of-CZ effect			Preferred-method estimates		
	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic
Intercept	162.6	NA	3.84 ***	209.0	NA	5.63 ***	192.4	NA	5.55 ***	NA	NA	NA
Demographic characteristics:												
Total urban population (1,000)	0.0318	0.015	0.56	-0.0096	-0.004	-0.27	-0.0095	-0.004	-0.29	-0.0263	-0.0122	-0.58
Percent African-American	-0.0981	-0.058	-1.86 +	-0.0836	-0.049	-2.08 *	-0.0921	-0.054	-2.39 *	-0.1920	-0.1127	-3.01 **
Percent Hispanic	-0.1547	-0.064	-2.63 **	-0.0928	-0.038	-1.82 +	-0.1112	-0.046	-2.54 *	-0.070	-0.0290	-0.66
Retirement county ▒	9.55	0.149	7.35 ***	8.481	0.132	8.65 ***	6.792	0.106	7.32 ***	4.50	0.0702	3.85 ***
Percent of population aged 25-64	0.477	0.053	2.12 *	0.344	0.038	1.92 +	0.308	0.034	1.77 +	0.119	0.0132	0.52
Labor market characteristics:												
Mean annual earnings (log)	-20.47	-0.153	-4.76 ***	-24.51	-0.183	-6.37 ***	-21.53	-0.160	-5.99 ***	-23.54	-0.1755	-5.25 ***
Right-to-work law ▒	1.67	0.032	1.23	0.24	0.005	0.24	-0.628	-0.012	-0.64	5.25	0.1010	2.26 *
Labor force participation rate	0.034	0.009	0.30	-0.0010	-0.000	-0.01	-0.0345	-0.009	-0.39	-0.039	-0.0097	-0.33
Education levels and activity:												
Percent high school graduates	-0.069	-0.033	-0.65	-0.0877	-0.042	-1.04	-0.0861	-0.041	-1.09	0.331	0.1571	2.61 **
Percent college graduates	-0.359	-0.060	-1.48	-0.103	-0.017	-0.52	0.012	0.002	0.06	-0.173	-0.0290	-0.71
Percent dropouts (aged 16-19)	0.1298	0.038	1.51	0.1269	0.038	1.79 +	0.077	0.023	1.17	-0.0266	-0.0079	-0.33
Local college enrollment ▒	3.28	0.041	1.89 +	3.93	0.050	3.31 ***	2.40	0.030	2.04 *	-0.43	-0.0054	-0.32
Local taxes and expenditures:												
Local tax level	0.116	0.014	0.59	-0.129	-0.016	-0.71	-0.005	-0.001	-0.03	-0.084	-0.0101	-0.40
Education spending per pupil	4.60	0.078	3.00 **	4.30	0.073	3.31 ***	3.72	0.063	2.99 **	3.77	0.0640	2.39 *
Transportation access:												
Highway interchanges	-0.005	-0.001	-0.03	0.086	0.010	0.75	0.189	0.022	1.75 +	0.425	0.0496	3.23 **
Highway intersections	3.17	0.017	1.00	3.07	0.017	1.59	2.14	0.012	1.03	0.27	0.0015	0.11
Airport in county ▒	-0.30	-0.003	-0.15	-0.62	-0.006	-0.50	-0.12	-0.001	-0.10	0.71	0.0070	0.48
Airport within 50 miles ▒	3.06	0.038	2.17 *	4.18	0.052	3.81 ***	2.77	0.034	2.63 **	3.36	0.0417	2.26 *

continued--

See notes at end of table.

Table 6--Comparison of parameter estimates for preferred method and selected alternative methods--continued

Variable	OLS estimates			WLS bounded-influence estimates ("robust weights")			WLS bounded-influence estimates with rest-of-CZ effect			Preferred-method estimates		
	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic
Business and banking structure:												
Small businesses as percent of all businesses in--												
Goods-producing industries	-0.2567	-0.105	-4.38 ***	-0.1575	-0.064	-3.19 **	-0.1327	-0.054	-2.84 **	-0.1076	-0.0439	-1.94 +
Producer service industries	0.2911	0.068	3.01 **	0.2677	0.062	3.36 ***	0.2088	0.049	2.83 **	0.1298	0.0303	1.45
Other service industries	0.0612	0.039	2.12 *	0.0594	0.038	2.48 *	0.0254	0.016	1.12	-0.0188	-0.0120	-0.73
Branch banking law	3.70	0.049	2.28 *	2.11	0.028	1.78 +	-0.30	-0.004	-0.27	-2.62	-0.0350	-0.98
Amenities:												
Climate quality index	1.338	0.098	2.84 **	1.745	0.128	4.71 ***	1.144	0.084	3.26 **	NA	NA	NA
Topography (mountainousness) index	-0.555	-0.025	-1.00	0.084	0.004	0.21	0.087	0.004	0.23	-1.321	-0.0605	-1.86 +
Water coverage index	0.523	0.020	0.92	0.883	0.034	2.07 *	0.589	0.023	1.44	-0.600	-0.0229	-1.07
Relationship to metro areas:												
Population of metro areas within 50 miles (in millions)	0.215	0.007	0.37	-0.392	-0.013	-0.97	-0.501	-0.017	-1.40	-0.301	-0.0100	-0.49
Economic base:												
Transfer payments ($1,000 per capita)	-13.42	-0.101	-4.16 ***	-17.39	-0.132	-6.70 ***	-16.70	-0.126	-6.94 ***	-15.67	-0.1186	-4.67 ***
Industrial diversification index	0.0461	0.030	0.98	0.0829	0.053	2.34 *	0.0848	0.055	2.54 *	0.0629	0.0406	1.48

See notes at end of table.

continued--

Table 6--Comparison of parameter estimates for preferred method and selected alternative methods--continued

Variable	OLS estimates			WLS bounded-influence estimates ("robust weights")			WLS bounded-influence estimates with rest-of-CZ effect			Preferred-method estimates		
	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic	Regression coefficient	Stdized. coefficient	T-statistic
Region:												
New England ▒	23.57	0.117	6.00 ***	19.90	0.098	7.65 ***	13.34	0.066	4.99 ***	NA	NA	NA
Middle Atlantic ▒	6.69	0.041	2.09 *	8.83	0.054	4.08 ***	5.91	0.036	3.14 **	NA	NA	NA
East North Central ▒	0.82	0.011	0.49	2.74	0.036	2.22 *	1.48	0.019	1.31	NA	NA	NA
West North Central ▒	-8.05	-0.132	-6.16 ***	-5.95	-0.098	-6.34 ***	-4.90	-0.081	-5.47 ***	NA	NA	NA
South Atlantic ▒	8.94	0.129	z	8.74	0.126	z	6.29	0.091	z	NA	NA	NA
East South Central ▒	2.91	0.037	1.73 +	3.48	0.045	2.82 **	2.50	0.032	2.22 *	NA	NA	NA
West South Central ▒	-1.47	-0.021	-0.99	-2.51	-0.036	-2.18 *	-1.39	-0.020	-1.24	NA	NA	NA
Mountain ▒	-2.69	-0.032	-1.27	-6.32	-0.076	-3.52 ***	-3.99	-0.048	-2.23 *	NA	NA	NA
Pacific ▒	-2.69	-0.019	-0.80	-7.30	-0.052	-2.79 **	-3.28	-0.023	-1.38	NA	NA	NA
Rest-of-CZ growth	NA	NA	NA	NA	NA	NA	0.3044	0.237	15.29 ***	NA	NA	NA
Isolated county ▒	NA	NA	NA	NA	NA	NA	-1.37	-0.010	-0.51	NA	NA	NA

z = This statistic cannot readily be computed when the modified Suits method is used to compute coefficients for variables related by additivity constraints.
NA = Not estimated or not applicable.
+ = Statistically significant at 10-percent level.
* = Statistically significant at 5-percent level.
** = Statistically significant at 1-percent level.
*** = Statistically significant at 0.1-percent level.
▒ = 0, 1 variable.

Source: U.S. Department of Agriculture, Economic Research Service.

estimates are, however, consistent with the hypothesis that a relatively large prime-aged population is conducive to earnings growth.

None of the methods indicates an association between urban population share and earnings growth that approaches significance.

Labor Market Characteristics

The preferred method and the three alternatives considered are highly consistent in identifying wages as a highly significant variable that had a strong inverse effect on earnings growth. Further, there was little variation in the estimated magnitude of the wage effect: a 10-percentage-point difference in initial wage levels was estimated to shift cumulative earnings growth by 2.05 to 2.45 percentage points across the four methods.

This result does not imply that areas with a low-wage, low-quality labor force had a competitive advantage during the 1980's. Rather, it indicates that after adjusting for labor force education, as well as the industrial and occupational mix of area employment, lower initial wages were a significant factor associated with subsequent earnings growth.

One should also note that the additional earnings growth yielded by this effect would not be nearly enough to compensate for the initial difference in economic well-being associated with a difference in 1976-78 wage levels. If two counties had the same number of workers and other characteristics in the late 1970's, except that the wage level and, hence, total earnings were 10 percent higher in the second, total earnings in the latter county would still have been 7 to 8 percent higher in 1989, despite slightly more rapid growth in the initially poorer county.

Results for the right-to-work effect are less consistent. Only the preferred method yielded a statistically significant effect. The alternative methods generated estimated effects that were small, not significant, and inconsistent in sign. These results suggest that the effect of right-to-work laws on growth may have been concentrated in counties near the boundaries between right-to-work States and other States, operating primarily to influence the location of businesses in cases where both right-to-work and non-right-to-work sites were available within the same labor market. As noted earlier (p. 36, footnote 42), the preferred-method estimates of the effect of State-level variables reflected only commuting zones that straddle State borders.

Estimated effects of the labor force participation rate were near zero for all four sets of results.

Educational Levels and Activity

The preferred-method estimates reveal a strong and significant positive relationship between high school completion rates and earnings growth. However, the corresponding coefficients for the other three methods are small, not significant, and negative. These results suggest, somewhat surprisingly, that local high school completion rates affected the distribution of earnings growth within a commuting zone, but had little effect on the distribution of earnings growth among commuting zones. Alternatively, those States and commuting zones where high school completion rates were high may have had characteristics not captured in the model that depressed earnings growth in the 1980's. (Unpublished results available from the authors also indicate that the preferred-method estimate of the high school completion effect is highly sensitive to the period over which growth is assessed.)

While the dummy variable for substantial in-county college enrollments has a near-zero effect in the preferred-method estimates, it has a significant positive effect in all of the alternative estimates

(although the OLS parameter estimate is significant only at the 10-percent level). The alternative estimates suggest that a county with such enrollments experienced 2.40 to 3.93 percentage points in additional earnings growth over the 1979-89 period. However, the failure of the preferred-method estimates to yield a similar figure suggests that either: (a) the earnings effects of higher education activity in a county tend to spill over into the other counties within a commuting zone, and so are revealed only by inter-commuting zone comparisons, or (b) counties with substantial higher education activity tend to be found in commuting zones that share some other unspecified characteristic(s) that stimulated earnings growth in the 1980's.

The marginal effect of college completion rates is not significant under any of the four methods, nor is the high school dropout rate among young adults, except for the robust-weights estimate of the dropout effect, which is positive and statistically significant at the 10-percent level.

Local Taxes and Expenditures

All four methods yield large and statistically significant positive estimates of the effect of local education spending on earnings growth rates. Furthermore, the four parameter estimates are similar in magnitude, with the estimated effect of a $1,000 difference in per-pupil spending on 1979-89 growth ranging from 3.72 to 4.60 percentage points for the four estimates.

The estimated effect of local tax levels is small and does not approach significance under any of the four methods.

Transportation Access

While the preferred-method estimates indicate that the number of highway interchanges in a county is a highly significant predictor of 1979-89 earnings growth, the alternative method estimates do not confirm this relationship. (The rest-of-CZ estimate of this effect, which is about half the magnitude of the preferred-method estimate, is significant at the 10-percent level.) Much as noted above for the high school completion effect, such results may imply that highway access primarily affects the distribution of earnings growth within commuting zones. Alternatively, while highways contributed to earnings growth at the county level, those commuting zones where some counties were well served by interstate highways tended to have unmeasured characteristics that diminished their earnings growth during the 1980's.

All four methods are, however, consistent in indicating that an airport within 50 miles had a modest but significant positive effect on earnings growth in the 1980's. The presence of such an airport increased earnings growth by 2.77 to 4.18 percentage points.

None of the four methods revealed a statistically significant effect from the presence of an airport within the county, or from having an intersection of interstate highways within the county.

Business and Banking Structure

The three alternative methods indicate that the small-business share of goods-producing establishments had a negative effect on 1980's earnings growth; the preferred-method estimate fell just short of significance at the 5-percent level. However, the OLS estimate, which suggests that a 10-percentage-point difference in this share corresponded to an inverse 2.57-percentage-point difference in 1980's earnings growth, was considerably larger in magnitude than the others. Corresponding values for the other three methods ranged from 1.08 to 1.58 percentage points.

Corresponding results for the small-business share of producer service establishments tend to suggest that a high proportion of small businesses in this sector was conducive to subsequent earnings growth. While the preferred-method estimate fell short of statistical significance, the other three estimates were statistically significant and were somewhat larger in magnitude. The estimated effect of a 10-percentage-point difference in this variable ranged from 1.30 percentage points of additional earnings growth for the preferred-method estimates to 2.91 percentage points for the OLS estimates.

Results are more mixed for the small-business share of other service establishments. While OLS and robust-weights estimates are both positive and statistically significant, the rest-of-CZ estimate is not significant, and the preferred-method estimate is not significant and is opposite in sign.

There is limited evidence for an effect from State branch banking laws. Only the OLS estimate was statistically significant at the 5-percent level, while the robust-weights estimate was significant at the 10-percent level. Further, both the preferred-method estimate and the rest-of-CZ estimate were opposite in sign from the OLS estimate.

Amenities and Proximity to Metro Areas

The climate quality index was not included in the model as estimated by the preferred method, because of the limited extent of intra-CZ climate variation. However, all three alternative methods provided evidence for a large and significant positive effect of climate quality on earnings growth. The OLS and rest-of-CZ estimates were significant at the 1-percent level, and the robust-weights estimate was significant at the 0.1-percent level. A difference of 1 unit in this index (which corresponds to a 1-standard-deviation difference in any one of the four component measures) was estimated to change earnings by 1.14 percentage points for the rest-of-CZ growth model (which is likely to understate the total effect of the climate variable, since rest-of-CZ growth rates will also be influenced by regional climate) to 1.74 percentage points for the robust-weights method. Note that each of these models also include dummy variables for the census divisions (New England, Middle Atlantic, South Atlantic, and other divisions). (See page 51 for a discussion of these effects.) Hence, the results just described indicate that differences in climate quality were associated with differences in earnings growth within these regions, and not merely between regions.

Estimates of the effects of the topography and water coverage indexes and the size of the metro population within 50 miles were inconsistent in sign and generally not significant.[57]

Economic Base

The negative relationship of per capita transfer payments and earnings growth was consistently strong and significant at the 0.1-percent level across all four methods. The estimated effect of a one-standard-deviation difference in the level of per capita transfers ranged from 2.62 to 3.40 percentage points. As suggested on page 41, the negative association of growth and transfer payments may reflect adverse economic conditions in some areas that led both to pre-1979 out-migration by labor force participants and to poor earnings growth during the 1980's.

---

57. The robust-weights estimate of the water coverage effect was positive and statistically significant. As noted in the previous section (see p. 40), the preferred-method estimate of the topography-index effect was negative and significant at the 10-percent level.

Although the effect of industrial diversification was not statistically significant in the preferred-method estimates, the robust-weights and rest-of-CZ estimates of this effect were positive and significant.[58] The estimated effect of a one-standard-deviation difference in this index ranged from 0.77 percentage points in earnings growth for the OLS coefficient to 1.41 percentage points for the rest-of-CZ coefficient. (A one-standard deviation increase in industrial diversification is equivalent to shifting 8.33 percent of total county employment from industries that are overrepresented in the county to industries that are underrepresented.)

Further research might seek to clarify the nature of the effects of industrial diversification, perhaps probing whether the alternative-method results reported here might be explained by the effect of diversification within particular parts of the local economy.

Industrial Mix Effects

The effects of individual industry employment share variables are reported in detail in Appendix C. Industry share variables that had positive and significant associations with earnings growth for all four methods include transport services, real estate, hotels, miscellaneous business services, education services, Federal military employment, and State and local government.[59] Industries with significant negative effects across the four methods include forestry; metal mining; oil and gas extraction; coal mining; heavy construction; lumber and wood products; stone, clay, and glass products; primary metal manufacturing; nonelectrical machinery; electrical machinery; railroads; and wholesale trade.[60] Six additional industries had statistically significant preferred-method effects: farming, printing and publishing, trucking, miscellaneous repair services, social services, and miscellaneous services.[61]

Industry effects that were both large and statistically significant across all four methods include the negative effects associated with metal mining, oil and gas extraction, coal mining, and railroads, as well as the positive effects associated with real estate and hotels.

Thus, as noted earlier, the industrial focus of a county does make a difference in terms of its expected growth. However, it is likely that some of the industries associated with growth in the 1990's will be different from those associated with growth in the 1980's. For example, while the Federal military sector was associated with improved earnings in the 1980's, it is reasonable to suspect that this same sector will be associated with decreased earnings during the military downsizing of the 1990's.

---

58. As noted in Appendix D, the industrial diversification effect is also statistically significant in a fixed-effects model where backward elimination has been used to shorten the list of independent variables.

59. Significant only at the 10-percent level were the robust-weights estimate of the miscellaneous business services effect and the preferred-method estimate of the military employment effect.

60. Eight of the 48 coefficient estimates for these 12 industries were significant only at the 10-percent level. These include the preferred-method estimates for stone, clay and glass products, nonelectrical machinery, and wholesale trade.

61. The printing and publishing effect was negative and significant for three of the four methods, with OLS the only exception. The social services effect was negative and significant for the preferred and rest-of-CZ methods, and significant at the 10-percent level for the robust-weights method.

Region and Rest-of-CZ Effects

While the preferred method incorporates controls for otherwise unexplained regional effects at the commuting zone level, the alternative methods incorporate more limited controls at the census division level. Parameter estimates associated with these census division controls indicate that there were large differences in regional performance not accounted for by the other variables in the model.

New England nonmetro counties grew much faster than comparable counties nationwide. The estimated effect of location in New England ranged from 13.34 percentage points in earnings growth for the rest-of-CZ method to 23.57 percentage points for the OLS method. Middle Atlantic nonmetro counties also grew relatively rapidly. Earnings growth in these counties was 5.91 percentage points greater than expected in the rest-of-CZ estimates, 6.69 points in the OLS estimates, and 8.83 points in the robust-weights estimates. And location in the East South Central States had a more modest but still significant positive effect, with estimates ranging from 2.50 to 3.48 percentage points (the OLS parameter was significant only at the 10-percent level).

On the other hand, West North Central nonmetro counties grew much more slowly than expected, even after their other characteristics were taken into account. The estimated effect of a West North Central location ranged from -4.90 percentage points in earnings growth for the rest-of-CZ method to -8.05 percentage points for the OLS method.

Estimated effects for other regions were less consistent across methods. However, the robust-weights method yielded significant positive effects for East North Central locations, and significant negative effects for West South Central, Mountain, and Pacific locations. (However, the point estimates of the East North Central and West South Central effects were small in absolute, as well as standardized magnitude. This suggests that the growth performance of nonmetro counties in these regions during the 1980's was largely driven by county and State characteristics, and not by any overriding regional trends.) The Mountain States effect was also significantly negative using the rest-of-CZ method.

Because of the data transformation applied to dummy variables in this report and described in Appendix B (pp. 12-13), the statistical significance of the coefficients for the South Atlantic division were not calculated. However, it appears that South Atlantic counties grew substantially faster than average; estimated coefficients for this region ranged from 6.29 to 8.94 percentage points.

In general, one would expect the rest-of-CZ estimates to understate these regional effects, since the rest-of-CZ growth rates are themselves influenced by these effects. Therefore, the robust-weights estimates may be a better guide to the power of unmeasured regional characteristics in the determination of earnings growth.

As noted above, including a rest-of-commuting-zone growth rate in the model had a substantial influence on parameter estimates. The rest-of-CZ growth rate itself was highly significant and had a large effect on earnings growth. The results indicate that, all else equal, each additional 1 percentage point in earnings in neighboring counties was associated with 0.30 percentage point in additional growth in the affected county. (The isolated-county variable is included in the rest-of-CZ model to allow for the fact that a county with no significant commuting ties is not equivalent to a county where neighboring counties experienced zero earnings growth. However, the difference between these two cases turns out not to be statistically significant.)

Conclusions

The large quantity of data brought to bear in this analysis has permitted clear identification of some of the important factors influencing nonmetro county earnings growth during the 1980's (table 7).

Major Influences on Earnings Growth

A handful of county characteristics were found to be statistically significant when we used the preferred method to estimate the growth model and also under all three of the alternative methods that we considered. Our findings with respect to these variables may be viewed as the most reliable positive findings of our report:

a. Past success in attracting retirees had a consistently strong and significant positive effect on earnings growth.

b. Higher initial wage levels had a significant negative effect on earnings growth in rural areas during the 1980's.

c. A consistently significant and strong positive relationship was observed between earnings growth and per-pupil public education spending in nonmetro areas.

d. Access to an airport with scheduled passenger service had a significant, with all four methods, positive effect on earnings growth.

e. All four models indicate that the small-business share of goods-producing establishments had a significant negative effect on earnings growth.[62]

f. Transfer payments per capita had, with all four methods of estimation, a strong and significant negative effect on earnings growth.

Cross-method comparisons also tend to support the preferred-method results indicating a significant negative association between areas of African-American population concentration and earnings growth.[63]

Further, most of the industry employment share variables identified as statistically significant by the preferred method were also significant under the alternative methods. Overall, the industrial composition of local employment had a strong influence on earnings growth, as did the share of employment in service occupations, which was negatively associated with growth.

---

62. Although the preferred-method coefficient is significant only at the 10-percent level.

63. The OLS estimate of this effect is significant at the 10-percent level; the other two alternative estimates were significant at the 5-percent level.

Table 7--Comparison of parameter estimates for preferred method and selected alternative methods, summary of results

Dependent variable for all models shown is 1979-89 real earnings growth.

	Significant coefficients				Large coefficients			
	Alternative			Preferred	Alternative			Preferred
Variable	OLS	WLS	Rest-of-CZ		OLS	WLS	Rest-of-CZ	
Demographic characteristics:								
Total urban population	N	N	N	N	N	N	N	N
Percent African-American	N	Y	Y	Y	Y	N	Y	Y
Percent Hispanic	Y	N	Y	N	Y	N	N	N
Retirement county ▒	Y	Y	Y	Y	Y	Y	Y	Y
Percent of population aged 25-64	Y	N	N	N	Y	N	N	N
Labor market characteristics:								
Mean annual earnings (log)	Y	Y	Y	Y	Y	Y	Y	Y
Right-to-work law ▒	N	N	N	Y	N	N	N	Y
Labor force participation rate	N	N	N	N	N	N	N	N
Education levels and activity:								
Percent high school graduates	N	N	N	Y	N	N	N	Y
Percent college graduates	N	N	N	N	Y	N	N	N
Percent dropouts (aged 16-19)	N	N	N	N	N	N	N	N
Local college enrollment ▒	N	Y	Y	N	N	N	N	N
Local taxes and expenditures:								
Local tax level	N	N	N	N	N	N	N	N
Education spending per pupil	Y	Y	Y	Y	Y	Y	Y	Y
Transportation access:								
Highway interchanges	N	N	N	Y	N	N	N	N
Highway intersections	N	N	N	N	N	N	N	N
Airport in county ▒	N	N	N	N	N	N	N	N
Airport within 50 miles ▒	Y	Y	Y	Y	N	Y	N	N
Business and banking structure:								
Small businesses as percent of all businesses in--								
Goods-producing industries	Y	Y	Y	N	Y	Y	Y	N
Producer service industries	Y	Y	Y	N	Y	Y	N	N
Other service industries	Y	Y	N	N	N	N	N	N
Branch banking law ▒	Y	N	N	N	N	N	N	N

See notes at end of table.

continued--

Table 7--Comparison of parameter estimates for preferred method and selected alternative methods, summary of results--continued

Dependent variable for all models shown is 1979-89 real earnings growth.

	Significant coefficients				Large coefficients			
	Alternative			Preferred	Alternative			Preferred
Variable	OLS	WLS	Rest-of-CZ		OLS	WLS	Rest-of-CZ	
Amenities:								
Climate quality index	Y	Y	Y	NA	Y	Y	Y	NA
Topography (mountainousness) index	N	N	N	N	N	N	N	Y
Water coverage index	N	Y	N	N	N	N	N	N
Relationship to metro areas:								
Population of metro areas within 50 miles	N	N	N	N	N	N	N	N
Economic base:								
Transfer payments	Y	Y	Y	Y	Y	Y	Y	Y
Industrial diversification index	N	Y	Y	N	N	Y	Y	N
Region:								
New England ▒	Y	Y	Y	NA	Y	Y	Y	NA
Middle Atlantic ▒	Y	Y	Y	NA	N	Y	N	NA
East North Central ▒	N	Y	N	NA	N	N	N	NA
West North Central ▒	Y	Y	Y	NA	Y	Y	Y	NA
South Atlantic ▒	NA	NA	NA	NA	Y	Y	Y	NA
East South Central ▒	N	Y	Y	NA	N	N	N	NA
West South Central ▒	N	Y	N	NA	N	N	N	NA
Mountain ▒	N	Y	Y	NA	N	Y	N	NA
Pacific ▒	N	Y	N	NA	N	Y	N	NA
Other:								
Rest-of-CZ growth	NA	NA	Y	NA	NA	NA	Y	NA
Isolated county ▒	NA	NA	N	NA	NA	NA	N	NA

▒ = 0, 1 variable.
NA = Not applicable/not estimated.
Y = Yes (significant/large).
N= No (not significant/not large).
OLS = Ordinary least-squares.
WLS = Weighted least-squares (with bounded-influence adjustment) (robust weights).
Rest-of-CZ = Weighted least-squares with bounded-influence adjustment and control for growth rate in remainder of commuting zone (CZ).

Source: U.S. Department of Agriculture, Economic Research Service.

In addition, results from the three alternative methods suggest a strong positive relationship between climate quality and earnings growth. This is a relationship that could not be effectively tested using the preferred method.[64]

Unmeasured factors that differ across space were clearly also important influences on growth. There were substantial and significant differences in earnings growth by census division, even after individual county characteristics were taken into account. When commuting-zone effects were taken into account, capturing the effects of any variations across commuting zones in unmeasured factors, the percentage of variance in growth explained by the model rose from 39.31 to 55.34. These are adjusted r^2 values for models without correction for heteroskedasticity or non-normal residuals, and computed after dropping 120 isolates (counties that are the only nonmetro county in their commuting zone) from the data. When preferred-method weights were used, the corresponding values are 67.00 and 51.51. Overall, the adjusted r^2 value for the preferred model, calculated with weighted or unweighted data, falls by more than a fourth if commuting-zone effects are dropped from that model, even when census division dummies and the climate index are retained to capture larger-scale inter-regional differences.

Variables with Little Influence on Earnings Growth

We also find that a few particular county characteristics clearly were not important contributors to local area earnings growth. Presence of an in-county airport did not have much effect on the distribution of earnings growth.[65] Nor did proximity to metro population concentrations.[66] Intersections between interstate highways also played at most a minor role in explaining the variation of nonmetro earnings.[67]

Variables with Uncertain Influence on Earnings Growth

In many other instances, the data are less clear. For these variables, our results are sensitive to the econometric methods used to derive them, the exact specification of our model, or the stringency of our statistical tests. Alternative methods fail to confirm preferred-method results that indicate significant positive earnings effects associated with right-to-work laws and high school completion rates. Only limited alternative-method support is found for the preferred-method results that showed that access to highway interchanges was associated with faster earnings growth. These preferred-method results may reflect processes that influence the distribution of growth within commuting zones but not among commuting zones.

---

64. We count as major influences those noncontrol variables identified as statistically significant by the preferred method, and as significant at least at the 10-percent level by all four methods. In addition, we count the effect of the small-business share of goods-producing industries, which fell just short of the threshold for significance at the 5-percent level in the preferred-method estimates, but was large and significant for all of the alternative methods. As noted in the text, we also include climate, although the effect of climate could not be assessed using the preferred method.

65. Note that the model already controls for the presence of an airport within 50 miles.

66. The latter is perhaps a surprising result, in light of observed rapid growth in some rural areas on the perimeters of existing metro areas (McGranahan and Salsgiver, 1993). It suggests that such growth may reflect other factors included in our model and associated with proximity to a metro area, such as access to nearby air service and to major highway systems, and perhaps the quality and effectiveness of local school systems, rather than distance from metro populations *per se*.

67. These are effects not large or significant for any of the four methods considered, and for which the 95-percent confidence interval around the preferred-method estimates does not include a range of large values.

On the other hand, results for the alternative methods provide some evidence for effects that are identified as smaller and not significant by the preferred-method results. These may be real effects that spill over into other counties within a commuting zone and, hence, are difficult to detect using the preferred method, or they may be apparent effects that arise because the variables are serving as proxies for other unmeasured variables for which the commuting-zone effects provide better controls.

Effects identified as statistically significant for earnings growth by at least two of the three alternative methods, but not by the preferred method, include a negative relationship with Hispanic population share, a positive relationship with local college enrollments, a positive relationship with the small-business share of the producer services sector, a positive relationship with the small-business share of the other services sector, and a positive relationship with industrial diversification.

While our best estimates indicated that these variables did not have a substantial effect on growth in the 1980's, alternative econometric models yield different conclusions. Hence, any future work with these data that uses econometric techniques not considered here might usefully attempt further assessment of the significance and magnitude of these effects.

Results for one alternative model also suggest that growth in any one county may have substantial and significant spillover effects in other counties in the same commuting zone, but unmeasured factors that have a common effect across the commuting zone could also explain this result.

Other variables were not statistically significant by any of the methods applied, yet the confidence intervals for the preferred-method coefficients associated with these variables do not allow us to reject the possibility that they actually did have relatively large effects on earnings growth. (See Appendix E for the confidence intervals for the preferred-method estimates.) Variables fitting this description include urban population, labor force participation,[68] college completion rate, young-adult high school dropout rate, local tax level, and the local topography index. The percentage of the population aged 25 to 64, the 1979 branch banking law variable, and the index of coverage by bodies of water might also be grouped with these variables, although the OLS estimates for the first two effects and the robust-weights estimate for the third effect are significant at the 5-percent level.

In summary, while we examined the likely effects on rural earnings growth of a large number of key variables, we were unable to settle all the questions posed by our model.

We have shed much light on the question of what factors contribute to earnings growth in rural counties, but further work will be needed to better respond to the concerns of rural development practitioners and policymakers.

---

68. Apparently, any advantages that might be associated with a large pool of potential additional workers were offset by unmeasured negative characteristics associated with a low labor force participation rate, or perhaps by the depressing effect such a pool might exert on wage growth. This result may reflect, in part, the limited extent of average rural earnings growth and the widespread occurrence of rural decline during the 1980's. Real earnings declined in more than half of the 2,346 study counties between 1979 and 1989, while only about 30 percent of these counties experienced growth of 10 percent or more. The distribution of potential workers might affect the location of growth during a period of widespread economic expansion, but the availability of potential workers seems less likely to have influenced the distribution of closures and down-sizing in an era characterized largely by retrenchment and declines in earnings and employment in much of rural America.

Comments and Suggestions for Further Research

Some of the results of our analysis were unexpected and may suggest the need for further study. We found that areas where transfer payments were high experienced diminished earnings growth. This may reflect persistent characteristics that led to out-migration of labor force participants and, hence, boosted the density of transfer recipients, but other possible explanations might also be explored. Our results also suggest that rural counties with concentrations of African-Americans experienced diminished growth, but do not tell us the reason for this outcome. All else equal, areas where the goods-producing sector of the economy was dominated by small firms did not fare as well during the 1980's, despite the emphasis in some past work on the vitality of smaller firms. On the other hand, under some assumptions, industrial diversification appears to have fostered growth in nonmetro counties.

The lack of statistically significant effects associated with certain variables, particularly local tax levels and the size of nearby metro populations, was also surprising. In addition, the strong negative relationship between service occupation employment and earnings growth was somewhat surprising, as were some of the industrial sector effects reported in Appendix C.

Limitations

Several limitations of the current report have been discussed in previous sections, and these should be noted again. Estimation of a single model at the national level precludes modeling of possible interregional differences in growth processes, such as those suggested by Quan and Beck (1987). This approach also made it infeasible to gather primary data on institutional or other variables that might have been of interest. Use of controls for commuting-zone fixed effects meant that effects could not be estimated accurately for variables whose effects tend to diffuse throughout a commuting zone, or variables with little intra-CZ variation. At the same time, there is still some potential for our results to be influenced by unmeasured intercounty differences. In most instances, we did not explore the possibility of nonlinear relationships between our independent variables and earnings growth. To reduce the likelihood that we might misidentify consequences of growth as determinants of growth, we also ruled out testing for the possible effects of contemporaneous change in other variables on earnings growth rates.

It is also important to note that the statistical uncertainty surrounding most of our coefficient estimates is great enough to render most of our conclusions tentative, even in the absence of any modeling or measurement errors, although this limitation is not peculiar to our report.

Finally, these results reflect the experience of one decade and one economic cycle. As supply-demand conditions, institutions, and other factors change, so will the relationship between local characteristics and earnings growth.

Directions for Future Research

Further research might yield more precise and reliable estimates of many of the effects considered in this report. Such research might reduce our uncertainty about the role of some county and State characteristics in fostering rural earnings growth, and might help to confirm, account for, or explain some of the unexpected effects described above.

More advanced modeling of spatial effects is one approach that might well prove useful, especially with regard to estimating the effects of State characteristics. Another strategy that might clarify some of the questions left open by this report would require compilation of extensive county-by-county longitudinal data to permit estimation of a panel data model. Such an approach would

control for all unchanging differences between counties, and would seek to explain the changes in growth rates for each county in successive periods in terms of changes in other county characteristics, such as tax levels and expenditures. Yet another approach might seek to estimate both own-county and rest-of-commuting-zone effects for each variable of interest, to capture likely spillover and backwash effects for individual county characteristics. (Because many commuting zones include both metro and nonmetro counties (in fact, there are only a handful of exclusively metro commuting zones), this would require either a narrower focus on exclusively nonmetro commuting zones, or compilation of all needed data for metro as well as nonmetro counties.)

The processes that lead to the overall earnings effects that we observed might be disentangled to some extent through estimation of distinct regression equations for employment growth and growth in earnings per job, perhaps in a simultaneous-equations framework. Population growth might also be included in such a multiequation model.

Within the framework of the current report, some refinements are also possible. These include use of more extensive data on county population age structure, information on the level and distribution of local noneducation expenditures, and a measure of earnings per job explicitly adjusted for differences in job and worker quality. The relationship between diversification and growth might be further explored through use of distinct diversification measures for major industrial sectors to assess whether, for example, a diverse trade sector has the same effect as a diverse manufacturing sector. The influence of nearby metropolitan areas on nonmetro growth might also be probed in more depth, through inclusion of nearby-metro growth rates or other characteristics among the potential factors associated with nonmetro growth. Some gains in accuracy might also be made (albeit perhaps at the cost of some generality) by estimating the model within subgroups of nonmetro counties and testing for differences among groups. For example, one might estimate the econometric model separately for each census division, or for metro-adjacent and nonadjacent counties.

While these and other approaches might address some of the questions left unanswered in this report, they are unlikely to resolve them all. The many local area characteristics that may influence earnings growth are related to each other in many ways. Even with a large set of data, it is difficult to disentangle their effects on growth or on each other. Further, the factors that influence growth in the 1990's and beyond may differ from those that influenced growth in the 1980's, with changes in worker characteristics, public services, or site characteristics demanded by business, as well as changes in the preferences and priorities of individual migrants whose choices also influence the location of growth.

In spite of the limitations of this report, we believe that we have made substantial progress in identifying some key factors that explain differences in earnings growth rates among rural counties.

References

Advisory Commission on Intergovernmental Relations. *Tax Capacity of the Fifty States: Methodology and Estimates*. Report No. M-134, Washington, DC, March 1982.

Bartik, Timothy J. "Small Business Start-Ups in the United States: Estimates of the Effects of the Characteristics of States," *Southern Economic Journal.* Vol. 55, No. 4 (April 1989), pp. 1,004-1,018.

Bartik, Timothy J. *Who Benefits from State and Local Economic Development Policies?*, W.E. Upjohn Institute for Employment Research, Kalamazoo, MI, 1991.

Beale, Calvin L., and Glenn V. Fuguitt. "Decade of Pessimistic Nonmetro Population Trends Ends on Optimistic Note," *Rural Development Perspectives*. Vol. 6, No. 3 (June-September 1990), pp. 14-18.

Bender, Lloyd D., Bernal L. Green, Thomas F. Hady, John A. Kuehn, Marlys K. Nelson, Leon B. Perkinson, and Peggy J. Ross. *The Diverse Social and Economic Structure of Nonmetropolitan America*. Rural Development Research Report No. 49, U.S. Dept. Agr., Econ. Res. Serv., September 1985.

Bentley, Susan E. *Transfer Payments and Investment Income in the Nonmetro United States*. Rural Development Research Report No. 71, U.S. Dept. Agr., Econ. Res. Serv., June 1988.

Bluestone, Herman, Celeste A. Long, and Shirley L. Porterfield. "Small Business Activity: Does It Make a Difference at the County Level?" Paper presented at the 28th Annual Meeting of the Southern Regional Science Association, Chapel Hill, NC, April 27-29, 1989.

Butler, Margaret A. *Rural-Urban Continuum Codes for Metro and Nonmetro Counties*. Staff Report No. AGES 9028, U.S. Dept. Agr., Econ. Res. Serv., April 1990.

Carlino, Gerald A., and Edwin S. Mills. "The Determinants of County Growth," *Journal of Regional Science*. Vol. 27, No. 1 (February 1987), pp. 39-54.

Conroy, Michael E. "The Concept and Measurement of Regional Industrial Diversification," *Southern Economic Journal*. Vol. 41, No. 3 (January 1975), pp. 492-505.

Dubin, Elliot. "Financing Rural Elementary and Secondary Education." Occasional Paper for the Appalachian Regional Education Laboratory, 1989.

Economic Policy Institute. "New EPI Book: Unions Are Key to Competitiveness Solution," *Economic Policy Institute Journal*. Vol. 2, No. 2 (July 1992).

Fox, William F. "Fiscal Differentials and Industrial Location: Some Empirical Evidence," *Urban Studies*. Vol. 18 (1981), pp. 105-111.

Freedman, David A., and Stephen C. Peters. "Bootstrapping a Regression Equation: Some Empirical Results," *Journal of the American Statistical Association*. Vol. 79, No. 385 (March 1984), pp. 97-106.

Glaeser, Edward L., Hedi D. Kallal, Jose A. Scheinkman, and Andrei Shleifer. "Growth in Cities," *National Bureau of Economic Research Working Paper*. No. 3787, National Bureau of Economic Research, July 1991.

Gracia-Diez, Mercedes. "Compositional Changes of the Labor Force and the Increase of the Unemployment Rate: An Estimate for the United States," *Journal of Business and Economic Statistics*. Vol. 7, No. 2 (April 1989), pp. 237-243.

Greene, William H. *LIMDEP: Version 5*, August 1988. Self-published.

Hady, Thomas F., and Peggy J. Ross. *An Update: The Diverse Social and Economic Structure of Nonmetropolitan America*. Staff Report No. AGES 9036, U.S. Dept. Agr., Econ. Res. Serv., May 1990.

Hallahan, Charles. "Introduction to the Bootstrap." Unpublished material presented at USDA, Economic Research Service Regression seminar, 1990.

Halstead, D. Kent. *Tax Wealth in Fifty States*. National Institute of Education, U.S. Department of Health, Education, and Welfare, 1978.

Hausman, J.A. "Specification Tests in Econometrics," *Econometrica*. Vol. 46, No. 6 (November 1978), pp. 1,251-1,271.

Judge, George G., R. Carter Hill, William E. Griffiths, Helmut Lütkepohl, and Tsoung-Chao Lee. *Introduction to the Theory and Practice of Econometrics*. New York: John Wiley & Sons, 1988.

Kassab, Cathy. "Studying Economic Change: Are Robust Regression Procedures Needed?" *Rural Sociology*. Vol. 55, No. 3 (Fall 1990), pp. 357-375.

Killian, Molly Sizer, and Timothy Parker. "Education and Local Employment Growth in a Changing Economy," *Education and Rural Economic Development: Strategies for the 1990's*. Staff Report No. AGES 9153, U.S. Dept. Agr., Econ. Res. Serv., September 1991.

Killian, Molly Sizer, John Redman, and David W. Sears. "Spatial Variations in the Growth and Stability of Rural Areas." Unpublished paper presented at the Eleventh Annual Research Conference of the Association for Public Policy Analysis and Management, Arlington, VA, November 1988.

Killian, Molly Sizer, and Charles M. Tolbert. "Mapping Social and Economic Space: The Delineation of Local Labor Markets in the United States," *Inequalities in Labor Market Areas*. (Joachim Singelman and Forrest A. Deseran, eds.), Westview Press, 1993.

Kusmin, Lorin D. *Factors Associated with the Growth of Local and Regional Economies: A Review of Selected Empirical Literature*. Staff Report AGES-9405, U.S. Dept. Agr., Econ. Res. Serv., March 1994.

Maddala, G.S. *Introduction to Econometrics*. New York: Macmillan Publishing Company, 1988.

McGranahan, David A. *Population Loss in Remote Rural Areas*. Agricultural Information Bulletin 664-70, U.S. Dept. Agr., Econ. Res. Serv., November 1993. Reprinted in *Issues for the 1990's*, Janet B. Stevens (ed.), Agriculture Information Bulletin 664, September 1994, pp. 153-154.

McGranahan, David A., and Linda M. Ghelfi. "The Education Crisis and Rural Stagnation in the 1980's," *Education and Rural Economic Development: Strategies for the 1990's*. Staff Report No. AGES 9153, U.S. Dept. Agr., Econ. Res. Serv., September 1991.

McGranahan, David A., and Jacqueline Salsgiver. "Recent Population Change in Adjacent Nonmetro Counties," *Rural Development Perspectives*. Vol. 8, No. 3 (June 1993), pp. 2-7.

McHugh, R.J., and J.T. Wilkinson. "A Random Effects Approach to Substate Growth Models: A Comment on 'The Determinants of County Growth,'" *Journal of Regional Science*. Vol. 28, No. 2 (May 1988), pp. 271-273.

Milkove, Daniel L., and Patrick J. Sullivan. *Deregulation and the Structure of Rural Financial Markets*. Rural Development Research Report No. 75, U.S. Dept. Agr., Econ. Res. Serv., February 1990.

Myers, Raymond B. *Classical and Modern Regression with Applications (2nd edition)*. Boston: PWS-KENT Publishing Company, 1990.

Newman, Robert J., and Dennis H. Sullivan. "Econometric Analysis of Business Tax Impacts on Industrial Location: What Do We Know and How Do We Know It?" *Journal of Urban Economics*. Vol. 23 (1988), pp. 215-234.

Plaut, Thomas R., and Joseph E. Pluta. "Business Climate, Taxes and Expenditures, and State Industrial Growth in the United States," *Southern Economic Journal*. Vol. 50 (July 1983), pp. 99-119.

Quan, Nguyen T., and John H. Beck. "Public Educational Expenditures and State Economic Growth: Northeast and Sunbelt Regions," *Southern Economic Journal*. Vol 54, No. 2 (October 1987), pp. 361-376.

SAS Institute Inc. *SAS User's Guide: Basics, Version 5 Edition*. Cary, NC: SAS Institute Inc., 1985.

Sears, David W., John M. Redman, Lorin D. Kusmin, and Molly S. Killian. *Growth and Stability of Rural Economies in the 1980s: Differences Among Counties*. Staff Report 9230, U.S. Dept. Agr., Econ. Res. Serv., December 1992.

Smith, Eldon D., Brady J. Deaton, and David R. Kelch. "Location Determinants of Manufacturing Industry in Rural Areas," *Southern Journal of Agricultural Economics*. July 1978, pp. 23-32.

Suits, Daniel B. "Dummy Variables: Mechanics v. Interpretation," *Review of Economics and Statistics*. Vol. 66, No. 1 (1984), pp. 177-180.

Tax Foundation. *Facts and Figures on Government Finance*. 21st Edition. Washington, DC, 1981.

Teixeira, Ruy A., and Paul L. Swaim. "Skill Demand and Supply in the New Economy: Issues for Rural Areas," *Education and Rural Economic Development: Strategies for the 1990's*. Staff Report No. AGES 9153, U.S. Dept. Agr., Econ. Res. Serv., September 1991.

Thompson, Wilbur R., and Philip R. Thompson. "National Industries and Local Occupational Strengths: The Cross-Hairs of Targetting," *Urban Studies*. Vol. 24 (1987), pp. 547-560.

Tolbert, Charles M., and Molly Sizer. *U.S. Commuting Zones and Labor Market Areas: A 1990 Update*. U.S. Dept. Agr., Econ. Res. Serv. (forthcoming).

U.S. Department of Commerce, Bureau of the Census. *Statistical Abstract of the United States, 111th edition*. Washington, DC, 1991.

U.S. Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Professions, Office of Health Professions Analysis and Research. *Area Resource File*.

Wasylenko, Michael. "Empirical Evidence on Interregional Business Location Decisions and the Role of Fiscal Incentives in Economic Development," in *Industry Location and Public Policy*. Henry W. Herzog, Jr. and Alan M. Schlottman (eds.), University of Tennessee Press, Knoxville, TN, 1991.

Appendix A. Sources of Data

Variable	Source
Growth index, 1979-1989	Bureau of Economic Analysis (BEA) County Income data file
Urban population, 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Population of nearby metro areas, 1980	*1983 City and County Data Book* (metro area populations) Adjacency and proximity based on coding done at ERS
Adult college attainment rate, 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Adult high school attainment rate, 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
High school dropout rate (young adults), 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Local college enrollment ratio, 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Percentage of population aged 25 to 64, 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Labor force participation measure, 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Percentage African-American (Black), 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Percentage Hispanic, 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Percentage of employment by occupation, 1980	1980 U.S. Census Summary Tape File 3 (STF 3)
Percentage of employment by industry, 1980	BEA County Employment data file: unpublished version provided to ERS with unsuppressed data on employment by industry for all industries and counties
Retirement county dummy variable	ERS analysis based on data from 1970 and 1980 U.S. Census. See Bender, *et al.* (1985); Hady and Ross (1990)
Transfer payments per capita	Bureau of Economic Analysis (BEA) County Income data file

Variable	Source
Locally raised general revenue as a percentage of local personal income (1972 and 1977)	Revenue data from 1972 and 1977 *Census of Governments*; local personal income from BEA County Income file
Local primary and secondary education expenditures per pupil, 1977	Expenditures from 1977 Census of Governments Finance Summary Statistics data file. Student population based on ERS estimates (Dubin, 1989).
Earnings per job, 1976-78 average	Total labor and proprietor income and total jobs from BEA County Income and Employment files.
Right-to-work law status, 1980	Communication from Employment Standards Administration, Department of Labor United States
Percentage of all establishments that are both small and independent: goods-producing sector; producer services sector; other services sector (1980)	Establishment and Enterprise Microdata File (USEEM), developed by the Small Business Administration from Dun and Bradstreet's Dun's Market Indicators file
Divided limited-access highway interchanges; interstate highway intersections	Unpublished tabulations done at ERS
Airports with scheduled passenger service: (a) in county (b) within 50 miles	Identified from information found in the *National Plan of Integrated Airport Systems (NPIAS) 1986-95*, published in November 1987 by the Federal Aviation Administration
State branch banking laws	Milkove and Sullivan (1990)
Index of climate quality	Area Resource File, Dept. of Health and Human Services
Index of topographical attractiveness	Area Resource File, Dept. of Health and Human Services (elevation); unpublished tabulations done at ERS (rugged terrain)
Index of water coverage	Census Bureau data on water coverage

Appendix B. Estimation Issues and Techniques

Multivariate Linear Regression

As noted in the text (p. 4), a linear model (of economic growth, for example) is one where the outcome of interest, such as the index of economic growth (G) in this report, responds to incremental differences in the level of various factors $X_1, X_{2,...,} X_K$ in a fashion that does not depend on the levels of G or the various X's.

Such a model of growth can be expressed in the form of an equation by

$$G = b_0 + b_1X_1 + b_2X_2 + ... + b_KX_K + u,$$

where, for example, b_2 represents the effect on growth of a difference of one unit in X_2 (which might be, for example, the percentage of the population with a college education). The term u represents variation in G that is not explained by any of the factors in the model.

When data are available on G and the various X's, the magnitude of b_1, b_2, etc. can be estimated using a linear regression model

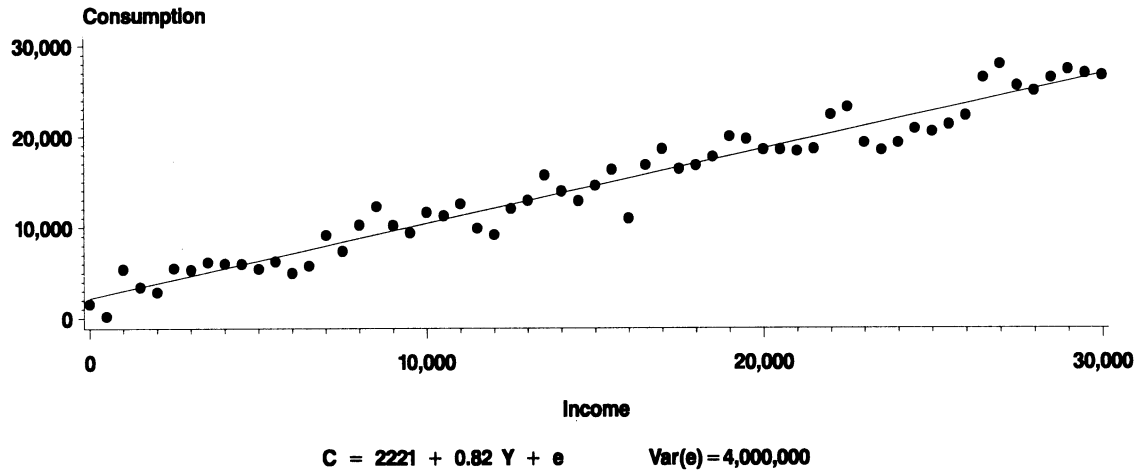
$$G = b_0 + b_1X_1 + b_2X_2 + ... + b_KX_K + e.$$

In this model, e is simply the discrepancy between whatever value of G is implied by all of the X's and the estimated b's (sometimes known as the fitted value of G) and the actual value of G; thus, its value differs for each county. This value e is often referred to as the error term or residual in the model.[1]

The variable on the left side of this equation (G in this example) is known as the dependent variable. The variables on the right side, the X's, which may influence the outcome of interest, may be known as the independent variables, explanatory variables, predictive variables, or regressors. The b's, which reflect the estimated relationship of the dependent variable to the regressors, may be known as the regression coefficients (or simply the "coefficients"); they may also be referred to as the parameters, or sometimes as the estimated parameters. The latter term emphasizes that, even if the model of growth that has been proposed is correct, the regression procedure yields only the best estimates of the values of each b, which need not be equal to the true values. An example of a linear regression model with only one independent variable is illustrated in figure B-1.

1. More precisely, the error term is the difference between the actual value and the fitted value that would be yielded by using the true coefficients (and is represented by u in the equation above), while the residuals are the differences between the actual value and the fitted values computed using the estimated coefficients (and these residuals are represented by e).

Figure B–1. Linear regression with one independent variable



$C = 2221 + 0.82\ Y + e$ $\quad Var(e) = 4{,}000{,}000$

Source: U.S. Department of Agriculture, Economic Research Service.

Data shown are hypothetical and for illustrative purposes.

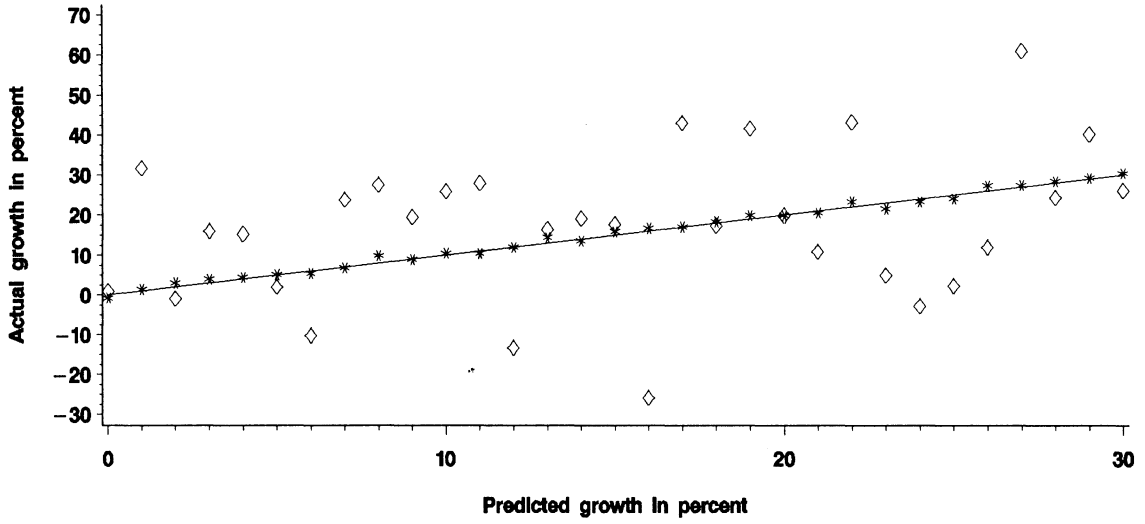
Ordinary Least-Squares (OLS)

The simplest method of regression, ordinary-least-squares (OLS), generates an estimate for each of the b's so that the sum of all the squared values of e is as small as possible. Thus, in estimating values of the b's using OLS, both positive and negative differences between the actual value of G and the value predicted by the model are penalized, and large "misses" are penalized much more than small ones.

Heteroskedasticity

As noted in the text (p. 23), if the variance of actual outcomes around their expected values is not the same for all observations, we have heteroskedasticity, as illustrated in figure B-2.

Figure B–2. Actual vs. expected growth for two types of counties



* * * = Type 1 county (low variance) ◇ ◇ ◇ = Type 2 county (high variance)
___ The sloped line connects points where actual growth equals predicted growth.

Note: Data shown are hypothetical and for illustrative purposes.

Source: U.S. Department of Agriculture, Economic Research Service.

For example, the growth rate for counties with small economies is likely to be more sensitive to the opening or closing of one or two employers, and hence will have a larger unpredictable component. Thus the regression model will not fit such counties as well. Perhaps more importantly, because the economic performance of such counties is more seriously affected by events that we can only model as random, their performance has less to tell us about the relationship between measurable variables and local economic growth. In this case, OLS may not be the best procedure to estimate the regression model. Instead, a regression technique that gives less weight to such counties is expected to yield better estimates of the underlying relationships.

Weighted Least-Squares (WLS) and Multiplicative Heteroskedasticity

The general term for an approach that assigns varying weights to observations in a regression model is weighted least-squares (WLS). Various techniques have been used to sort observations according to their expected variance in the context of heteroskedasticity. One of the more flexible, because it involves relatively nonrestrictive assumptions about the source and form of any differences in expected variance, is to assume multiplicative heteroskedasticity.

In this technique, the residuals are taken from an OLS estimate of the regression model. The logarithm of the squared values of these residuals is then regressed on variables that may be associated with heteroskedasticity. If there are clear *a priori* grounds to identify particular variables that might be associated with heteroskedasticity, these should be included in the regressor list at this step. (In this report, all variables in the growth model are treated as potential predictors of heteroskedasticity.) If particular observation characteristics are associated with a large variance of observed outcomes around

observed outcomes around the expected outcomes, then these observations will have positive coefficients in the resulting regression.[2] Fitted values from this intermediate regression are then used to compute weights for each observation, with lower weights assigned to cases in which the expected magnitude of the residual is greater.

This technique can correct for heteroskedasticity that may be associated with many characteristics of the observation, provided that these interact in a multiplicative fashion to determine the observation variance (Judge and others, 1988, pp. 365-371) or that their interaction can reasonably be approximated by the multiplicative form.

The overall significance of the regression used to explain variation in the logged squared residuals is one indicator of the presence of heteroskedasticity. In this report, this regression is significant at the 0.0001 level; this is true whether or not the residuals were generated by an econometric method that controlled for fixed effects.[3] While this does not imply that the heteroskedasticity-correcting weights generated by this technique are optimal, it does tell us that there is heteroskedasticity in the data, and that it is significantly associated with the variables in the growth model.

Robust Regression

As noted in the text (p. 23), OLS assumes that the differences between actual and expected outcomes follow a normal statistical distribution, and this also holds for WLS if the data are properly weighted. If this assumption does not hold, the results of OLS or WLS may be inefficient and misleading.

In particular, if the tails of the residual distribution (or the distribution of weighted residuals in the case of WLS) are "fatter" than the tails of a normal distribution, then OLS or WLS will give too much weight to those cases with large residuals.[4] Robust regression techniques are intended to

---

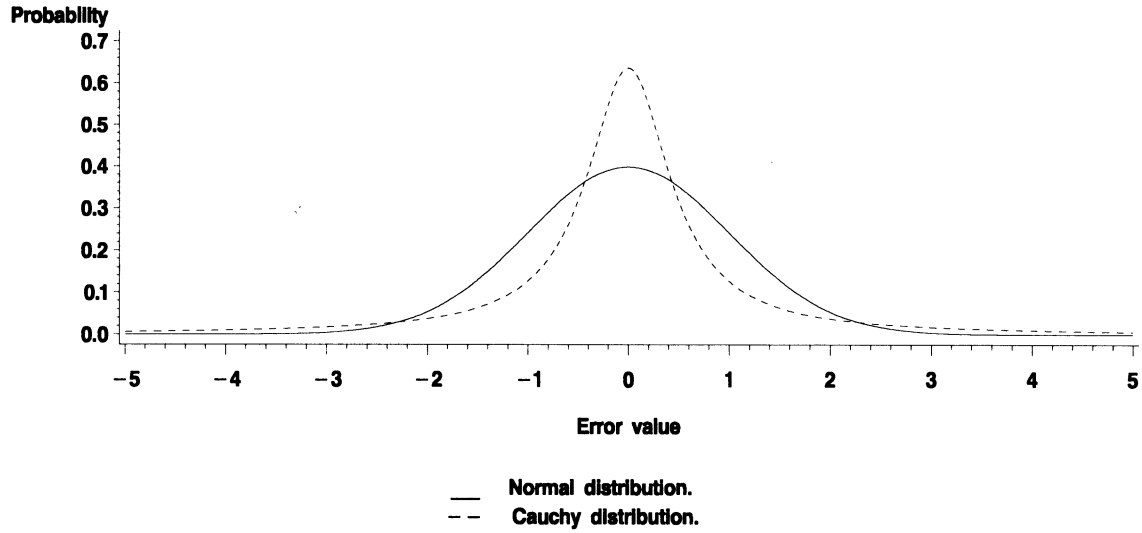
2. The usefulness of the technique is not dependent on the precision of the individual coefficient estimates in this intermediate regression; what is relevant is their joint effectiveness in assigning appropriate variances to individual observations.

3. With residuals from a simple OLS estimate of the growth model, the F-statistic associated with this heteroskedasticity-fitting regression is 2.714, with 116 and 2,229 degrees of freedom. When the residuals are generated by a method that controlled for fixed effects, the F statistic associated with the heteroskedasticity-fitting regression is 1.729, with 116 and 2,109 degrees of freedom. Both values are statistically significant at the 0.01-percent level.

4. The shape of the tails of the residual distribution indicates how rapidly the probability of observing a particular residual value falls off with the magnitude of that value. If this probability falls off rapidly with the magnitude of the residual (as is true for the normal distribution beyond a certain point), then we should use an estimation technique that gives a heavy weight to minimizing particularly large observations. OLS, which minimizes the sum of squared residuals, has this attribute. However, if relatively large residuals are not that unusual, as indicated by a distribution with fat tails, then less emphasis should be given to minimizing the largest residuals when the best model is estimated.

The distribution of residuals might have relatively fat tails because the underlying error process generates a substantial number of extreme values. If there are large errors in measurement, for some observations, of either the dependent variable or important independent variables, this might also generate some large residuals that have no implications for the correctness of the model.

Figure B–3. Two error distributions: Normal and Cauchy



Source: U.S. Department of Agriculture, Economic Research Service.

Data shown are hypothetical and for illustrative purposes.

diminish the weight given to those cases. Bounded-influence regression is one of those techniques.

Bounded-Influence Regression

In this report, bounded-influence regression is used to reduce the influence of unusual observations. This technique takes into account the values of the independent variables as well as the size of regression residuals in identifying cases that should be given a lower weight.

In this technique, an influence statistic $DFFITS_i$ is computed for each observation I in a regression. $DFFITS_i$ is "a measure of the standardized change in the least-squares fit when the ith case is deleted from the OLS analysis" (Kassab, 1990). An upper influence bound,

$$c = v\,[(J + 1) / I]^{0.5},$$

is set on the influence that any one observation can have on the results, where J equals the number of slope parameters in the model, I equals the number of observations, and v is a robustness parameter. The parameter v is usually set to 2.0, but can also be set to 1.5 or 1.0. Lower values of v mean that a reduced weight will be given to more cases, and that the reductions in the weights of these cases will be greater. (Where the bounded-influence technique is applied in this report, v is set equal to 2.0.)

The regression is then re-estimated using WLS, with lower weights being given to those observations whose influence exceed the influence bound. That is, less importance is attributed to

observations not consistent with the relationship of dependent and independent variables suggested by all other observations (Kassab, 1990; and Kassab, personal communication). If the regression has already been weighted to correct for heteroskedasticity, then the weights that are used in computing the final results are the products of the heteroskedasticity-correcting weights and those computed to satisfy the influence bound.

In our report, testing for non-normality of the (weighted) residuals from the multiplicative-heteroskedasticity-corrected econometric model yields a value of 0.06294 for Kolomogorov's D. For the fixed-effect, multiplicative-heteroskedasticity-corrected econometric model, the corresponding value is 0.08108. Both indicate that the hypothesis of normality can be rejected at the 1-percent level (SAS Institute, 1985, p. 1,187).

In this report, 6.3 percent of observations (148/2346) exceed the influence bound when fixed effects are not included in the computations, and 6.4 percent of observations (143/2226) when they are included. On average, the bounded-influence adjustment reduces the weight of these observations by about one-fourth.

Dummy Variables

Categorical variables such as region of the country, or presence/absence of an airport, are normally represented in regression analysis by dichotomous variables, commonly referred to as dummy variables. These are variables that take on only the values 0 or 1, depending on whether an observation satisfies a particular condition. A categorical variable with two categories (for example, airport vs. no airport) can be represented with 1 dummy variable (1 = airport, 0 = no airport); a categorical variable with n categories (for example, the nine census divisions) can be represented with n-1 independent dummy variables. Inclusion of a dummy variable for the nth category, for example, inclusion of separate 0,1 dummy variables for airport and no airport, would add redundant information to the list of independent variables.

Omitted Variables

As noted in the text (p. 24), omitted variable bias can lead to erroneous conclusions about the influence of particular variables. The problem of omitted variable bias can be illustrated by a simple example. Suppose that wages in a company depend on age, veteran status, and gender:

$$WAGE = b_0 + b_1*AGE + b_2*VETERAN + b_3*MALE + e,$$

with higher wages being paid to older workers, veterans, and males. Assume that males in this company are more likely than females to be veterans, and that males in this company are also older on average than their female colleagues. Now suppose that the coefficients of the following regression are estimated:

$$WAGE = b_0 + b_1*AGE + b_2*VETERAN + e.$$

Some of the difference in wages that is in reality a function of gender will be attributed in this regression to age and to veteran status, so that the estimated values of the coefficients on the latter two variables will be biased by the omission of the gender variable. In this example, both coefficients are biased upwards, because all three variables have a positive effect on wages, and because both age and veteran status are positively correlated with the omitted variable. In general, omission of a variable that should be included can also bias other coefficients downwards. The effects of omitted variable bias on the values of the remaining coefficients can be large if the

omitted variables: (a) have a large effect on the dependent variable, and (b) are highly correlated with some of the variables that are included in the model.

Multicollinearity

Another concern noted in the text (p. 24) is that of multicollinearity. If two variables are highly correlated with each other, they are said to be highly collinear. In this case, it may be quite difficult to distinguish the effects of one from the effects of the other. In the example discussed just above, if all males are veterans, and all veterans are male, it will be impossible to distinguish the effects of gender and veteran status. The two variables are then said to be perfectly collinear. Even if there are a few female veterans or male nonveterans, they may not provide enough information to clearly distinguish the effects of the two variables.

Even if no independent variable is highly correlated with any other single variable, multicollinearity may arise in a model with many independent variables when some variables are highly correlated with (linear) combinations of other variables. For example, in a population of professional males, it is likely that age is highly correlated with the sum of years of work experience and years of education. Thus, it might be difficult to distinguish the wage effects of age, years of education, and years of experience in a single model.

Multicollinearity will be reflected in relatively large standard errors and small t-statistics for the affected parameters. Thus, even relatively large effects may be poorly estimated or identified as not significant in the presence of multicollinearity. In addition, the adverse consequences of any errors in the measurement of the independent variables will be exacerbated when multicollinearity is present.[5] However, multicollinearity in itself does not lead to bias in the coefficient estimates.

Some analysts treat multicollinearity as a rationale for excluding variables from their model, even when theory argues for their inclusion. However, this is likely to lead to omitted variable bias. When variables that have a role in the model in principle are dropped because they have a high correlation with other variables that remain in the model, it is especially likely that the coefficients of those other variables will be biased.

In addition, if variables are omitted to reduce multicollinearity, the standard errors of the coefficients on the remaining variables will be lower than they should be, falsely suggesting that the effects of these variables are known with precision. For example, in the example given earlier, if veteran status and male gender are highly correlated, it may be that there is not enough data to tell which actually accounts for higher wages. However, if one of the two is dropped from the model to eliminate multicollinearity, it will falsely appear that the effect of the other is unambiguously significant.

Accordingly, in developing the models emphasized in this report, multicollinearity was not treated as a justification for leaving variables from the model.[6]

---

5. If one variable X_1 is highly collinear with some combination of other variables, the estimated coefficient for X_1 will reflect the small differences between X_1 and that combination. If X_1 has been measured with some errors, those errors may be large relative to the unique component of X_1, even if they are small relative to X_1 itself.

6. Except that in a few instances where multiple measures of the same underlying concept were available, only one was chosen. In addition, some variables that were included in preliminary versions of the model have been dropped from the version presented here. A brief discussion of these variables and the reasons for their exclusion is available on request from the authors.

Tolerances and Variance Inflation Factors

We do, however, compute tolerances and variance inflation factors (VIF's) for each of the variables included in the model as an indicator of the degree of multicollinearity associated with each variable. The tolerance for any one variable is equal to 1 minus the value of r^2 when that variable is regressed on all of the other independent variables in the model. The VIF is the reciprocal of the tolerance, and indicates by what factor the variance of the estimated coefficient error is greater than it would be for a variable with a comparable distribution that was uncorrelated with all of the other independent variables. These are reported in Appendix G.

These values are applicable to parameter estimates for the OLS and robust-weights methods only. It would be moderately difficult to compute the tolerances and VIF's for the preferred method; furthermore, it is not clear what interpretation could be given to the results. Many county characteristics are distributed in such a way that fixed commuting-zone effects might explain much of their variance (that is, county characteristics are correlated within commuting zones), so that their VIF's would be much higher if computed explicitly for the fixed effects model (where each fixed effect should be treated as a distinct independent variable). However, it seems implausible that this offers a rationale either for omitting potentially significant variables from the model or for disregarding the power of unmeasured inter-CZ differences that are captured by fixed commuting-zone effects. In general, while an examination of the tolerances may offer some insight into why a variable thought to be important fails to show a statistically significant effect, it will not provide a rationale for omitting conceptually distinct variables from the model.

The computed VIF's are less than 3 for most of the noncontrol variables in the model. Those with higher values include the high school attainment rate (9.95 to 11.22), college attainment rate (6.47 to 8.56), mean annual earnings (4.01 to 4.70), percentage African-American (3.74 to 4.64), labor force participation rate (3.21 to 3.69), and climate quality index (4.62 to 4.55), as well as urban population (2.75 to 3.07) and industrial diversification (3.59 to 2.77).[7] It may be useful to note that several of these variables are identified as statistically significant in the text despite these relatively high VIF's.

While VIF's for many of the control variables are also moderate, they are higher in a majority of cases, and very high in some instances (the most extreme instance is the VIF for farming employment, which exceeds 100). This should not be surprising, in light of the relationships among these variables: if not for the omitted category assigned to each group of control variables, any one variable in the group could be perfectly predicted by the others, and the associated VIF's would be infinite. This is one reason that we emphasize the role of these variables as controls, and give less emphasis to their individual coefficients in the text.

Model Developed Using Backward Elimination of Variables

As noted in the text (p. 24), the authors employ a backward elimination procedure to derive an alternative, simpler fixed-effect backward elimination econometric model that reduces the extent of multicollinearity without arbitrarily dropping statistically significant variables from the model. Backward elimination drops variables with the least explanatory power from the model until only statistically significant variables are left. Stepwise regression (which was also applied to a preliminary version of our model, with similar results) starts with a single variable, then adds variables with the most explanatory power until no variable excluded would be significant if added (meanwhile dropping variables that are nonsignificant after others are added). The resulting models

7. The first and second values given are for unweighted and weighted data, respectively.

represent the smallest models that can be constructed without excluding some statistically significant variable.

This was done primarily to assess the seriousness of the effects of multicollinearity on the results of the preferred methods and, in particular, to assess whether the larger standard errors associated with multicollinearity lead to a failure to identify many potentially significant influences on growth. This econometric model also reflects controls for fixed commuting-zone effects, heteroskedasticity and for non-normality. Results for this model are presented and discussed briefly in Appendix D. Results from this econometric model suggest that multicollinearity has had only a modest influence on our results.

Fixed Spatial Effects

If there are unmeasured influences that affect local economic growth, vary across geographic areas, and are correlated with other variables included as regressors in the model, then coefficient estimates from any of the techniques described above will be biased and inconsistent. An econometric method that incorporates fixed effects for distinct geographic areas may control for such influences.[8] Such an approach assumes a common unmeasured effect across all observations in a single area; this is equivalent to a model that has a dummy variable for each distinct area.

This bulletin reports results for estimates of the earnings growth model with and without fixed commuting-zone effects (see Commuting zones boxed text). We also compare the fixed-effects results against results that do not account for such effects.

Commuting zones

A commuting zone is a cluster of contiguous counties grouped together on the basis of commuting patterns; thus a commuting zone may be viewed as being a regional labor market. Commuting zones may contain both metropolitan and nonmetropolitan (rural) counties. In all, the 3,069 counties in the United States (exclusive of Alaska) could be grouped into 763 commuting zones (based on 1980 Census data); individual commuting zones in the United States contained anywhere from one to 22 counties, with an average of 4.02 counties. (Single-county commuting zones, also known as commuting isolates, are counties without strong commuting ties to any other county.) Of these 763 commuting zones, 746 include one or more of the 2,346 rural counties that were covered in our report. See Killian and Tolbert (1993) for more on the definition and identification of commuting zones. Commuting-zone identifiers were taken from a data file provided by Molly Sizer Killian, formerly of ERS. (An update of commuting-zone boundaries that reflects 1990 data has been done, and a report on this update is currently in draft form (Tolbert and Sizer, forthcoming.))

In this report, fixed commuting-zone effects are found to be highly significant statistically, and hence are taken into account in our preferred method.[9]

8. Fixed-effect techniques are also commonly used in the study of time-series-cross-section (panel) data, where data for the same individuals or jurisdictions are available for two or more points in time. In these cases, a fixed effect is estimated for each individual or jurisdiction.

9. With 120 commuting zones that had only a single nonmetro county excluded from the analysis, so that 2,226 counties and 626 commuting zones remain, F-statistics for the joint significance of the remaining commuting zones ranged from 2.20 (if the data are not weighted) to 2.57 (if weights are used that correct for heteroskedasticity and non-normality in the context of fixed effects.) Both values are far above the threshold of statistical significance for an F-statistic with 625 and 1,484 degrees of freedom. (These values reflect the inclusion of commuting-zone effects in a model that also includes census division dummies and a climate variable. In the model presented in the text, the latter variables are dropped when commuting-zone effects are included.) With no weighting of the data, the adjusted proportion of variance explained (adjusted r^2) rises from 0.393 to 0.553 when fixed commuting-zone effects are added to the OLS model. If preferred-method weights are used, the corresponding values are 0.515 and 0.669.

When fixed effects are assumed, all differences between geographic areas can be explained in terms of the fixed effects.[10] Thus, it becomes much more difficult (impossible) to estimate parameters for regressors with little (no) variation within individual regions, so that, for example, the effects of State characteristics could not be estimated if fixed State effects were assumed.

Random Effects

In this report, commuting-zone effects are modeled as fixed effects. An alternative random effects approach exists that imposes the assumption that the regional effects are uncorrelated with the independent variables in the model, as well as the assumption that those effects are normally distributed. This technique has been termed the "error components" technique. It can also be viewed as a statistical model in which the errors for each observation (county) have two components: a normally distributed disturbance e_i, with variance σ_i^2, which is independent from one observation (county) to the next, and a second normally distributed disturbance u_k, with variance σ_k^2, which is the same for all observations (counties) in group (geographic area) k, but is independent across groups. Both error terms are independent of the independent variables in the model.

It has been noted in Judge and others (1988) that the parameters estimated by the random effects technique are a matrix-weighted average of the fixed-effect parameter estimates and the estimates produced by the between estimator, which uses only data on differences between geographic areas in the mean values for the dependent value and regressors. This will permit (more efficient) estimation of parameters for variables that vary entirely (primarily) between geographic areas (Judge and others, 1988, pp. 479-491; Greene, 1988), relative to the fixed-effect approach. McHugh and Wilkinson (1988) use this approach to estimate a model of the determinants of county growth.

However, the assumptions that geographic area effects are independent of the regressors and are normally distributed are restrictions that must be tested. An appropriate test has been developed by Hausman (Hausman, 1978; Judge and others, 1988). When the random effects technique was applied to a preliminary version of the model of county growth used in this report, application of the Hausman test clearly indicated that the assumptions required to justify use of the random effects technique were not valid. Hence, we do not report random effects results.

Bootstrapping

As noted in the text (p. 30, footnote 36), the multiplicative heteroskedasticity and bounded-influence regression techniques used in this report require estimation of observation weights from the underlying data. These estimates are then treated as fixed when they are used to compute the final results. This strategy may lead to more efficient coefficient estimates if the underlying assumptions are correct. However, the resulting standard error estimates will be incorrect, and are likely to understate the true standard errors.

Bootstrapping techniques are appropriate under these conditions (Freedman and Peters, 1984; Kassab, 1990). Under the bootstrapping approach, the process whereby a sample is drawn from a population, or whereby one realization of a stochastic process is drawn from the population of possible outcomes, is simulated by using random sampling with replacement from the original sample or study population to create a series of B bootstrap samples of the same size as the

10. As noted earlier (p. 25, footnote 33), the other estimated parameters may explain some inter-regional differences, but inter-regional differences will not contribute any information to the estimation of those parameters.

original. The procedure of interest is run on each of these samples, and the empirical distribution of the B sets of parameter estimates is used to estimate the parameter standard errors.[11]

Under the random regressor approach, observations are drawn randomly from the study population or sample to construct each bootstrap sample. This is appropriate when the original data represent a random sample from a larger population. The distribution of the dependent variable and of each regressor will be slightly different in each bootstrap sample.

Under the fixed regressor bootstrapping approach, residuals from the regression of interest are sampled to generate a bootstrap sample, rather than entire observations. The bootstrap samples are created by replacing the original values of the dependent variable with synthetic values; the latter are generated by adding to the predicted value of the dependent variable for each observation a randomly drawn residual.[12] Thus, all the values of the independent variables are the same in each bootstrap sample when the fixed-regressor approach is used (Freedman and Peters, 1984; Hallahan, 1990). The fixed regressor approach is the one that was tested for this report.

Bootstrapping was used during the course of this report to estimate unbiased standard errors for the robust weights coefficient estimates for a preliminary version of the model. We found that, on average, the bootstrap standard error was 8.3 percent greater than the conventional standard error. The difference between the two ranged from -9.8 to 20 percent of the conventional values for individual coefficients, with a standard deviation of 5.5 percent.

Programming constraints precluded the use of bootstrapping with the preferred method. For this reason and because the differences between the two sets of standard errors for the robust-weights method were modest, bootstrap techniques have not been used to compute standard errors for the final model estimates presented in this report, and only conventionally computed standard errors are reported.

Lagged Earnings Growth

If there are unmeasured differences between counties that influence growth and persist over time, then earnings growth in the previous period might have a statistically significant effect on earnings growth in the study period, if a measure of such lagged earnings growth were included in the model. In addition, past growth or decline may have effects that tend to lead to continued growth or decline. Failure to account for such lagged growth effects could bias estimates of the effect of other county characteristics, if those characteristics are correlated with past growth rates. However, a model of earnings growth that includes lagged growth rates may understate the total effect on current growth of measured county characteristics that persist over time, if part of that effect is an indirect one operating through past growth.

In light of these considerations, the effects of including lagged earnings growth in the model have been tested using the preferred method and three alternative methods. The variable used was logged earnings change from 1969 to 1978. (The terminal year 1978 has been chosen, rather than

---

11. If the distribution of parameter estimates is not normal, confidence intervals for the estimates may be computed from the complete empirical distribution of parameter estimates, rather than from their standard errors alone. However, we found no evidence of significant departures from normality in the distribution of bootstrap parameter estimates that were done for this report.

12. The procedure is somewhat altered when heteroskedasticity is present, and the expected magnitude of the residuals is dependent on other characteristics of the observation. In this case, standardized residuals are drawn randomly, and are then inflated by an observation-specific expected variance.

1979, to prevent bias arising from errors in the measurement of 1979 earnings.) The effects are notably modest. In the preferred-method estimates, the coefficient on lagged growth is less than 0.013, and does not approach statistical significance. The OLS estimates of the same effect are even closer to zero. In the rest-of-CZ model, the coefficient on lagged growth is somewhat larger at 0.048 and is statistically significant, but inclusion of the lagged growth variable in the model has only modest effects on other parameter estimates or t-tests. Only when the robust-weights method is used do we find the assessed significance of some variables of interest changed by inclusion of the lagged growth variable. Even in this last case, the estimated effect of lagged growth on current growth is only 0.073, indicating a very moderate degree of persistence in growth trends between the two decades, and the changes in other parameter estimates when lagged growth is included are not dramatic.[13]

In view of these results, we do not report detailed results for models incorporating a lagged growth measure in this report.

Transformation of Selected Variables

As noted in the text, when categorical variables are included in a regression model, they are represented by dummy variables. When a categorical variable has several possible values, one category must be omitted to identify the model. In many cases, the choice of the omitted category is arbitrary, but the interpretation of the results may be influenced by that choice.

Suits (1984) proposed an alternative, whereby the sum of all categorical effects is constrained to equal zero, so that each effect can be interpreted as a deviation from the average of all effects. Gracia-Diez (1989) describes a transformation of the data that imposes this constraint implicitly. For a variable with J categories, and assuming that the last is the category to be omitted, the 0,1 dummy variables D_1, D_2,..., D_{J-1} are replaced by the variables $D^*_1 = D_1 - D_J$, $D^*_2 = D_2 - D_J$, $D^*_{J-1} = D_{J-1} - D_J$. The coefficient on the Jth category will not be explicitly reported when the regression is estimated, but can be calculated from the coefficients on the first J-1 transformed variables.

This technique constrains the sum of the coefficients for all values of the categorical variable to equal zero. However, it does not constrain the average effect of the categorical variable to equal zero in the population or sample of interest. If we wish to report coefficients that can be interpreted as deviations from the population or sample average effect, the original dummy variables D_j (j = 1,..J) must instead be replaced by:

$$D_j^{**} = D_j - (\sum_i D_j(i) / \sum_i D_J(i)) * D_J,$$

where the values of D_j and D_J are summed over all i observations in the population or sample.

Finally, we have generalized this technique beyond simple categorical variables. Where a group of continuous variables X_j (j = 1,..,J) is related in such a fashion that the sum of their values is fixed (for example, the percentage of total employment in each industry j must sum to 100 across all

13. Five variables, including two industrial mix control variables, are statistically significant for one specification, but are only significant at the 10-percent level for the other specification. Four additional variables, including three industrial mix control variables, are statistically significant at the 10-percent level for one specification but are not significant for the other specification. Even for these variables, the differences in assessed significance do not reflect dramatic differences between the two specifications in parameter estimates. The difference is less than 10 percent for four of these variables, between 10 and 20 percent for another four, and just over 20 percent of the smaller value in the ninth case.

industries 1,..J), then ordinarily one of the variables would be omitted from the regression specification, implicitly setting a zero constraint on the coefficient for that variable. As an alternative, the variables X_j can be replaced by:

$$X_j^{**} = X_j - (\overline{X_j}/\overline{X_J}) * X_J,$$

where the values of X_j and X_J are averaged over all i observations in the population or sample. Including the transformed variables in place of the original variables in the regression specification constrains the mean combined effect of all variables in the group, including the omitted variable, over all observations to equal zero. As before, the implied coefficient of the omitted variable can be computed from the reported results.[14]

In this report, the transformations described above are applied to the regional dummy variables (included in some specifications), to the percentage of employment in each industry, and to the percentage of employment in each occupation. Thus, reported regression coefficients for variables in each of these groups represent deviations from an average effect of all variables in the group. The implied coefficients for the Jth category in each group were computed and are also reported; standard errors for these implied coefficients were not, however, computed. Those variables not explicitly included in the regression models are the South Atlantic dummy variable, the percentage of employment in retail trade, and the percentage of the employed in sales and clerical occupations.[15]

---

14. Where a weighted sum of all effects is constrained to equal zero, rather than a simple sum of coefficients, it may be easier to compute the coefficient of the omitted category by running the regression with both original and transformed data and comparing coefficients of the included variables. This is the technique used in the current report.

15. The standard error for the implied coefficient can be computed fairly readily in the case where the simple sum of all dummy variable coefficients is constrained to equal zero. However, we have not worked out the appropriate formulae to use for the more complicated constraints imposed here, and hence we cannot report standard errors or t-statistics for these effects.

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods

Dependent variable for all methods shown is 1979-89 real county earnings growth.

	Ordinary least-squares (OLS) estimates			Estimates reflecting corrections for heteroskedasticity and non-normality (no commuting zone effects)		
Variable	Regression coefficient	Standardized coefficient	T-statistic	Regression coefficient	Standardized coefficient	T-statistic
Intercept	162.6	NA	3.84 ***	209.0	NA	5.63 ***
Demographic characteristics:						
Total urban population (1,000)	0.0318	0.0148	0.56	-0.0096	-0.0045	-0.27
Percent African-American	-0.0981	-0.0576	-1.86 +	-0.0836	-0.0491	-2.08 *
Percent Hispanic	-0.1547	-0.0636	-2.63 **	-0.0928	-0.0382	-1.82 +
Retirement county▒	9.55	0.1491	7.35 ***	8.481	0.1324	8.65 ***
Percent of population aged 25-64	0.477	0.0529	2.12 *	0.344	0.0382	1.92 +
Labor market characteristics:						
Mean annual earnings (log)	-20.47	-0.1526	-4.76 ***	-24.51	-0.1827	-6.37 ***
Right-to-work law▒	1.67	0.0321	1.23	0.24	0.0047	0.24
Labor force participation rate	0.034	0.0086	0.30	-0.0010	-0.0003	-0.01
Education levels and activity:						
Percent high school graduates	-0.069	-0.0326	-0.65	-0.0877	-0.0416	-1.04
Percent college graduates	-0.359	-0.0601	-1.48	-0.103	-0.0172	-0.52
Percent dropouts (aged 16-19)	0.1298	0.0384	1.51	0.1269	0.0375	1.79 +
Local college enrollment▒	3.28	0.0415	1.89 +	3.93	0.0497	3.31 ***
Local taxes and expenditures:						
Local tax level	0.116	0.0140	0.59	-0.129	-0.0156	-0.71
Education spending per pupil (K-12)	4.60	0.0781	3.00 **	4.30	0.0730	3.31 ***
Transportation access:						
Highway interchanges	-0.005	-0.0005	-0.03	0.086	0.0100	0.75
Highway intersections	3.17	0.0172	1.00	3.07	0.0166	1.59
Airport in county▒	-0.30	-0.0030	-0.15	-0.62	-0.0062	-0.50
Airport within 50 miles▒	3.06	0.0380	2.17 *	4.18	0.0519	3.81 ***
Business and banking structure:						
Small businesses as percent of all businesses in--						
Goods-producing industries	-0.2567	-0.1048	-4.38 ***	-0.1575	-0.0643	-3.19 **
Producer service industries	0.2911	0.0679	3.01 **	0.2677	0.0625	3.36 ***
Other service industries	0.0612	0.0392	2.12 *	0.0594	0.0381	2.48 *
Branch banking law▒	3.70	0.0494	2.28 *	2.11	0.0281	1.78 +

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

	Ordinary least-squares (OLS) estimates			Estimates reflecting corrections for heteroskedasticity and non-normality (no commuting zone effects)		
Variable	Regression coefficient	Standardized coefficient	T-statistic	Regression coefficient	Standardized coefficient	T-statistic
Amenities:						
Climate quality index	1.338	0.0978	2.84 **	1.745	0.1276	4.71 ***
Topography (mountainousness) index	-0.555	-0.0254	-1.00	0.084	0.0039	0.21
Water coverage index	0.523	0.0200	0.92	0.883	0.0338	2.07 *
Relationship to metro areas:						
Population of metro areas within 50 miles (in millions)	0.215	0.0071	0.37	-0.392	-0.0130	-0.97
Economic base:						
Transfer payments ($1,000 per capita)	-13.42	-0.1015	-4.16 ***	-17.39	-0.1315	-6.70 ***
Industrial diversification index	0.046068	0.0298	0.98	0.0829	0.0535	2.34 *
Percentage of county employment in occupation:						
Professional	0.043	0.0046	0.17	-0.177	-0.0191	-0.77
Managerial	0.679	0.0542	2.19 *	0.540	0.0431	2.08 *
Technical	2.556	0.0895	4.55 ***	2.083	0.0730	4.58 ***
Precision/crafts	-0.342	-0.0469	-2.01 *	-0.412	-0.0564	-2.85 **
Operators and laborers	-0.050	-0.0156	-0.50	-0.072	-0.0222	-0.88
Services	-0.827	-0.1098	-4.66 ***	-0.625	-0.0829	-4.06 ***
Sales and clerical	0.341	0.0466	z	0.481	0.0656	z
Percentage of county employment in industry:						
Agriculture, forestry, and fisheries--						
Farming	-0.027	-0.0121	-0.39	-0.084	-0.0377	-1.39
Agricultural services	-0.705	-0.0411	-2.05 *	-0.142	-0.0083	-0.43
Forestry	-2.819	-0.0376	-2.10 *	-1.853	-0.0247	-1.69 +
Fisheries	-0.624	-0.0213	-1.21	-0.669	-0.0228	-1.28
Mining--						
Metal mining	-0.988	-0.1202	-6.07 ***	-0.588	-0.0716	-2.10 *
Oil & gas extraction	-1.323	-0.1854	-7.84 ***	-1.204	-0.1687	-8.57 ***
Non-metal mining	-0.198	-0.0127	-0.74	-0.287	-0.0184	-1.12
Coal mining	-0.681	-0.1109	-4.41 ***	-0.524	-0.0854	-3.64 ***

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

Variable	Ordinary least-squares (OLS) estimates: Regression coefficient	Standardized coefficient	T-statistic	Estimates reflecting corrections for heteroskedasticity and non-normality (no commuting zone effects): Regression coefficient	Standardized coefficient	T-statistic
Percentage of county employment in industry:						
Construction--						
Construction-general	-1.061	-0.0839	-4.80 ***	-0.304	-0.0240	-1.09
Heavy construction	-0.391	-0.0463	-2.59 **	-0.478	-0.0566	-2.27 *
Construction-special	0.233	0.0126	0.70	0.349	0.0189	1.13
Manufacturing--						
Food manufacturing	0.437	0.0507	2.87 **	0.418	0.0484	3.57 ***
Tobacco manufacturing	0.093	0.0007	0.04	1.017	0.0071	0.71
Textiles	-0.156	-0.0271	-1.29	-0.047	-0.0082	-0.52
Apparel	-0.016	-0.0025	-0.11	0.141	0.0220	1.22
Lumber & wood products	-0.403	-0.0661	-3.15 **	-0.175	-0.0287	-1.79 +
Furniture & fixtures	-0.067	-0.0055	-0.32	-0.050	-0.0041	-0.37
Paper & allied products	0.369	0.0333	1.86 +	0.359	0.0324	2.61 **
Printing & publishing	0.741	0.0255	1.48	0.677	0.0233	2.46 *
Chemicals	-0.144	-0.0149	-0.85	-0.037	-0.0038	-0.27
Petroleum & coal products	-0.804	-0.0198	-1.17	-0.327	-0.0081	-0.54
Rubber & plastic products	0.081	0.0056	0.33	0.151	0.0104	0.85
Leather	-0.217	-0.0138	-0.78	0.037	0.0023	0.18
Stone, clay, & glass	-0.481	-0.0289	-1.69 +	-0.448	-0.0270	-2.55 *
Primary metal industries	-0.634	-0.0634	-3.57 ***	-0.443	-0.0443	-2.99 **
Fabricated metal products	0.259	0.0211	1.20	0.307	0.0250	1.79 +
Machinery (nonelectrical)	-0.365	-0.0369	-2.05 *	-0.385	-0.0389	-3.28 **
Electrical machinery	-0.307	-0.0294	-1.65 +	-0.243	-0.0232	-2.14 *
Motor vehicles	0.109	0.0058	0.35	-0.023	-0.0013	-0.10
Other transport equipment	0.479	0.0287	1.74 +	0.332	0.0199	1.26
Instruments	0.381	0.0152	0.92	0.460	0.0183	1.81 +
Miscellaneous manufacturing	0.478	0.0199	1.21	0.551	0.0229	1.75 +
Transportation and utilities--						
Railroads	-1.343	-0.0672	-3.88 ***	-1.171	-0.0586	-4.13 ***
Passenger transit	-0.822	-0.0066	-0.37	0.489	0.0040	0.29
Trucking & warehousing	-0.304	-0.0136	-0.77	-0.112	-0.0050	-0.32
Water transportation	-2.379	-0.0417	-2.32 *	-1.810	-0.0317	-2.40 *
Air transportation	3.056	0.0138	0.76	1.902	0.0086	0.51
Pipelines	0.233	0.0017	0.10	1.497	0.0107	0.90
Transport services	5.883	0.0351	2.08 *	4.095	0.0244	1.97 *
Communication	-0.087	-0.0022	-0.12	0.608	0.0154	1.05
Utilities	0.387	0.0210	1.18	0.398	0.0215	1.38

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

Variable	Ordinary least-squares (OLS) estimates: Regression coefficient	Standardized coefficient	T-statistic	Estimates reflecting corrections for heteroskedasticity and non-normality (no commuting zone effects): Regression coefficient	Standardized coefficient	T-statistic
Percentage of county employment in industry:						
Wholesale trade	-0.641	-0.0505	-2.29 *	-0.623	-0.0491	-2.72 **
Retail trade--						
Building materials	-0.068	-0.0016	-0.09	0.628	0.0146	1.09
General retail	0.570	0.0222	z	0.074	0.0029	z
Food stores	0.512	0.0207	1.13	0.549	0.0222	1.62
Auto dealers	0.670	0.0287	1.59	0.101	0.0043	0.26
Apparel stores	-2.671	-0.0447	-2.08 *	-2.252	-0.0377	-2.29 *
Furniture stores	0.150	0.0019	0.10	-0.710	-0.0090	-0.56
Restaurants, etc.	0.481	0.0439	1.93 +	0.283	0.0258	1.26
Miscellaneous retail	-0.144	-0.0044	-0.23	-0.011	-0.0004	-0.02
Finance, insurance, and real estate--						
Banking	-0.538	-0.0174	-0.81	-0.447	-0.0145	-0.83
Brokers	-9.167	-0.0181	-1.04	-6.142	-0.0121	-0.85
Insurance carriers	0.145	0.0034	0.20	0.365	0.0087	0.76
Insurance agents	0.095	0.0013	0.07	-0.634	-0.0088	-0.52
Real estate	2.411	0.1039	5.11 ***	2.930	0.1262	6.00 ***
Misc. combined finance	0.461	0.0019	0.11	-1.545	-0.0063	-0.49
Holding companies	3.400	0.0164	0.99	3.634	0.0176	1.65 +
Services--						
Hotels/lodging	0.863	0.0812	3.64 ***	0.719	0.0678	3.86 ***
Personal services	-0.920	-0.0142	-0.68	0.472	0.0073	0.48
Misc. business services	0.764	0.0532	3.08 **	0.673	0.0469	1.92 +
Auto repair	3.421	0.0421	2.39 *	3.108	0.0382	2.83 **
Miscellaneous repair	0.134	0.0016	0.09	1.333	0.0155	0.96
Motion pictures	-0.595	-0.0041	-0.22	-0.708	-0.0049	-0.34
Amusements, n.e.c.	0.420	0.0256	1.38	0.858	0.0523	2.45 *
Health services	-0.256	-0.0292	-1.37	-0.208	-0.0237	-1.37
Legal services	-1.090	-0.0087	-0.48	-3.657	-0.0293	-1.85 +
Educational services	0.902	0.0484	2.61 **	0.710	0.0381	2.90 **
Social services	-0.191	-0.0119	-0.69	-0.442	-0.0274	-1.94 +
Museums, zoos, etc.	-4.059	-0.0172	-1.04	-2.624	-0.0111	-1.51
Membership organizations	0.018	0.0009	0.05	-0.122	-0.0063	-0.44
Private household services	-0.102	-0.0081	-0.35	-0.436	-0.0346	-1.80 +
Miscellaneous services	0.579	0.0115	0.61	0.883	0.0175	1.15

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

Variable	Ordinary least-squares (OLS) estimates: Regression coefficient	Standardized coefficient	T-statistic	Estimates reflecting corrections for heteroskedasticity and non-normality (no commuting zone effects): Regression coefficient	Standardized coefficient	T-statistic
Percentage of county employment in industry:						
Government--						
Federal government (civilian)	-0.664	-0.0813	-4.42 ***	-0.245	-0.0300	-1.21
Federal government (military)	0.348	0.0489	2.69 **	0.264	0.0372	2.93 **
State & local government	0.389	0.0759	3.74 ***	0.254	0.0496	3.04 **
Region:						
New England	23.57	0.1166	6.00 ***	19.90	0.0985	7.65 ***
Middle Atlantic	6.69	0.0412	2.09 *	8.83	0.0544	4.08 ***
East North Central	0.82	0.0106	0.49	2.74	0.0356	2.22 *
West North Central	-8.05	-0.1323	-6.16 ***	-5.95	-0.0978	-6.34 ***
South Atlantic	8.94	0.1286	z	8.37	0.1205	z
East South Central	2.91	0.0372	1.73 +	3.48	0.0446	2.82 **
West South Central	-1.47	-0.0208	-0.99	-2.51	-0.0355	-2.18 *
Mountain	-2.69	-0.0322	-1.27	-6.32	-0.0756	-3.52 ***
Pacific	-2.69	-0.0192	-0.80	-7.30	-0.0522	-2.79 **
Rest-of-commuting-zone growth	NA	NA	NA	NA	NA	NA
Isolated county	NA	NA	NA	NA	NA	NA

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

Dependent variable for all methods shown is 1979-89 real county earnings growth.

	Estimates reflecting rest-of-commuting zone growth rate effect and corrections for heteroskedasticity and non-normality			Estimates reflecting corrections for heteroskedasticity, non-normality, and fixed commuting zone effects (preferred method)		
Variable	Regression coefficient	Standardized coefficient	T-statistic	Regression coefficient	Standardized coefficient	T-statistic
Intercept	192.4	NA	5.55 ***	NA	NA	NA
Demographic characteristics:						
Total urban population (1,000)	-0.0095	-0.0044	-0.29	-0.0263	-0.0122	-0.58
Percent African-American	-0.0921	-0.0541	-2.39 *	-0.1920	-0.1127	-3.01 **
Percent Hispanic	-0.1112	-0.0458	-2.54 *	-0.070	-0.0290	-0.66
Retirement county▒	6.792	0.1060	7.32 ***	4.50	0.0702	3.85 ***
Percent of population aged 25-64	0.308	0.0342	1.77 +	0.119	0.0132	0.52
Labor market characteristics:						
Mean annual earnings (log)	-21.53	-0.1605	-5.99 ***	-23.54	-0.1755	-5.25 ***
Right-to-work law▒	-0.628	-0.0121	-0.64	5.25	0.1010	2.26 *
Labor force participation rate	-0.0345	-0.0086	-0.39	-0.039	-0.0097	-0.33
Education levels and activity:						
Percent high school graduates	-0.0861	-0.0408	-1.09	0.331	0.1571	2.61 **
Percent college graduates	0.012	0.0020	0.06	-0.173	-0.0290	-0.71
Percent dropouts (aged 16-19)	0.077	0.0226	1.17	-0.0266	-0.0079	-0.33
Local college enrollment▒	2.40	0.0304	2.04 *	-0.43	-0.0054	-0.32
Local taxes and expenditures:						
Local tax level	-0.005	-0.0006	-0.03	-0.084	-0.0101	-0.40
Education spending per pupil (K-12)	3.72	0.0631	2.99 **	3.77	0.0640	2.39 *
Transportation access:						
Highway interchanges	0.189	0.0221	1.75 +	0.425	0.0496	3.23 **
Highway intersections	2.14	0.0116	1.03	0.27	0.0015	0.11
Airport in county▒	-0.12	-0.0012	-0.10	0.71	0.0070	0.48
Airport within 50 miles▒	2.77	0.0344	2.63 **	3.36	0.0417	2.26 *
Business and banking structure:						
Small businesses as percent of all businesses in--						
Goods-producing industries	-0.1327	-0.0542	-2.84 **	-0.1076	-0.0439	-1.94 +
Producer service industries	0.2088	0.0487	2.83 **	0.1298	0.0303	1.45
Other service industries	0.0254	0.0163	1.12	-0.0188	-0.0120	-0.73
Branch banking law▒	-0.30	-0.0041	-0.27	-2.62	-0.0350	-0.98

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

Variable	Estimates reflecting rest-of-commuting zone growth rate effect and corrections for heteroskedasticity and non-normality			Estimates reflecting corrections for heteroskedasticity, non-normality, and fixed commuting zone effects (preferred method)		
	Regression coefficient	Standardized coefficient	T-statistic	Regression coefficient	Standardized coefficient	T-statistic
Amenities:						
Climate quality index	1.144	0.0836	3.26 **	NA	NA	NA
Topography (mountainousness) index	0.087	0.0040	0.23	-1.321	-0.0605	-1.86 +
Water coverage index	0.589	0.0225	1.44	-0.600	-0.0229	-1.07
Relationship to metro areas:						
Population of metro areas within 50 miles (in millions)	-0.501	-0.0166	-1.40	-0.301	-0.0100	-0.49
Economic base:						
Transfer payments ($1,000 per capita)	-16.70	-0.1263	-6.94 ***	-15.67	-0.1186	-4.67 ***
Industrial diversification index	0.0848	0.0547	2.54 *	0.0629	0.0406	1.48
Percentage of county employment in occupation:						
Professional	-0.050	-0.0054	-0.23	-0.097	-0.0105	-0.37
Managerial	0.397	0.0317	1.59	0.192	0.0153	0.66
Technical	1.427	0.0500	3.30 ***	0.634	0.0222	1.20
Precision/crafts	-0.179	-0.0245	-1.30	-0.043	-0.0059	-0.26
Operators and laborers	-0.078	-0.0241	-1.01	-0.060	-0.0187	-0.62
Services	-0.503	-0.0668	-3.51 ***	-0.442	-0.0586	-2.56 *
Sales and clerical	0.294	0.0391	z	0.311	0.0425	z
Percentage of county employment in industry:						
Agriculture, forestry, and fisheries--						
Farming	-0.105	-0.0474	-1.90 +	-0.144	-0.0645	-2.13 *
Agricultural services	-0.128	-0.0075	-0.42	0.388	0.0226	0.93
Forestry	-1.907	-0.0254	-1.74 +	-2.320	-0.0309	-2.00 *
Fisheries	-0.538	-0.0184	-1.31	0.128	0.0044	0.22
Mining--						
Metal mining	-0.695	-0.0846	-2.54 *	-1.013	-0.1232	-3.63 ***
Oil & gas extraction	-1.129	-0.1582	-7.98 ***	-0.682	-0.0956	-3.96 ***
Non-metal mining	-0.322	-0.0207	-1.36	-0.482	-0.0309	-1.55
Coal mining	-0.450	-0.0733	-3.36 ***	-0.657	-0.1070	-4.77 ***

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

Variable	Estimates reflecting rest-of-commuting zone growth rate effect and corrections for heteroskedasticity and non-normality			Estimates reflecting corrections for heteroskedasticity, non-normality, and fixed commuting zone effects (preferred method)		
	Regression coefficient	Standardized coefficient	T-statistic	Regression coefficient	Standardized coefficient	T-statistic
Percentage of county employment in industry:						
Construction--						
Construction-general	-0.263	-0.0208	-0.92	-0.079	-0.0063	-0.32
Heavy construction	-0.611	-0.0723	-3.04 **	-0.624	-0.0738	-3.01 **
Construction-special	0.020	0.0011	0.07	0.365	0.0198	1.06
Manufacturing--						
Food manufacturing	0.274	0.0318	2.42 *	0.106	0.0123	0.87
Tobacco manufacturing	0.831	0.0058	0.50	-1.534	-0.0106	-0.82
Textiles	-0.079	-0.0136	-1.05	-0.162	-0.0281	-1.30
Apparel	0.041	0.0064	0.38	-0.024	-0.0037	-0.17
Lumber & wood products	-0.227	-0.0373	-2.40 *	-0.338	-0.0555	-2.49 *
Furniture & fixtures	-0.039	-0.0032	-0.28	-0.011	-0.0009	-0.08
Paper & allied products	0.261	0.0235	1.79 +	0.066	0.0059	0.39
Printing & publishing	0.646	0.0222	3.11 **	0.566	0.0194	2.05 *
Chemicals	-0.137	-0.0141	-0.98	-0.177	-0.0183	-1.21
Petroleum & coal products	-0.329	-0.0081	-0.61	-0.441	-0.0109	-0.68
Rubber & plastic products	0.209	0.0144	1.39	0.265	0.0182	1.40
Leather	0.008	0.0005	0.05	-0.041	-0.0026	-0.16
Stone, clay, & glass	-0.448	-0.0270	-2.89 **	-0.379	-0.0228	-1.78 +
Primary metal industries	-0.465	-0.0465	-2.64 **	-0.503	-0.0503	-2.68 **
Fabricated metal products	0.260	0.0213	1.58	0.142	0.0116	0.78
Machinery (nonelectrical)	-0.353	-0.0357	-2.80 **	-0.269	-0.0272	-1.83 +
Electrical machinery	-0.240	-0.0230	-2.05 *	-0.425	-0.0407	-3.14 **
Motor vehicles	-0.036	-0.0019	-0.16	0.017	0.0009	0.07
Other transport equipment	0.120	0.0072	0.40	0.111	0.0066	0.35
Instruments	0.513	0.0204	2.24 *	0.159	0.0063	0.47
Miscellaneous manufacturing	0.337	0.0140	1.26	0.203	0.0084	0.71
Transportation and utilities--						
Railroads	-1.230	-0.0616	-4.65 ***	-1.219	-0.0610	-3.75 ***
Passenger transit	-0.656	-0.0053	-0.41	0.694	0.0056	0.33
Trucking & warehousing	-0.238	-0.0107	-0.77	-0.709	-0.0318	-2.36 *
Water transportation	-1.948	-0.0341	-1.89 +	-1.225	-0.0215	-1.27
Air transportation	1.848	0.0084	0.58	-2.036	-0.0092	-0.54
Pipelines	1.455	0.0104	0.95	-0.572	-0.0041	-0.28
Transport services	3.667	0.0219	2.28 *	4.024	0.0240	2.00 *
Communication	0.356	0.0090	0.62	-0.729	-0.0185	-1.22
Utilities	0.286	0.0155	1.05	0.157	0.0085	0.54

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

Variable	Estimates reflecting rest-of-commuting zone growth rate effect and corrections heteroskedasticity and non-normality			Estimates reflecting corrections for heteroskedasticity, non-normality, and fixed commuting zone effects (preferred method)		
	Regression coefficient	Standardized coefficient	T-statistic	Regression coefficient	Standardized coefficient	T-statistic
Percentage of county employment in industry:						
Wholesale trade	-0.501	-0.0395	-2.37 *	-0.413	-0.0326	-1.73 +
Retail trade--						
Building materials	0.656	0.0152	1.22	0.661	0.0153	0.96
General retail	0.132	0.0051	z	0.247	0.0096	z
Food stores	0.548	0.0222	1.61	0.203	0.0082	0.54
Auto dealers	0.189	0.0081	0.49	0.295	0.0127	0.73
Apparel stores	-0.299	-0.0050	-0.35	-0.215	-0.0036	-0.21
Furniture stores	-0.288	-0.0036	-0.24	-0.651	-0.0082	-0.49
Restaurants, etc.	0.377	0.0343	1.71 +	0.297	0.0271	1.25
Miscellaneous retail	-0.146	-0.0045	-0.32	-0.511	-0.0157	-0.94
Finance, insurance, and real estate--						
Banking	-0.661	-0.0214	-1.29	-0.708	-0.0229	-1.20
Brokers	-1.487	-0.0029	-0.19	-1.360	-0.0027	-0.18
Insurance carriers	0.092	0.0022	0.21	0.005	0.0001	0.01
Insurance agents	-0.896	-0.0124	-0.78	-1.127	-0.0156	-1.10
Real estate	2.846	0.1226	5.70 ***	1.576	0.0679	2.79 **
Misc. combined finance	0.165	0.0007	0.06	-0.530	-0.0022	-0.17
Holding companies	2.844	0.0138	1.40	2.620	0.0127	1.82 +
Services--						
Hotels/lodging	0.597	0.0562	3.29 ***	0.543	0.0511	2.16 *
Personal services	-0.229	-0.0035	-0.25	0.125	0.0019	0.11
Misc. business services	0.668	0.0465	2.10 *	0.599	0.0417	4.91 ***
Auto repair	2.485	0.0305	2.39 *	-0.101	-0.0012	-0.08
Miscellaneous repair	2.085	0.0242	1.56	2.813	0.0326	1.96 *
Motion pictures	0.481	0.0033	0.23	-0.436	-0.0030	-0.17
Amusements, n.e.c.	0.404	0.0246	1.24	-0.270	-0.0164	-0.64
Health services	-0.062	-0.0071	-0.44	0.228	0.0260	1.45
Legal services	-3.068	-0.0245	-1.66 +	0.776	0.0062	0.35
Educational services	0.893	0.0480	3.82 ***	0.878	0.0471	4.32 ***
Social services	-0.510	-0.0316	-2.53 *	-0.433	-0.0268	-2.08 *
Museums, zoos, etc.	-0.655	-0.0028	-0.28	-1.099	-0.0047	-0.38
Membership organizations	0.040	0.0021	0.16	0.312	0.0160	0.89
Private household services	-0.202	-0.0161	-0.89	-0.066	-0.0052	-0.25
Miscellaneous services	0.756	0.0150	1.05	1.606	0.0318	2.01 *

See notes at end of table.

continued--

Appendix table C-1--Complete parameter estimates for preferred method and selected alternative methods --continued

Variable	Estimates reflecting rest-of-commuting zone growth rate effect and corrections for heteroskedasticity and non-normality			Estimates reflecting corrections for heteroskedasticity, non-normality, and fixed commuting zone effects (preferred method)		
	Regression coefficient	Standardized coefficient	T-statistic	Regression coefficient	Standardized coefficient	T-statistic
Percentage of county employment in industry:						
Government--						
Federal government (civilian)	-0.291	-0.0357	-1.65 +	-0.322	-0.0394	-1.89 +
Federal government (military)	0.259	0.0364	3.37 ***	0.177	0.0248	1.82 +
State & local government	0.254	0.0497	3.34 ***	0.363	0.0708	3.95 ***
Region:						
New England▒	13.34	0.0660	4.99 ***	NA	NA	NA
Middle Atlantic▒	5.91	0.0364	3.14 **	NA	NA	NA
East North Central▒	1.48	0.0193	1.31	NA	NA	NA
West North Central▒	-4.90	-0.0806	-5.47 ***	NA	NA	NA
South Atlantic▒	6.29	0.0906	z	NA	NA	NA
East South Central▒	2.50	0.0319	2.22 *	NA	NA	NA
West South Central▒	-1.39	-0.0197	-1.24	NA	NA	NA
Mountain▒	-3.99	-0.0478	-2.23 *	NA	NA	NA
Pacific▒	-3.28	-0.0234	-1.38	NA	NA	NA
Rest-of-commuting-zone growth	0.3044	0.2369	15.29 ***	NA	NA	NA
Isolated county▒	-1.37	-0.0098	-0.51	NA	NA	NA

z = This statistic cannot readily be computed when the modified Suits method is used to compute coefficients for variables related by additivity constraints.
n.e.c. = Not elsewhere classified.
+ = Statistically significant at the 10-percent level.
* = Statistically significant at the 5-percent level.
** = Statistically significant at the 1-percent level.
*** = Statistically significant at the 0.1- percent level.
NA = Not estimated or not applicable.
▒ = 0, 1 variable.

Source: U.S. Department of Agriculture, Economic Research Service.

Appendix D. Comparison of Preferred-Method Results with Backward Elimination Results

The model of earnings growth presented in this report includes many variables, particularly when those variables introduced to control for industry and employment mix are taken into account. Further, significant degrees of multicollinearity exist among some of the variables included in the model. For these reasons, some reviewers of an earlier version of the report expressed concern that our results might be unduly influenced by multicollinearity; the analysis in this appendix seeks to address that concern.

As noted in Appendix B (pp. B-8 to B-9), a backward elimination process has been applied to the basic model to drop variables that were not statistically significant. Variables left in the resulting model are those significant at least at the 10-percent level. In identifying variables to be included in this fixed-effects backward elimination model, the data are transformed prior to application of the backward elimination algorithm to reflect both the fixed effects and the weighting for heteroskedasticity and non-normality incorporated in the preferred method.

In general, the backward elimination algorithm cannot be relied upon to identify the right variables from a set of collinear independent variables. To see this, it is useful to note that, at the extreme, such an algorithm would be expected to pick out a small number of statistically significant variables from a large number of independent variables that had no true relationship (direct or indirect) with the dependent variable.

However, examination of the backward elimination results may still be useful as a check on our preferred-method results. If many variables that had not been significant in the original preferred-method estimates are identified as statistically significant after using the backward elimination process, this might suggest that multicollinearity had obscured some important relationships, even though we could not be confident that the algorithm had correctly picked out the independent variables involved in those relationships.

But in fact, a comparison of the backward elimination results with the preferred-method results (app. table D-1) reveals that the original preferred-method estimates and the backward elimination estimates identify nearly the same set of significant variables, and that in most cases the parameter estimates for the two models are quite similar.

1. Considering industrial and occupational mix variables together with those of substantive interest, we find that 30 variables are identified as significant in both sets of estimates. Four additional variables are identified as significant at the 10-percent level in both sets.

   For 26 of these 34 variables, the two parameter estimates are within 10 percent of each other; in six cases, the difference is between 10 and 20 percent (of the smaller value). The remaining two cases include the transfer-payment effect (the original preferred-method parameter estimate is 20.03 percent greater than the corresponding backward elimination estimate) and the lumber and wood products employment effect, where the preferred-method estimate is 27 percent greater.

2. In four cases (three industry share variables, and the elevation-topography index), an effect is significant in the backward elimination estimates, but is only significant at the 10-percent level in the original preferred-method estimates. The preferred-method and backward elimination coefficient estimates differ by less than 10 percent in three of these four cases, and by 11 percent in the other case.

Appendix table D-1--Parameter estimates when preferred method is supplemented with backward elimination of variables

Variable	Backward elimination results: Regression coefficient	Backward elimination results: T-statistic	Preferred-method results: Regression coefficient	Preferred-method results: T-statistic
Intercept	NA	NA	NA	NA
Demographic characteristics:				
Total urban population (1,000)	0	0	-0.0263	-0.58
Percent African-American	-0.2088	-3.79 ***	-0.1920	-3.01 **
Percent Hispanic	0	0	-0.070	-0.66
Retirement county	4.77	4.34 ***	4.50	3.85 ***
Percent of population aged 25-64	0	0	0.119	0.52
Labor market characteristics:				
Mean annual earnings (log)	-22.81	-6.28 ***	-23.54	-5.25 ***
Right-to-work law	5.63	2.53 *	5.25	2.26 *
Labor force participation rate	0	0	-0.039	-0.33
Education levels and activity:				
Percent high school graduates	0.2939	3.42 ***	0.331	2.61 **
Percent college graduates	0	0	-0.173	-0.71
Percent dropouts (aged 16-19)	0	0	-0.0266	-0.33
Local college enrollment	0	0	-0.43	-0.32
Local taxes and expenditures:				
Local tax level	0	0	-0.084	-0.40
Education spending per pupil (K-12)	3.29	2.42 *	3.77	2.39 *
Transportation access:				
Highway interchanges	0.436	3.68 ***	0.425	3.23 **
Highway intersections	0	0	0.27	0.11
Airport in county	0	0	0.71	0.48
Airport within 50 miles	2.95	2.06 *	3.36	2.26 *
Business and banking structure:				
Small businesses as percent of all businesses in--				
Goods-producing industries	-0.0911	-1.77 +	-0.1076	-1.94 +
Producer service industries	0.1441	1.77 +	0.1298	1.45
Other service industries	0	0	-0.0188	-0.73
Branch banking law	0	0	-2.62	-0.98

See notes at end of table.

continued--

Appendix table D-1--Parameter estimates when preferred method is supplemented with backward elimination of variables--continued

Variable	Backward elimination results: Regression coefficient	Backward elimination results: T-statistic	Preferred-method results: Regression coefficient	Preferred-method results: T-statistic
Amenities:				
Climate quality index	NA	NA	NA	NA
Topography (mountainousness) index	-1.365	-2.01 *	-1.321	-1.86 +
Water coverage index	0	0	-0.600	-1.07
Relationship to metro areas:				
Population of metro areas within 50 miles (in millions)	0	0	-0.301	-0.49
Economic base:				
Transfer payments ($1,000 per capita)	-13.06	-4.50 ***	-15.67	-4.67 ***
Industrial diversification index	0.0777	2.36 *	0.0629	1.48
Percentage of county employment in occupation:				
Professional	0	0	-0.097	-0.37
Managerial	0	0	0.192	0.66
Technical	0	0	0.634	1.20
Precision/crafts	0	0	-0.043	-0.26
Operators and laborers	0	0	-0.060	-0.62
Services	-0.418	-3.09 **	-0.442	-2.56 *
Sales and clerical	z	z	0.311	z
Percentage of county employment in industry:				
Agriculture, forestry, and fisheries--				
Farming	-0.087	-1.68 +	-0.144	-2.13 *
Agricultural services	0	0	0.388	0.93
Forestry	-2.278	-2.05 *	-2.320	-2.00 *
Fisheries	0	0	0.128	0.22
Mining--				
Metal mining	-1.002	-3.77 ***	-1.013	-3.63 ***
Oil & gas extraction	-0.656	-4.21 ***	-0.682	-3.96 ***
Non-metal mining	0	0	-0.482	-1.55
Coal mining	-0.644	-5.45 ***	-0.657	-4.77 ***

See notes at end of table.

continued--

Appendix table D-1--Parameter estimates when preferred method is supplemented with backward elimination of variables--continued

Variable	Backward elimination results		Preferred-method results	
	Regression coefficient	T-statistic	Regression coefficient	T-statistic
Percentage of county employment in industry:				
Construction--				
Construction-general	0	0	-0.079	-0.32
Heavy construction	-0.605	-3.11 **	-0.624	-3.01 **
Construction-special	0	0	0.365	1.06
Manufacturing--				
Food manufacturing	0	0	0.106	0.87
Tobacco manufacturing	0	0	-1.534	-0.82
Textiles	0	0	-0.162	-1.30
Apparel	0	0	-0.024	-0.17
Lumber & wood products	-0.266	-2.14 *	-0.338	-2.49 *
Furniture & fixtures	0	0	-0.011	-0.08
Paper & allied products	0	0	0.066	0.39
Printing & publishing	0.601	2.29 *	0.566	2.05 *
Chemicals	0	0	-0.177	-1.21
Petroleum & coal products	0	0	-0.441	-0.68
Rubber & plastic products	0	0	0.265	1.40
Leather	0	0	-0.041	-0.16
Stone, clay, & glass	-0.370	-1.87 +	-0.379	-1.78 +
Primary metal industries	-0.518	-2.92 **	-0.503	-2.68 **
Fabricated metal products	0	0	0.142	0.78
Machinery (nonelectrical)	-0.268	-1.96 *	-0.269	-1.83 +
Electrical machinery	-0.389	-3.09 **	-0.425	-3.14 **
Motor vehicles	0	0	0.017	0.07
Other transport equipment	0	0	0.111	0.35
Instruments	0	0	0.159	0.47
Miscellaneous manufacturing	0	0	0.203	0.71
Transportation and utilities--				
Railroads	-1.160	-3.79 ***	-1.219	-3.75 ***
Passenger transit	0	0	0.694	0.33
Trucking & warehousing	-0.592	-2.12 *	-0.709	-2.36 *
Water transportation	0	0	-1.225	-1.27
Air transportation	0	0	-2.036	-0.54
Pipelines	0	0	-0.572	-0.28
Transport services	3.730	1.96 *	4.024	2.00 *
Communication	0	0	-0.729	-1.22
Utilities	0	0	0.157	0.54

See notes at end of table. continued--

Appendix table D-1--Parameter estimates when preferred method is supplemented with backward elimination of variables--continued

Variable	Backward elimination results: Regression coefficient	Backward elimination results: T-statistic	Preferred-method results: Regression coefficient	Preferred-method results: T-statistic
Percentage of county employment in industry:				
Wholesale trade	-0.403	-1.90 +	-0.413	-1.73 +
Retail trade--				
Building materials	0	0	0.661	0.96
General retail	z	z	0.247	z
Food stores	0	0	0.203	0.54
Auto dealers	0	0	0.295	0.73
Apparel stores	0	0	-0.215	-0.21
Furniture stores	0	0	-0.651	-0.49
Restaurants, etc.	0	0	0.297	1.25
Miscellaneous retail	0	0	-0.511	-0.94
Finance, insurance, and real estate--				
Banking	0	0	-0.708	-1.20
Brokers	0	0	-1.360	-0.18
Insurance carriers	0	0	0.005	0.01
Insurance agents	0	0	-1.127	-1.10
Real estate	1.567	3.09 **	1.576	2.79 **
Misc. combined finance	0	0	-0.530	-0.17
Holding companies	2.835	2.13 *	2.620	1.82 +
Services--				
Hotels/lodging	0.523	2.44 *	0.543	2.16 *
Personal services	0	0	0.125	0.11
Misc. business services	0.612	5.75 ***	0.599	4.91 ***
Auto repair	0	0	-0.101	-0.08
Miscellaneous repair	2.801	2.04 *	2.813	1.96 *
Motion pictures	0	0	-0.436	-0.17
Amusements, n.e.c.	0	0	-0.270	-0.64
Health services	0	0	0.228	1.45
Legal services	0	0	0.776	0.35
Educational services	0.816	4.88 ***	0.878	4.32 ***
Social services	-0.475	-2.45 *	-0.433	-2.08 *
Museums, zoos, etc.	0	0	-1.099	-0.38
Membership organizations	0	0	0.312	0.89
Private household services	0	0	-0.066	-0.25
Miscellaneous services	1.838	2.48 *	1.606	2.01 *

See notes at end of table.

continued--

Appendix table D-1--Parameter estimates when preferred method is supplemented with backward elimination of variables--continued

Variable	Backward elimination results		Preferred-method results	
	Regression coefficient	T-statistic	Regression coefficient	T-statistic
Percentage of county employment in industry:				
Government--				
Federal government (civilian)	-0.287	-1.85 +	-0.322	-1.89 +
Federal government (military)	0.196	2.31 *	0.177	1.82 +
State & local government	0.368	5.82 ***	0.363	3.95 ***

z = This statistic cannot readily be computed when the modified Suits method is used to compute coefficients for variables related by additivity constraints.
n.e.c. = Not elsewhere classified.
+ = Statistically significant at the 10-percent level.
* = Statistically significant at the 5-percent level.
** = Statistically significant at the 1-percent level.
*** = Statistically significant at the 0.1- percent level.
NA = Not estimated or not applicable.
▒ = 0, 1 variable.

Source: U.S. Department of Agriculture, Economic Research Service.

3. In one additional case, the farming industry employment share effect is only significant at the 10-percent level in the backward elimination estimates, while it is significant at the 5-percent level, and 65 percent greater in magnitude in the original estimates.

4. If no sharp distinction is made between those coefficient estimates that are significant at the 5-percent level and those that are significant only at the 10-percent level, then there are only two variables for which the preferred-method estimates and the backward elimination estimates yield notably different conclusions:

   a) industrial diversification, which is not statistically significant in the original model, has a statistically significant positive effect on earnings growth in the backward elimination model, and

   b) the small-firm share of producer service establishments, not statistically significant in the original model, has a positive effect on earnings growth that is significant at the 10-percent level in the backward elimination model.

Overall, these results suggest that inclusion of a relatively large set of variables in the model has interfered little with the identification or measurement of important relationships.

Appendix table E-1--Preferred-method parameter estimates, with 95-percent confidence intervals

Variable	Regression coefficient	95% confidence interval		Standardized values		
		Lower bound	Upper bound	Regression coefficient	Confidence interval Lower bound	Confidence interval Upper bound
Intercept	NA	NA	NA	NA	NA	NA
Demographic characteristics:						
Total urban population (1,000)	-0.026	-0.115	0.063	-0.0122	0.0290	-0.0534
Percent African-American	-0.192	-0.317	-0.067	-0.1127	-0.0393	-0.1861
Percent Hispanic	-0.070	-0.279	0.139	-0.0290	0.0570	-0.1150
Retirement county	4.50	2.21	6.79	0.0702	0.1060	0.0345
Percent of population aged 25-64	0.119	-0.329	0.566	0.0132	0.0628	-0.0365
Labor market characteristics:						
Mean annual earnings (log)	-23.5	-32.3	-14.8	-0.1755	-0.1100	-0.2410
Right-to-work law	5.25	0.70	9.80	0.1010	0.1886	0.0134
Labor force participation rate	-0.039	-0.273	0.194	-0.0097	0.0481	-0.0676
Education levels and activity:						
Percent high school graduates	0.331	0.083	0.580	0.1571	0.2751	0.0391
Percent college graduates	-0.173	-0.652	0.305	-0.0290	0.0511	-0.1091
Percent dropouts (aged 16-19)	-0.027	-0.184	0.131	-0.0079	0.0388	-0.0545
Local college enrollment	-0.43	-3.06	2.20	-0.0054	0.0279	-0.0388
Local taxes and expenditures:						
Local tax level	-0.084	-0.494	0.326	-0.0101	0.0395	-0.0597
Education spending per pupil (K-12)	3.77	0.68	6.87	0.0640	0.1165	0.0115
Transportation access:						
Highway interchanges	0.425	0.167	0.682	0.0496	0.0797	0.0195
Highway intersections	0.27	-4.55	5.09	0.0015	0.0276	-0.0247
Airport in county	0.71	-2.18	3.59	0.0070	0.0357	-0.0216
Airport within 50 miles	3.36	0.45	6.27	0.0417	0.0778	0.0055
Business and banking structure:						
Small businesses as percent of all businesses in--						
Goods-producing industries	-0.108	-0.216	0.001	-0.0439	0.0005	-0.0883
Producer service industries	0.130	-0.046	0.305	0.0303	0.0712	-0.0106
Other service industries	-0.019	-0.069	0.032	-0.0120	0.0203	-0.0444
Branch banking law	-2.62	-7.87	2.62	-0.0350	0.0350	-0.1049

See notes at end of table.

continued--

Appendix table E-1--Preferred-method parameter estimates, with 95-percent confidence intervals --continued

Variable	Regression coefficient	95% confidence interval		Standardized values		
		Lower bound	Upper bound	Regression coefficient	Confidence interval Lower bound	Confidence interval Upper bound
Amenities:						
Climate quality index	NA	NA	NA	NA	NA	NA
Topography (mountainousness) index	-1.32	-2.71	0.07	-0.0605	0.0033	-0.1244
Water coverage index	-0.60	-1.70	0.50	-0.0229	0.0191	-0.0649
Relationship to metro areas:						
Population of metro areas within 50 miles (in millions)	-0.30	-1.51	0.90	-0.0100	0.0300	-0.0500
Economic base:						
Transfer payments ($1,000 per capita)	-15.7	-22.3	-9.1	-0.1186	-0.0688	-0.1683
Industrial diversification index	0.063	-0.020	0.146	0.0406	0.0943	-0.0132
Percentage of county employment in occupation:						
Professional	-0.10	-0.61	0.42	-0.0105	0.0450	-0.0660
Managerial	0.19	-0.38	0.76	0.0153	0.0608	-0.0302
Technical	0.63	-0.40	1.67	0.0222	0.0585	-0.0141
Precision/crafts	-0.04	-0.37	0.28	-0.0059	0.0386	-0.0504
Operators and laborers	-0.06	-0.25	0.13	-0.0187	0.0403	-0.0776
Services	-0.44	-0.78	-0.10	-0.0586	-0.0137	-0.1035
Sales and clerical	0.31	z	z	0.0425	z	z
Percentage of county employment in industry:						
Agriculture, forestry, and fisheries--						
Farming	-0.14	-0.28	-0.01	-0.0645	-0.0052	-0.1239
Agricultural services	0.39	-0.43	1.21	0.0226	0.0703	-0.0250
Forestry	-2.32	-4.59	-0.05	-0.0309	-0.0006	-0.0612
Fisheries	0.13	-1.01	1.27	0.0044	0.0432	-0.0345
Mining--						
Metal mining	-1.01	-1.56	-0.47	-0.1232	-0.0567	-0.1898
Oil & gas extraction	-0.68	-1.02	-0.34	-0.0956	-0.0483	-0.1429
Non-metal mining	-0.48	-1.09	0.13	-0.0309	0.0082	-0.0701
Coal mining	-0.66	-0.93	-0.39	-0.1070	-0.0630	-0.1510

See notes at end of table.

continued--

Appendix table E-1--Preferred-method parameter estimates, with 95-percent confidence intervals --continued

Variable	95% confidence interval		Standardized values		
				Confidence interval	
	Regression coefficient	Lower bound	Upper bound	Regression coefficient	Lower bound	Upper bound
Construction--						
Construction-general	-0.08	-0.57	0.41	-0.0063	0.0321	-0.0447
Heavy construction	-0.62	-1.03	-0.22	-0.0738	-0.0257	-0.1218
Construction-special	0.36	-0.31	1.04	0.0198	0.0564	-0.0168
Manufacturing--						
Food manufacturing	0.11	-0.13	0.34	0.0123	0.0400	-0.0154
Tobacco manufacturing	-1.53	-5.20	2.13	-0.0106	0.0148	-0.0361
Textiles	-0.16	-0.41	0.08	-0.0281	0.0143	-0.0704
Apparel	-0.02	-0.30	0.25	-0.0037	0.0393	-0.0467
Lumber & wood products	-0.34	-0.60	-0.07	-0.0555	-0.0118	-0.0991
Furniture & fixtures	-0.01	-0.29	0.27	-0.0009	0.0221	-0.0240
Paper & allied products	0.07	-0.26	0.40	0.0059	0.0357	-0.0238
Printing & publishing	0.57	0.02	1.11	0.0194	0.0380	0.0009
Chemicals	-0.18	-0.46	0.11	-0.0183	0.0113	-0.0478
Petroleum & coal products	-0.44	-1.71	0.83	-0.0109	0.0205	-0.0422
Rubber & plastic products	0.26	-0.11	0.64	0.0182	0.0437	-0.0073
Leather	-0.04	-0.54	0.46	-0.0026	0.0293	-0.0345
Stone, clay, & glass	-0.38	-0.80	0.04	-0.0228	0.0023	-0.0479
Primary metal industries	-0.50	-0.87	-0.14	-0.0503	-0.0135	-0.0870
Fabricated metal products	0.14	-0.21	0.50	0.0116	0.0407	-0.0175
Machinery (nonelectrical)	-0.27	-0.56	0.02	-0.0272	0.0019	-0.0564
Electrical machinery	-0.43	-0.69	-0.16	-0.0407	-0.0153	-0.0661
Motor vehicles	0.02	-0.45	0.48	0.0009	0.0255	-0.0238
Other transport equipment	0.11	-0.51	0.73	0.0066	0.0437	-0.0305
Instruments	0.16	-0.50	0.82	0.0063	0.0326	-0.0200
Miscellaneous manufacturing	0.20	-0.36	0.76	0.0084	0.0317	-0.0149
Transportation and utilities--						
Railroads	-1.22	-1.86	-0.58	-0.0610	-0.0291	-0.0929
Passenger transit	0.69	-3.43	4.82	0.0056	0.0389	-0.0277
Trucking & warehousing	-0.71	-1.30	-0.12	-0.0318	-0.0054	-0.0583
Water transportation	-1.23	-3.12	0.67	-0.0215	0.0117	-0.0546
Air transportation	-2.04	-9.43	5.35	-0.0092	0.0242	-0.0426
Pipelines	-0.57	-4.58	3.43	-0.0041	0.0246	-0.0328
Transport services	4.02	0.08	7.97	0.0240	0.0475	0.0005
Communication	-0.73	-1.90	0.44	-0.0185	0.0112	-0.0481
Utilities	0.16	-0.41	0.73	0.0085	0.0393	-0.0223

See notes at end of table.

continued--

Appendix table E-1--Preferred-method parameter estimates, with 95-percent confidence intervals --continued

Variable	Regression coefficient	95% confidence interval		Standardized values		
		Lower bound	Upper bound	Regression coefficient	Confidence interval Lower bound	Confidence interval Upper bound
Wholesale trade	-0.41	-0.88	0.05	-0.0326	0.0043	-0.0695
Retail trade--						
Building materials	0.66	-0.69	2.01	0.0153	0.0466	-0.0160
General retail	0.25	z	z	0.0096		
Food stores	0.20	-0.53	0.94	0.0082	0.0380	-0.0216
Auto dealers	0.30	-0.50	1.09	0.0127	0.0467	-0.0214
Apparel stores	-0.22	-2.22	1.79	-0.0036	0.0300	-0.0372
Furniture stores	-0.65	-3.26	1.95	-0.0082	0.0247	-0.0411
Restaurants, etc.	0.30	-0.17	0.76	0.0271	0.0695	-0.0154
Miscellaneous retail	-0.51	-1.58	0.55	-0.0157	0.0171	-0.0485
Finance, insurance, and real estate--						
Banking	-0.71	-1.87	0.45	-0.0229	0.0145	-0.0604
Brokers	-1.36	-16.17	13.45	-0.0027	0.0266	-0.0320
Insurance carriers	0.01	-0.99	1.00	0.0001	0.0234	-0.0232
Insurance agents	-1.13	-3.14	0.88	-0.0156	0.0122	-0.0435
Real estate	1.58	0.47	2.68	0.0679	0.1156	0.0202
Misc. combined finance	-0.53	-6.64	5.58	-0.0022	0.0228	-0.0271
Holding companies	2.62	-0.20	5.44	0.0127	0.0263	-0.0010
Services--						
Hotels/lodging	0.54	0.05	1.04	0.0511	0.0975	0.0047
Personal services	0.13	-2.11	2.36	0.0019	0.0357	-0.0319
Misc. business services	0.60	0.36	0.84	0.0417	0.0583	0.0250
Auto repair	-0.10	-2.58	2.38	-0.0012	0.0293	-0.0317
Miscellaneous repair	2.81	0.00	5.63	0.0326	0.0653	0.0000
Motion pictures	-0.44	-5.46	4.59	-0.0030	0.0315	-0.0375
Amusements, n.e.c.	-0.27	-1.09	0.56	-0.0164	0.0339	-0.0668
Health services	0.23	-0.08	0.54	0.0260	0.0611	-0.0091
Legal services	0.78	-3.57	5.12	0.0062	0.0410	-0.0286
Educational services	0.88	0.48	1.28	0.0471	0.0685	0.0257
Social services	-0.43	-0.84	-0.03	-0.0268	-0.0015	-0.0521
Museums, zoos, etc.	-1.10	-6.77	4.57	-0.0047	0.0194	-0.0287
Membership organizations	0.31	-0.37	1.00	0.0160	0.0513	-0.0193
Private household services	-0.07	-0.58	0.45	-0.0052	0.0358	-0.0463
Miscellaneous services	1.61	0.04	3.17	0.0318	0.0629	0.0008

See notes at end of table.

continued--

Appendix table E-1--Preferred-method parameter estimates, with 95-percent confidence intervals --continued

Variable	Regression coefficient	95% confidence interval		Standardized values		
		Lower bound	Upper bound	Regression coefficient	Confidence interval	
					Lower bound	Upper bound
Government--						
Federal government (civilian)	-0.32	-0.66	0.01	-0.0394	0.0015	-0.0802
Federal government (military)	0.18	-0.01	0.37	0.0248	0.0515	-0.0019
State & local government	0.36	0.18	0.54	0.0708	0.1059	0.0357

▒ = 0,1 variable.
NA = Not estimated or not applicable.
n.e.c. = Not elsewhere classified.
z = This statistic cannot readily be computed when the modified Suits method is used to compute coefficients for variables related by additivity constraints.

Source: U.S. Department of Agriculture, Economic Research Service.

Appendix table F-1--Variables included in the growth model: Employment by industry and occupation

Variable name	Mean	Standard deviation
Percentage of county employment in occupation:		
Professional	10.73	2.79
Managerial	8.30	2.06
Technical	2.20	0.91
Precision/crafts	16.14	3.54
Operators and laborers	25.07	8.03
Services	15.18	3.43
Sales and clerical	22.39	3.53
Percentage of county employment in industry:		
Agriculture, forestry, and fisheries--		
Farming	17.09	11.62
Agricultural services	1.14	1.51
Forestry	0.076	0.344
Fisheries	0.087	0.882
Mining--		
Metal mining	0.39	3.14
Oil & gas extraction	1.18	3.62
Non-metal mining	0.40	1.66
Coal mining	0.81	4.21
Construction--		
Construction-general	1.79	2.04
Heavy construction	1.65	3.06
Construction-special	2.07	1.40
Manufacturing--		
Food manufacturing	1.61	3.00
Tobacco manufacturing	0.016	0.179
Textiles	1.15	4.47
Apparel	2.20	4.03
Lumber & wood products	2.22	4.24
Furniture & fixtures	0.51	2.12
Paper & allied products	0.57	2.33
Printing & publishing	0.55	0.89
Chemicals	0.55	2.66
Petroleum & coal products	0.100	0.637
Rubber & plastic products	0.56	1.78

See notes at end of table. continued--

Appendix table F-1--Variables included in the growth model: Employment by industry and occupation--continued

Variable name	Mean	Standard deviation
Percentage of county employment in industry (cont.):		
Manufacturing (cont.)--		
Leather	0.39	1.65
Stone, clay, & glass	0.64	1.56
Primary metal industries	0.61	2.58
Fabricated metal products	0.83	2.11
Machinery (nonelectrical)	1.23	2.61
Electrical machinery	0.93	2.48
Instruments	0.19	1.03
Miscellaneous manufacturing	0.27	1.08
Motor vehicles	0.35	1.38
Other transport equipment	0.27	1.55
Transportation and utilities--		
Railroads	0.670	1.294
Passenger transit	0.114	0.209
Trucking & warehousing	1.146	1.161
Water transportation	0.083	0.453
Air transportation	0.038	0.117
Pipelines	0.036	0.185
Transport services	0.038	0.154
Communication	0.782	0.654
Utilities	1.063	1.398
Wholesale trade	3.415	2.040
Retail trade--		
Building materials	0.995	0.600
General retail	1.285	1.008
Food stores	2.734	1.045
Auto dealers	2.544	1.110
Apparel stores	0.597	0.433
Furniture stores	0.464	0.326
Restaurants, etc.	4.110	2.356
Miscellaneous retail	1.550	0.794
Finance, insurance, and real estate--		
Banking	2.385	0.837
Brokers	0.015	0.051
Insurance carriers	0.238	0.612
Insurance agents	0.419	0.358
Real estate	0.632	1.114
Misc. combined finance	0.042	0.106
Holding companies	0.030	0.125

See notes at end of table. continued--

Appendix table F-1--Variables included in the growth model: Employment by industry and occupation--continued

Variable name	Mean	Standard deviation
Percentage of county employment in industry:		
Services--		
Hotels/lodging	1.34	2.43
Personal services	0.67	0.40
Misc. business services	0.58	1.80
Auto repair	0.38	0.32
Miscellaneous repair	0.21	0.30
Motion pictures	0.12	0.18
Amusements, n.e.c.	0.58	1.58
Health services	4.41	2.94
Legal services	0.32	0.21
Educational services	0.57	1.39
Social services	1.04	1.60
Museums, zoos, etc.	0.011	0.109
Membership organizations	1.78	1.33
Private household services	2.97	2.05
Miscellaneous services	0.37	0.51
Government--		
Federal government (civilian)	1.82	3.17
Federal government (military)	1.53	3.63
State & local government	13.46	5.05

n.e.c. = Not elsewhere classified.
Values are for 2,346 counties covered in regression analyses.

Source: U.S. Department of Agriculture, Economic Research Service.

Appendix table G-1--Tolerances and Variance Inflation Factors for Independent Variables Included in Model

Variable	Tolerances		Variance inflation factors	
	OLS	Robust weights	OLS	Robust weights
Intercept	NA	NA	NA	NA
Demographic characteristics:				
Total urban population (1,000)	0.364	0.326	2.75	3.07
Percent African-American	0.267	0.216	3.74	4.64
Percent Hispanic	0.437	0.497	2.29	2.01
Retirement county	0.622	0.591	1.61	1.69
Percent of population aged 25-64	0.412	0.368	2.43	2.72
Labor market characteristics:				
Mean annual earnings (log)	0.249	0.213	4.01	4.70
Right-to-work law	0.375	0.345	2.67	2.90
Labor force participation rate	0.311	0.271	3.21	3.69
Education levels and activity:				
Percent high school graduates	0.100	0.089	9.95	11.22
Percent college graduates	0.155	0.117	6.47	8.56
Percent dropouts (aged 16-19)	0.398	0.367	2.51	2.72
Local college enrollment	0.529	0.451	1.89	2.22
Local taxes and expenditures:				
Local tax level	0.450	0.434	2.22	2.30
Education spending per pupil (K-12)	0.377	0.342	2.65	2.93
Transportation access:				
Highway interchanges	0.659	0.636	1.52	1.57
Highway intersections	0.868	0.799	1.15	1.25
Airport in county	0.648	0.581	1.54	1.72
Airport within 50 miles	0.831	0.823	1.20	1.21
Business and banking structure:				
Small businesses as percent of all businesses in--				
Goods-producing industries	0.447	0.421	2.24	2.38
Producer service industries	0.502	0.423	1.99	2.37
Other service industries	0.752	0.708	1.33	1.41
Branch banking law	0.545	0.384	1.84	2.61

See notes at end of table.

continued--

Appendix table G-1--Tolerances and Variance Inflation Factors for Independent Variables Included in Model--continued

Variable	Tolerances		Variance inflation factors	
	OLS	Robust weights	OLS	Robust weights
Amenities:				
Climate quality index	0.216	0.220	4.62	4.55
Topography (mountainousness) index	0.394	0.452	2.54	2.21
Water coverage index	0.542	0.563	1.85	1.78
Relationship to metro areas:				
Population of metro areas within 50 miles (in millions)	0.684	0.651	1.46	1.54
Economic base:				
Transfer payments ($1,000 per capita)	0.430	0.406	2.33	2.46
Industrial diversification index	0.279	0.360	3.59	2.77
Percentage of county employment in occupation:				
Professional	0.188	0.143	5.31	6.99
Managerial	0.261	0.242	3.84	4.14
Technical	0.573	0.521	1.74	1.92
Precision/crafts	0.203	0.216	4.93	4.62
Operators and laborers	0.0625	0.0496	16.00	20.15
Services	0.193	0.163	5.17	6.14
Sales and clerical	z	z	z	z
Percentage of county employment in industry:				
Agriculture, forestry, and fisheries--				
Farming	0.00480	0.00661	208.20	151.30
Agricultural services	0.201	0.295	4.98	3.39
Forestry	0.700	0.700	1.43	1.43
Fisheries	0.412	0.542	2.42	1.85
Mining--				
Metal mining	0.0571	0.291	17.51	3.44
Oil & gas extraction	0.0448	0.0573	22.32	17.44
Non-metal mining	0.188	0.278	5.31	3.60
Coal mining	0.0329	0.0598	30.38	16.72

See notes at end of table. continued--

Appendix table G-1--Tolerances and Variance Inflation Factors for Independent Variables Included in Model--continued

Variable	Tolerances		Variance inflation factors	
	OLS	Robust weights	OLS	Robust weights
Construction--				
Construction-general	0.131	0.266	7.66	3.76
Heavy construction	0.0628	0.192	15.93	5.22
Construction-special	0.230	0.287	4.35	3.49
Manufacturing--				
Food manufacturing	0.0661	0.0659	15.13	15.17
Tobacco manufacturing	0.883	0.793	1.13	1.26
Textiles	0.0303	0.0231	32.97	43.26
Apparel	0.0365	0.0351	27.43	28.49
Lumber & wood products	0.0336	0.0311	29.72	32.13
Furniture & fixtures	0.122	0.0865	8.23	11.57
Paper & allied products	0.101	0.0814	9.93	12.29
Printing & publishing	0.422	0.291	2.37	3.43
Chemicals	0.0814	0.0907	12.29	11.03
Petroleum & coal products	0.566	0.630	1.77	1.59
Rubber & plastic products	0.161	0.145	6.21	6.91
Leather	0.183	0.168	5.48	5.95
Stone, clay, & glass	0.201	0.137	4.97	7.29
Primary metal industries	0.0863	0.104	11.59	9.66
Fabricated metal products	0.124	0.128	8.09	7.78
Machinery (nonelectrical)	0.0832	0.0620	12.02	16.12
Electrical machinery	0.0912	0.0585	10.97	17.08
Motor vehicles	0.242	0.242	4.13	4.14
Other transport equipment	0.209	0.302	4.79	3.32
Instruments	0.367	0.272	2.72	3.68
Miscellaneous manufacturing	0.362	0.385	2.76	2.60
Transportation and utilities--				
Railroads	0.257	0.285	3.89	3.51
Passenger transit	0.742	0.697	1.35	1.43
Trucking & warehousing	0.309	0.373	3.24	2.68
Water transportation	0.642	0.675	1.56	1.48
Air transportation	0.769	0.721	1.30	1.39
Pipelines	0.800	0.734	1.25	1.36
Transport services	0.846	0.845	1.18	1.18
Communication	0.507	0.522	1.97	1.92
Utilities	0.232	0.290	4.32	3.44

See notes at end of table. continued--

Appendix table G-1--Tolerances and Variance Inflation Factors for Independent Variables Included in Model--continued

Variable	Tolerances		Variance inflation factors	
	OLS	Robust weights	OLS	Robust weights
Wholesale trade	0.117	0.122	8.52	8.20
Retail trade--				
Building materials	0.544	0.529	1.84	1.89
General retail	z	z	z	z
Food stores	0.351	0.353	2.85	2.84
Auto dealers	0.316	0.387	3.17	2.58
Apparel stores	0.465	0.453	2.15	2.21
Furniture stores	0.566	0.555	1.77	1.80
Restaurants, etc.	0.0945	0.109	10.59	9.14
Miscellaneous retail	0.415	0.450	2.41	2.22
Finance, insurance, and real estate--				
Banking	0.351	0.358	2.85	2.79
Brokers	0.844	0.774	1.18	1.29
Insurance carriers	0.559	0.509	1.79	1.97
Insurance agents	0.702	0.688	1.43	1.45
Real estate	0.293	0.351	3.41	2.85
Misc. combined finance	0.855	0.816	1.17	1.22
Holding companies	0.915	0.889	1.09	1.13
Services--				
Hotels/lodging	0.0864	0.0945	11.58	10.58
Personal services	0.486	0.505	2.06	1.98
Misc. business services	0.159	0.355	6.30	2.82
Auto repair	0.725	0.733	1.38	1.36
Miscellaneous repair	0.708	0.749	1.41	1.33
Motion pictures	0.739	0.714	1.35	1.40
Amusements, n.e.c.	0.194	0.342	5.15	2.92
Health services	0.0639	0.0680	15.66	14.70
Legal services	0.734	0.728	1.36	1.37
Educational services	0.230	0.209	4.35	4.78
Social services	0.200	0.230	5.01	4.35
Museums, zoos, etc.	0.917	0.781	1.09	1.28
Membership organizations	0.243	0.249	4.11	4.01
Private household services	0.115	0.127	8.67	7.87
Miscellaneous services	0.564	0.582	1.77	1.72

See notes at end of table.

continued--

Appendix table G-1--Tolerances and Variance Inflation Factors for Independent Variables Included in Model--continued

Variable	Tolerances		Variance inflation factors	
	OLS	Robust weights	OLS	Robust weights
Government--				
Federal government (civilian)	0.0589	0.127	16.96	7.84
Federal government (military)	0.0469	0.0286	21.33	34.98
State & local government	0.0244	0.0232	40.98	43.15
Region:				
New England▒	0.542	0.448	1.84	2.23
Middle Atlantic▒	0.464	0.339	2.15	2.95
East North Central▒	0.195	0.150	5.13	6.67
West North Central▒	0.140	0.122	7.13	8.21
South Atlantic▒	z	z	z	z
East South Central▒	0.423	0.372	2.37	2.69
West South Central▒	0.280	0.303	3.57	3.30
Mountain▒	0.185	0.220	5.40	4.55
Pacific▒	0.302	0.251	3.31	3.99

▒ = 0, 1 variable.
NA = Not estimated or not applicable.
n.e.c. = Not elsewhere classified.
z = Not explicitly included in model (omitted category for estimation purposes).
Note: Tolerance = 1/ (Variance inflation factor).

Source: U.S. Department of Agriculture, Economic Research Service.

Appendix H. Adjusted R-Square and Explanatory Power of Control and Noncontrol Variables

Appendix table H-1 shows the adjusted r^2 values for the preferred and alternative econometric models that were reported in the text in table 2.[1]

Also shown, for weighted data only, is the component of adjusted r^2 that is associated exclusively with each type of variable; that is, the drop in adjusted r^2 when that type of variable is omitted from the model. These values show that commuting-zone effects account for a substantial component of earnings variance that cannot be explained by either noncontrol variables or by industry and occupational mix. Industrial and occupational mix account for a more modest proportion of variance that cannot be explained by commuting-zone effects or other variables, while we find that only a relatively small component of the total variance in earnings growth rates can be explained exclusively by noncontrol variables (that is, variables other than regional dummies and industrial-occupational employment share variables).

Note, however, that these components do not add up to total adjusted r^2, as much of the variance in earnings growth across counties can be accounted for by more than one type of variable, depending on which types are included in the model. Therefore, appendix table H-1 also indicates the components of earnings growth variance that can be explained by more than one of the three groups of independent variables indicated.

1. Differences in the dependent variable, the units of observation, the universe observed, the period of observation, and the estimation techniques used make it difficult to compare these results with those for other studies, such as those reviewed in Kusmin (1994).

Appendix table H-1--Adjusted r-squared and explanatory power of control and noncontrol variables

	Preferred method	Robust weights	OLS	Rest-of-CZ
Adjusted-r-squared	0.6700	0.5107	0.4006	0.5486
Accounted for by:				
Geographic variables only	0.1866	0.0313	0.0213	0.0779
Industry/ occupation controls only	0.0373	0.1268	0.1295	0.0686
Noncontrol variables only	0.0229	0.0594	0.0466	0.0394
Noncontrol or geographic variables	0.0601	0.0295	0.0249	0.0532
Noncontrol or industry / occupation variables	0.0576	0.1208	0.0924	0.0850
Geographic or industry/occupation variables	0.0887	-0.0112	-0.0155	0.0458
Any one of these three groups	0.2169	0.1541	0.1014	0.1787
Adjusted- r-squared for unweighted estimates	0.5534	0.4006	0.4006	0.4458

Source: U.S. Department of Agriculture, Economic Research Service.