# Explaining Rural-Urban Earnings Differentials in the U.S.

Douglas Rhoades
S&P-McGraw Hill, Inc.

Mitch Renkow
North Carolina State University

Prepared for the
1998 American Agricultural Economics Association Annual Meeting
Salt Lake City, Utah

# Explaining Rural-Urban Earnings Differentials in the U.S.

## Introduction

The persistent gap between the economic performance of rural and urban places in the U.S. has been widely noted, and remains a continuing empirical puzzle to social scientists interested in rural development. In 1977 real per capita earnings (in 1988 dollars) in metropolitan commuting zones were about $2600 greater than in nonmetropolitan commuting zones. By 1995 – and despite having fallen somewhat since 1988 – this gap had grown to nearly $3500.

In a recent paper, Renkow (1996) sought to quantify the magnitude of rural-urban[1] differences in the response of earnings to key economic variables, using county-level data from North Carolina. Three empirical findings of interest emerged from his analysis. First, he found that returns to schooling are significantly smaller in rural areas than in urban areas. Second, he found that local economic shocks had a much more profound impact on earnings in rural areas than they did in urban areas. Finally, he found no appreciable difference in the impact of macroeconomic forces on rural areas *vis-a-vis* urban areas. These results are potentially quite important in informing debates over (and the design of) national policies to boost rural economic performance. However, the fact that the data were drawn from only one state – one whose economic performance generally outpaced the rest of the nation in most respects during the time period studied – limits the extent to which one can comfortably generalize these findings to the rest of the nation.

In this paper, we estimate the response of earnings to key economic variables using county-level data from all 48 of the continental United States over the period 1977 to 1995. The breadth

---

[1] In this paper, we use the Bureau of Economic Analysis' definition of metropolitan counties, and use the terms "metropolitan," "metro," and "urban" interchangeably.

and length of the data set provides a rich testing ground for quantifying the underlying forces determining earnings, and how the response to these forces differs between rural and urban areas.

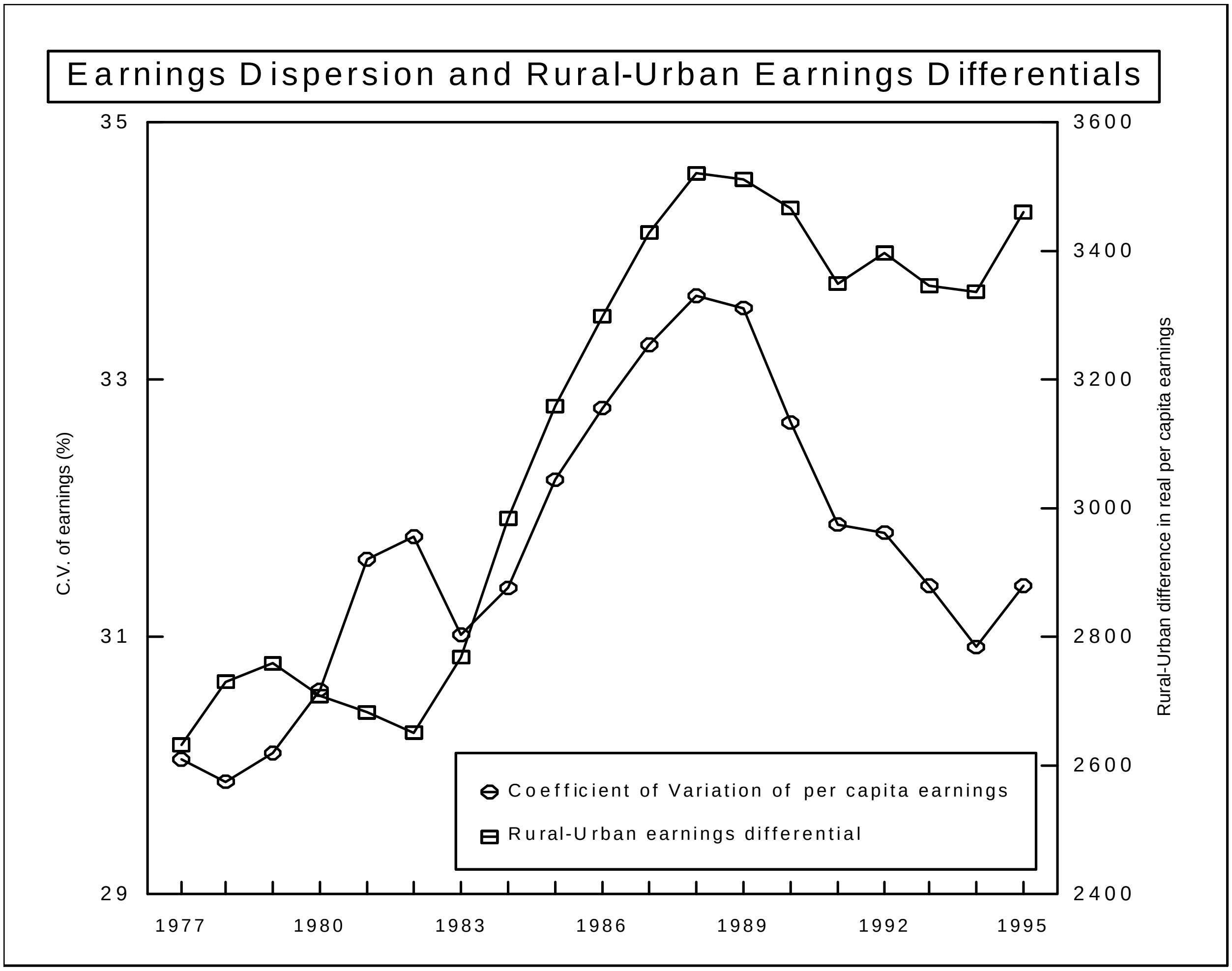
The paper is organized as follows. We first describe the behavior of earnings in rural and urban places across the U.S. over the past two decades. Succeeding sections describe the analytical framework for the empirical model and provide information on data used and some estimation issues. We then present and interpret the econometric results. The final section contains concluding remarks and a summary of important findings.

## Empirical Background

A fair amount of concern continues to be focused on the relative economic conditions of rural areas *vis-a-vis* urban areas. Between 1977 and 1995, real earnings per capita grew in metropolitan commuting zones at 1.33% per year compared to 1.14% in nonmetropolitan commuting zones.[2] The earnings gap widened markedly each year during the period 1982 through 1988 (see Figure 1). This provided an important impetus for a series of government initiatives providing support to farming communities and rural industrial centers – for example, Empowerment Zone and Enterprise Community block grants programs and the establishment of Rural Development Councils. Despite having fallen modestly in five of the seven years after 1988, the rural-urban earnings gap was more than 30% greater in 1995 than it was in 1977.

Figure 1 also indicates that the spatial dispersion of earnings – as measured by the coefficient of variation of earnings per capita across commuting zones – followed a strikingly similar pattern to that of the rural-urban earnings differential, particularly since 1983 (about the

[2] The 1990 definition of commuting zones developed by Tolbert and Sizer (1996) are the primary geographical unit of observation in this paper. Tolbert and Sizer organized commuter flows indicated by the Census Bureau's 1990 journey-to-work data, then employed a hierarchical cluster analyses to group counties on the basis of employment and residential association. Essentially, commuting zones are the best available delineation of discrete spatial labor markets.




time of the large run-up in the earnings gap). Both time series rose strongly throughout the most of 1980s, then tapered off (although earnings dispersion fell much more dramatically than did the earnings gap after 1988), and jumped markedly in 1995.

In previous empirical studies using time periods prior to 1990, incomes were shown to be diverging across countries, regions of the U.S., and states, at least since about the mid-1970s (Barro and Sala-I-Martin, 1990; Carlino, 1992; Coughlin and Mandelbaum, 1988). In contrast, the pattern in Figure 1 appears to indicate a general return to the historic tendency toward income convergence predicted by neoclassical growth theory. Nonetheless, it is interesting to note the apparent connection between earnings dispersion and the economic performance of rural and urban areas. The striking co-movement of these two series suggests that ascertaining the factors underlying the differential economic performance of rural areas *vis-a-vis* urban areas may hold a key to understanding the empirical puzzle embodied by the dynamics of income dispersion over the past two decades. It begs the questions of what are the underlying forces determining real earnings growth, and how does the response to these forces differ between rural and urban areas.

## Analytical Framework

The analysis presented here concerns itself with per capita labor earnings, the largest component of personal income. Earnings ($Y_E$) are the product of labor force participation ($L$) and wages ($W$).[3] Following Tokle and Huffman (1991), we take labor and wages in any geographic area to depend on human capital ($H$); permanent (anticipated) local labor market conditions ($W$); transitory (unanticipated) local labor market conditions ($w$); macroeconomic conditions ($m$); locational amenities ($a$); the age distribution of the population ($d$); and other miscellaneous socioeconomic variables ($q$) such as the ethnic composition of the local population and the industrial or sectoral composition of the local economy. The reduced form equation of per capita earnings can be written as:

$$Y_E = W(H, W, w, m, a, d, q) \times L(H, W, w, m, a, d, q) = Y_E(H, W, w, m, a, d, q) \qquad \textbf{(1)}$$

Estimation of the above function will inform discussion of the effects of key economic variables on earnings, including the significance, magnitude and direction of those effects. In addition, some light may be shed as to the best explanation for rural-urban differences in the intertemporal pattern of earnings indicated above. A possible explanation for these differences is deficiencies in schools and other institutions supporting human capital development in rural

[3] Wages and labor force participation are likely to be affected differently by the same variable. However, it seems reasonable to assume that the *direction* of effects of the key variables of interest here — human capital stock, local labor market conditions and macroeconomic conditions — will be the same.

areas (Ross and Rosenfeld, 1988). This explanation would be supported by differences in the response of $Y_E$ to $H$. Differences in the sensitivities of rural and urban areas to negative macroeconomic phenomena such as price shocks and recessions would be indicated by differences in the response of $Y_E$ to $m$. Sensitivity of earnings to local labor market dynamics indicated by significant impacts of $W$ and $w$ on $Y_E$ .

## Data and Econometric Issues

We estimated equation (1) using county-level data for the 48 contiguous U.S. states, aggregated up to the commuting zone level. Earnings data (wages plus salaries) for period 1976 to 1995 were obtained from the Bureau of Economic Analysis' (BEA) Regional Economic Information System. These were divided by mid-year population estimates to yield per capita earnings. A chain type GNP deflator was used to transform nominal values into real values with a base year of 1992.

We used the proportion of individuals aged 25 and up who had completed a high school degree as a proxy for human capital stock. Educational attainment data are only available for census years (e.g., 1970, 1980, 1990). We generated education data for the intercensal years (1971 through 1979 and 1981 through 1989) by linear interpolation; likewise, education data for the period 1991 through 1995 were linearly forecast based on the 1980 – 1990 trend. This technique is likely to have resulted in unobservable measurement error. To correct for inconsistency and bias due to this errors-in-variables problem, we employed an instrumental variables (IV) estimation approach. It has been shown that if the chosen instruments for the error-ridden variables are correlated with the true independent variable but uncorrelated with the measurement error, the resulting IV estimates will be consistent (Carter and Fuller, 1980).[4] The instruments selected were the U.S. inflation rate and the population density. The education variable was regressed against these variables and then the predicted value were used on the right hand side of the final regression equation.

County-level unemployment data were obtained from the Bureau of Labor Statistic and, as with all other variables, aggregated up to commuting zone level. One-step-ahead ARMA forecasts were made for each commuting zone's unemployment rate. This expected unemployment rate was used as an indicator of permanent (anticipated) local labor market conditions. Correspondingly, unexpected (transitory) labor market shocks are represented by the difference between the forecast and realized value of the unemployment rate.[5]

A number of demographic and sectoral variables were included in our regressions to control for unobservable fixed effects. Commuting zone employment shares of various one-digit SIC categories (for 1990) were used to account for differences in the sectoral composition of local economies. These data were from the BEA. Demographic variables included the proportion of working age males and females in three age ranges (15-24, 25-54, and 55-64); the proportion of children under 15; the proportion of the elderly (aged 65 and older); and the proportion of minorities in the population. The population share data were computed from U.S. Census Bureau county-level annual population estimates.

Additional independent variables included the real U.S. Gross Domestic Product (as a proxy for macroeconomic conditions) and a dummy variable with a value of 1 for metro counties and 0 for nonmetro counties. To test for differences in the impacts of key economic variables between

[4] The estimates using a particular set of instruments need not yield the minimum asymptotic variance, however (Judge, et al., pg. 533).

[5] This procedure is in the same spirit as that used by Tokle and Huffman (1996), but differs in that it uses time-series methods to generate forecasts and residuals instead of fitting quadratic trend equations to the data.

rural and urban areas, interaction (slope) dummies were created by multiplying the metro dummy by the education, expected unemployment, unemployment shock, and GDP variables.

## Results

Our earnings regression was estimated using 19 years of data from each of 722 commuting zones.[6] Given the time-series cross sectional nature of the data, there was ample reason to believe *a priori* that significant correlation among disturbances existed, both contemporaneously (across commuting zones) and serially (within commuting zones). The econometric model was therefore estimated using the TSCSREG procedure of SAS, a generalized least squares estimator that accounts for both forms of error correlation.

Table 1 presents the regression results. The dependent variable was the natural logarithm of per capita earnings. The logs of several righthand side variables of interest (high school graduates, U.S. GDP, and their interaction terms) were also used, primarily to reduce multicollinearity. Most of the key economic variables are of the expected sign, and significant at the 5% level or better. High school education, U.S. GDP, and urban status have a positive impact on earnings, while expected unemployment and unemployment shocks have a negative effect on earnings. The estimated coefficient of the minority population variable was (unexpectedly) positive, although not statistically significant. The coefficients on all demographic variables was negative, which is not surprising given that the omitted demographic category was prime working age males. The negative coefficients on the variables representing sectoral composition of local economies imply that earnings tended to be lower in areas that were more heavily dependent on jobs in the agricultural, manufacturing, government, mining, and service related

[6] One year of data was lost in generating the forecasts of expected unemployment for each commuting zone.

industries. This result is presumably a reflection of the decline of agriculture and traditional manufacturing and service industries relative to (omitted) high technology and financial sectors.

Rural-urban differences in earnings response are reflected in the parameter estimates for the interaction terms. All but one of these – that associated with unemployment shocks – are strongly significant. The magnitude of these rural-urban differences are most readily understood by comparing the implied elasticities for rural and urban commuting zones (see Table 2). Education's impact on earnings is substantially greater in urban areas: the urban elasticity of earnings with respect to education is about 40% larger than that of rural areas (.248 vs. .178). This result is similar to (but not as large as) that found by Renkow (1996) for North Carolina. On the other hand, earnings are strikingly more sensitive to macroeconomic fluctuations in rural areas than urban areas: the rural earnings elasticity with respect to U.S. GDP is nearly four times as great as in urban areas (.109 vs. .029). This result is consistent with the arguments of several authors who assert that income and employment is more strongly affected by business cycle trends in rural areas (Mally and Hady, 1988; Deavers, 1987). Finally, the results indicate that rural areas slightly more sensitive to expected unemployment than urban areas. However, while this difference is statistically significant, it is quite small in magnitude.

The above results seem to fit well with the literature and priori expectations, and are reasonably consistent with the findings of Renkow (1996), who used data only from North Carolina commuting zones. When the same type of model is estimated by census region, however, the results in some cases become difficult to explain (see Table 3). Most notably, the parameter estimates for the education variable vary widely by region and rural-urban status. For example, the results imply that educational attainment has a negative and statistically significant impact on earnings in rural commuting zones in the Northeast and in urban commuting zones in the West. One possible explanation for these results is that high school educational attainment is not the appropriate measurement of the human capital stock in these regions; college level attainment might be a better proxy. We plan to pursue this notion in future research. The rest of the results for the regional regressions are consistent with those of the (pooled) national model, except for the negative coefficient on U.S. GDP in the urban south. Explaining this puzzling result also merits further research.

## Concluding Remarks

In this study we have estimated an econometric model of earnings in attempt to explain some of the reasons for differences in rural and urban earnings. Briefly, the econometric results indicate that *(a)* earnings response to schooling is generally significantly lower in rural areas than in urban areas; *(b)* the impact of macroeconomic forces on rural areas is significantly greater than the impact on urban areas; *(c)* there is no appreciable difference in the impact of local unemployment shocks on rural areas *vis-a-vis* urban areas; and *(d)* there are some interesting regional differences in rural-urban earnings response that merit further research.

**References**


Barro, Robert and Xavier Sala-I-Martin. 1990. "Economic Growth and Convergence across the United States," National Bureau of Economic Research, NBER Working Paper No. 3419.

Carlino, Gerald. 1992. "Are Regional Per Capita Earnings Diverging?" *Federal Reserve Bank of Philadelphia Business Review*, March/April: 3-12.

Coughlin, Cletus and Thomas Mandelbaum. 1988. "Why Have State Per Capita Incomes Diverged Recently?" *Federal Reserve Bank of St. Louis Review*, September/October: 24-36.

Carter, R.L. and Wayne Fuller. 1980 . "Instrumental Variable Estimation of the Simple Errors-in-Variables Model," *Journal of the American Statistical Association*, Sept. 1980, 687-92.

Deavers, Kenneth. 1987. "Rural Economic Conditions and Development Policy for the 1980s and 1990s," in Gene Summers, John Bryden, Kenneth Deavers, Howard Newby, and Susan Sechler (eds.), *Agriculture and Beyond: Rural Economic Development*, Madison, WI: University of Wisconsin, pp. 113-24.

Judge, George, William Griffith, R. Carter Hill, and Tsoung-Chao Lee. 1980. *The Theory and Practice of Econometrics*. New York: Wiley and Sons.

Mally, James and Thomas Hady. 1988. "The Impact of Macroeconomic Policies on Rural Employment," in David Brown et al. (eds.), *Rural Economic Development in the 1980s: Prospects for the Future*, Washington, D.C.: U.S. Department of Agriculture, pp. 221-34.

McGranahan, David and Linda Ghelfi. 1991. "The Education Crisis and Rural Stagnation in the 1980's," in *Education and Rural Economic Development*. ERS Staff Report No. AGES 9153. Washington: U.S. Department of Agriculture, 40-91

Renkow, Mitch. 1996 . "Income Non-Convergence and Rural Urban Earnings Differentials: Evidence from North Carolina," *Southern Economic Journal* 62, 1017-1028.

Ross, Peggy and Stuart Rosenfeld. 1988. "Human Resource Policies and Economic Development," in David Brown et al. (eds.), *Rural Economic Development in the 1980s: Prospects for the Future*, Washington, D.C.: U.S. Department of Agriculture, pp. 333-58.

Tolbert, Charles M. and Molly Sizer. 1996 . "U.S. Commuting Zones and Labor Market Areas: A 1990 Update," ERS Staff Paper No. AGES-9614. Washington: U.S. Department of Agriculture.

**Table 1. Regression results[a]**

| Variable | Estimate | | Std. Error |
|---|---|---|---|
| *ln* (High school grads) | 0.178 | *** | 0.012 |
| *ln* (High school grads) × metro dummy | 0.070 | *** | 0.020 |
| *ln* (U.S. GDP) | 0.103 | ** | 0.046 |
| *ln* (U.S. GDP) × metro dummy | -0.075 | *** | 0.028 |
| Expected unemployment | -2.213 | *** | 0.036 |
| Expected unemployment × metro dummy | 0.095 | *** | 0.034 |
| Unemployment shock | -1.398 | *** | 0.019 |
| Unemployment shock × metro dummy | 0.029 | | 0.032 |
| Metro dummy | 0.788 | *** | 0.246 |
| Non-white population (%) | 0.016 | | 0.063 |
| Males 15-24 (%) | -0.608 | | 0.686 |
| Males 55-64 (%) | -1.583 | | 1.973 |
| Females 15-24 (%) | -1.343 | ** | 0.568 |
| Females 25-54 (%) | -1.219 | ** | 0.802 |
| Females 55-64 (%) | -6.646 | *** | 1.592 |
| Elderly (%) | -2.779 | *** | 0.371 |
| Children (%) | -1.940 | *** | 0.382 |
| Agricultural employment share | -2.191 | *** | 0.290 |
| Service & retail employment share | -1.270 | *** | 0.309 |
| Manufacturing employment share | -0.796 | *** | 0.191 |
| Government employment share | -1.429 | *** | 0.243 |
| Mining employment share | -0.791 | | 0.629 |
| Intercept | 10.381 | *** | 0.584 |
| $R^2$ | .987 | | |
| Sample size | 13,718 | | |

a. IV estimates using population density and U.S. inflation as instruments for the education variable (see text). Dependent variable is the log of real earnings per capita. Standard errors of estimates are found in parentheses. ** and *** indicate significance at the 5% and 1% levels, respectively.

**Table 2. Earnings elasticities with respect to key variables**

| Variable | Metro | Rural |
|---|---|---|
| High school graduates | .248 | .178 |
| U.S. GDP | .029 | .103 |
| Expected unemployment rate[a] | -.149 | -.158 |

a. Evaluated at the sample mean.

**Table 3. Earnings elasticities for different regions[a]**

| Variable | Metro | Rural |
|---|---|---|
| ***Northeast*** | | |
| High school graduates | .002 *** | -.047 *** |
| U.S. GDP | .046 *** | .144 *** |
| Expected unemployment rate[b] | -.118 *** | -.100 *** |
| ***Midwest*** | | |
| High school graduates | .034 *** | .105 *** |
| U.S. GDP | .008 *** | .104 ** |
| Expected unemployment rate[b] | -.137 *** | -.142 *** |
| ***South*** | | |
| High school graduates | .315 *** | .082 *** |
| U.S. GDP | -.106 *** | .122 *** |
| Expected unemployment rate[b] | -.190 *** | -.151 *** |
| ***West*** | | |
| High school graduates | -.292 *** | -.027 [NS] |
| U.S. GDP | .284 *** | .094 ** |
| Expected unemployment rate[b] | -.143 *** | -.198 *** |

a. **, and *** denote significance at the .05 and .01 level. "NS" indicates not statistically significant.

b. Evaluated at the sample mean.