WORKING PAPER SERIES

ON THE EMPIRICAL FINDING OF A HIGHER RISK OF POVERTY IN RURAL AREAS: IS RURAL RESIDENCE ENDOGENOUS TO POVERTY?

Monica Fisher

RPRC Working Paper No. 04-09

August 2004

Rural Poverty Research Center

http://www.rprconline.org/

RUPRI Rural Poverty Research Center
214 Middlebush Hall
University of Missouri
Columbia MO 65211-6200
PH 573 882-0316

RUPRI Rural Poverty Research Center
Oregon State University
213 Ballard Hall
Corvallis OR 97331-3601
PH 541 737-1442

ON THE EMPIRICAL FINDING OF A HIGHER RISK OF POVERTY IN RURAL AREAS: IS RURAL RESIDENCE ENDOGENOUS TO POVERTY?

MONICA FISHER

IS RURAL RESIDENCE ENDOGENOUS TO POVERTY?

August 23, 2004

Monica Fisher is Research Associate at University of Missouri, Truman School of Public Policy and Rural Poverty Research Center, Oregon State University. This research was supported by the Rural Poverty Research Center of the Rural Policy Research Institute, Columbia, Missouri. I thank Ken Foster, Linda Lobao, Andrew Plantinga, and Bruce Weber for valuable comments. The author alone is responsible for any substantive or analytic errors.

ON THE EMPIRICAL FINDING OF A HIGHER RISK OF POVERTY IN RURAL AREAS: IS RURAL RESIDENCE ENDOGENOUS TO POVERTY?

Abstract

Research shows households are more likely to be poor in rural versus urban America. Does this phenomenon partly reflect that people who choose rural residence have unmeasured attributes related to human impoverishment? To address this, two models are estimated using Panel Study of Income Dynamics data. A single equation Probit model of household poverty replicates the well-documented finding of higher poverty risk in rural places. However, a two-stage instrumental variables approach accounting for residential choice finds no measured effect of rural location on poverty. Results suggest failure to correct for endogenous rural residence leads to over-estimation of the "rural effect".

Keywords: endogeneity, households, instrumental variables, poverty, rural

ON THE EMPIRICAL FINDING OF A HIGHER RISK OF POVERTY IN RURAL AREAS: IS RURAL RESIDENCE ENDOGENOUS TO POVERTY?

Introduction

The incidence of poverty is higher in nonmetropolitan (nonmetro) than metropolitan (metro) areas and, as shown in Figure 1, this phenomenon is not new. Regression analyses also document a rural welfare disadvantage.[1] In the literature rural-urban differences are identified by including a binary variable for nonmetro residence as a regressor in empirical models of individual or household poverty (e.g. Brown and Hirschl; Cotter; Haynie and Gorman; Snyder and McLaughlin; Thompson and McDowell). Analysts then attempt to control for important individual/household characteristics (e.g. race and education of the household head; family structure) and contextual factors (e.g. county unemployment rate and region of residence) that influence well-being. Extant research shows that the odds of being poor are between 1.2 to 2.3 times higher for individuals or households residing in rural compared with urban areas (Weber and Jensen, Table 1).

While a strong positive correlation between rural residence and poverty is well documented in the literature, the direction of causation is indeterminate. Existing work assumes a causal link that goes from rural residence to poverty. Yet the positive association could instead reflect that poor people are attracted to rural areas, or otherwise reluctant (or unable) to leave them. Furthermore, there may be unmeasured factors that determine both poverty and rural residence and, subsequently, the measured impact of rural residence may be partly spurious. In sum, the nature of the link between poverty and rural residence is an open question and one that merits further empirical exploration given its relevance to policy.

This study asks if the estimated rural effect on poverty partly reflects residential selection bias. Current models of rural poverty treat nonmetro residence as an exogenous variable. The validity of this assumption is questionable, because people have some degree of freedom to choose where they live. If people who decide to live in rural areas have unmeasured attributes that are related to human impoverishment, estimates of a rural effect can be biased. Consider, for example, that poverty models rarely control for whether an individual is geographically mobile. It is plausible that, compared to urban people, rural people are less mobile, having a preference for living close to their extended family and childhood friends. Geographic mobility may also be negatively correlated with poverty; those who are willing to move in search of employment may be less likely to be unemployed and poor. If mobility is negatively correlated with both poverty and rural residence, then the effect on poverty of living in a rural area could be overstated if one does not include a proxy variable for mobility in the empirical model.

In this paper I use Panel Study of Income Dynamics (PSID) data to investigate the extent to which endogenous rural residence biases estimates of a rural effect on poverty. Two empirical models are estimated for comparative purposes. First I estimate a single equation Probit model of household poverty in order to replicate the well-documented finding of a higher risk of being poor in rural compared to urban places. I then use an instrumental variables approach to account for the possibility that rural residence is endogenous to poverty. The basic question addressed is whether a rural effect persists when one accounts for rural residential choice; or, does such an effect disappear suggesting it may be an artifact of residential sorting. While the paper focuses on links between rural residence and the risk of being poor, findings are relevant to a general body of research that measures place effects on individual behavior and well-being.

Data and Sampling Issues

This study uses data from the 1993 and 1994 waves of the Panel Study of Income Dynamics (PSID), a longitudinal survey that has followed a representative sample of about 5,000 families and their descendents since 1968.[2] The PSID family and individual files contain data on a wide range of topics including family structure and demographics, socio-economic background, geographic mobility, employment, earnings, income, wealth, welfare participation, housework time, health, and food security. The dataset is particularly useful for the analyses of this paper because it is one of only two nationally-representative datasets that provides, for public use, information on nonmetro/metro residence for certain years.[3]

The analyses focus on a sample of 6,750 households in 1993. I choose 1993 as the analysis year because it is the most recent year for which all of the required data for the analyses are available. In selecting this analysis year, two main factors are relevant. First, a variable for nonmetro/metro residence is not available for all years; such a variable is provided in 1968-1993, 1999, and 2001. Second, structural condition variables such as the county unemployment rate are only provided until 1993.

Several criteria were used to arrive at the final sample of 6,750 households. First, I focus on households headed by individuals whose main race is either black or white, because the PSID contains insufficient numbers of individuals of other racial and ethnic groups to enable systematic study. Second, only households that resided in the United States during the survey year are included in the sample. Finally, I drop all observations with incomplete data for the variables used in the analyses.

Empirical Analysis

Base Model Results for Household Poverty

I begin by estimating an empirical model of household poverty in which nonmetro residence is assumed to be an exogenous variable, paralleling the common practice of existing work. The model is a single equation Probit model of the form

(1) $$p = \alpha_0 + \alpha_1 x + + \alpha_2 u + \alpha_3 r + \alpha_4 n + \varepsilon_1 .$$

Dependent variable *p* is a binary variable indicating whether the household is poor, defined as having before-tax cash income less than or equal to 100 % of the family-size conditioned official poverty thresholds. Explanatory variables are individual and household characteristics **x** (including the age of the youngest child and number of household members as well as the householder's age, race, gender, marital status, education, work experience, and disability status), the county unemployment rate *u*, indicator variables for the region of residence **r**, and a binary variable *n* indicating whether the county of residence is nonmetro. Table 1 provides summary statistics for each of the variables used in the analysis.

The empirical model captures the main determinants of human impoverishment highlighted by poverty researchers (e.g. Rank, Yoon, and Hirschl; Schiller). One common view is that specific attributes of poor people, such as low levels of education or lack of competitive labor market skills, have brought about their poverty. From this individualist perspective, poverty is a consequence of individual decisions related to education, employment, and family structure; these decisions in turn have implications for economic well-being. Other observers argue that poverty is mainly the result of restricted educational, economic, and political opportunities. Restricted opportunities may be related to one's place of residence (e.g. neighborhood, county, region), or they may originate from discrimination on the basis of gender, race, or class. Thus, according to the restricted opportunity viewpoint, poverty is conditioned by forces beyond the control of individuals and households. These two explanations of poverty are here considered complementary as reflected in equation 1.

Table 2 presents Probit results for the single equation model of household poverty. The table reports coefficients, Huber/White robust standard errors, and marginal effects. Note that for binary variables, the marginal effects are interpreted as the percentage point change in the probability of poverty resulting from a discrete change in the explanatory variable. At standard test levels, most of the parameter estimates are individually significant and a Wald test indicates joint significance of explanatory variables. Parameter estimates for the variables age and age squared indicate that age of the household head is negatively correlated with poverty until the age of 43 years, at which point the correlation becomes positive. Results show a negative correlation between household poverty and the household head's education and work experience. Findings suggest that households are more likely to be poor if they are headed by an individual who has a temporary or permanent disability, is African American, and is not married. The risk of poverty is lower for households with more adult members and higher for households with more children. Finally, place of residence appears to matter to poverty outcomes. Households are less likely to be poor if they reside in the Northeast compared with the South. And, findings show that household poverty is negatively correlated with the county unemployment rate and nonmetro residence. Results reported in table 2 are consistent with empirical findings elsewhere (e.g. Brown and Hirschl; Cotter; Thompson and McDowell). Importantly, I confirm the result of previous studies that living in a rural area increases the risk of poverty, all else being equal. Nonmetro households have a 19 % higher probability of being poor compared with households in metro areas.[4]

Accounting for Endogeneity of Rural Residence

A maintained hypothesis is that current estimates of a rural effect on poverty are biased because residence in a nonmetro area is a choice influenced by unobserved individual characteristics. There are two main ways that one can deal with this residential selection problem. The first, instrumental variables, or two-stage least squares, identifies and exploits an exogenous source of variation in residential choice; the second, fixed-effects strategies, involves introducing controls for individual heterogeneity (Weinberg, Reagan, and Yankow). Each approach has advantages and drawbacks but, for the purposes here, instrumental variables is more appropriate.[5] Below I turn to a simultaneous equation model of poverty, asking whether and to what extent failure to account for rural residential choice biases the measured rural effect.

A two-stage instrumental variables approach is used to account for potential endogeneity of rural residence. In the first stage, a Probit model of nonmetro residence is estimated. The model assumes the probability an individual resides in a nonmetro location n is a function of individual-level variables that determine poverty **x**, the county unemployment rate u, region of residence **r**, and a set of identifying instruments **z** assumed to affect residential choice but not whether an individual is poor. The residential choice model is

(2) $$n = \beta_0 + \beta_1 x + \beta_2 u + \beta_3 r + \beta_4 z + \varepsilon_2 .$$

In the second stage, the Probit model of household poverty (equation 1) is estimated, where I replace observed nonmetro residence with predicted nonmetro residence, purged of its potentially spurious correlation with omitted variables.

Table 3 presents results from the first-stage regression. Findings indicate a negative correlation between nonmetro residence and the education of the household head, consistent with Census Bureau data documenting lower educational attainment in nonmetro compared to metro areas. Results in table 3 suggest that nonmetro households are less likely to be headed by an individual who is African American and unmarried. The number of adult household members is negatively correlated with nonmetro residence, while the number of children is positively correlated with nonmetro residence. Findings show a positive association between rural residence and the county unemployment rate. Finally, results indicate that southern households are more likely to live in nonmetro areas compared with households in the northeastern and western parts of the United States.

Two binary variables are used as instruments to identify the poverty equation. The first instrument is an indicator variable for whether the household head grew up in a city of any size, as opposed to growing up in a rural area, town, suburb, or combination of places. The second instrument is a binary variable for whether the household head is atheist or agnostic, religious preferences that are not well represented in rural America. Both of the instruments are strong predictors of nonmetro residence. The validity of the chosen instruments is discussed in a later section of the paper.

Table 4 presents second-stage Probit results for household poverty. Focusing first on all explanatory variables other than the nonmetro residence binary variable, we see that coefficient estimates are very similar for the two models. The sign on each of these variables is the same across equations and differences in magnitude are quite small. In addition, the set of statistically significant explanatory variables is essentially the same for the base model and two-stage Probit model (the exception is that the binary variable for residence in the west is statistically significant in the two-stage Probit model, but not the base model).

Turning to the parameter estimates for the nonmetro residence variable, findings are quite striking. Whereas the rural effect is positive and statistically significant in the base model, in the two-stage Probit model this effect disappears. Predicted nonmetro residence has an unexpected negative sign in table 4 and is statistically insignificant. That is, accounting for endogenous rural residence, there is no measured effect of living in a nonmetro location on household poverty. This finding parallels results reported by Evans, Oates, and Schwab who find that when one accounts for peer group choice in empirical models of teenage pregnancy and school dropout behavior, estimated peer group effects vanish. The finding here of no measured rural effect is at odds with a large literature that documents a rural disadvantage in poverty outcomes (see Weber and Jensen for a review). It is thus imperative to evaluate the instrumental variables method to assess the extent to which findings reported in table 4 are reliable.

Evaluating the Instrumental Variable Strategy

In this section I discuss some possible concerns with the two-stage Probit strategy. One issue is whether rural residence is in fact endogenous. Although economic theory suggests location of residence is a choice, this assumption should be tested. To explore this issue, I conduct an exogeneity test, using the approach proposed by Smith and Blundell for simultaneous limited dependent variable models. This test is essentially one for exclusion of residuals from an auxiliary regression of rural residence on all exogenous variables and instruments. As above, the instruments are a binary variable indicating whether the householder grew up in a city and a dummy variable for whether the household head has a religious preference for a religion not well represented in rural areas. I employ a Stata program "probexog" to implement the Smith-Blundell test. Findings indicate that one can reject statistical exogeneity of nonmetro residence at the 0.05 probability level. Thus it appears that instrumental variables methods are warranted for estimating the relationship between nonmetro residence and household poverty.

A second concern relates to the validity of the chosen instruments, because the two-stage instrumental variables strategy only corrects for endogeneity to the extent that good instruments are employed (Dietz; Newhouse and McClellan). To be valid the instruments should be highly correlated with rural residence and uncorrelated with the error term of the household poverty model. Results in table 3 indicate that the instruments are valid to the first-stage regression; the instruments are individually and jointly significant in the nonmetro residence model.

The second condition for instrument validity is more difficult to assess. A Sargan test of overidentifying restrictions is often implemented for such purpose. Under the null hypothesis that the instruments are uncorrelated with the error term of the poverty equation, the Sargan statistic is distributed as a χ^2 with degrees of freedom equal to the number of overidentifying restrictions (the number of instruments less the number of regressors) (Davidson and MacKinnon).[6] Unfortunately, the Sargan test is not directly applicable to the two-stage Probit model. To conduct the test, I re-estimate the simultaneous model using two-stage least squares. The calculated Sargan statistic is 0.153, while the critical χ^2 (1) at the 0.05 probability level is 3.84. Thus, the null hypothesis can not be rejected. In sum, the instruments appear to be valid.[7]

Conclusion

This study used Panel Study of Income Dynamics data (N = 6,750) to examine whether living in a rural area is associated with a higher risk of poverty in the United States. Two estimation strategies were employed. One, I estimated a single-equation Probit model of the probability of being poor as a function of characteristics of the household, county unemployment rate, region of residence, and nonmetro residence. This model assumes that nonmetro residence is an exogenous variable, paralleling common practice in the rural poverty literature. Empirical results confirm the well-documented finding of a higher risk of poverty in rural places. I find that nonmetro households have a 19 % higher probability of being poor compared with metro households.

The second modeling approach involved estimating a two-stage Probit model that relaxes the assumption of rural residence exogeneity. In the first stage, nonmetro residence was modeled as a function of household characteristics, county unemployment rate, region of residence, and a set of identifying instruments. In the second stage, predicted nonmetro residence from the first-stage regression replaced observed rural residence in the model of poverty probability. Findings of the two-stage Probit model indicate that nonmetro residence has no measured effect on household poverty, all else being equal. Tests for the validity of instruments used to identify the poverty model provide some support for the choice of instruments. A Smith-Blundell test for exogeneity of rural residence indicates that living in a nonmetro area is a choice, suggesting that the two-stage Probit model provides a more reliable means to measure the links between rural residence and household poverty compared with the single equation Probit model. In tandem, empirical findings show that failure to account for residential endogeneity in empirical modeling leads to an over-estimation of the effect of rural residence on household poverty.

Clearly, empirical findings of the present study should not be taken as definitive. Instead the paper provides mainly methodological insights. I respond to Weber and Jensen's call for more carefully specified modeling of the causal effects of rural residence on individual behavior and well-being. The authors argue that until endogenous nonmetro residence and other methodological challenges are addressed, it is not possible to make broad conclusions about rural-urban differences in welfare participation, employment, and poverty. Results here do not rule out the possibility that living in a rural area is a factor that causes poverty in the United States. Rather the study contributes to existing work mainly by highlighting the need to test and, if necessary, correct for endogeneity in the econometric measurement of the effects of rural residence on poverty outcomes. Future work using other nationally-representative datasets, covering a range of analysis years, and employing alternative estimation strategies to correct for residential endogeneity will enable an improved assessment of the extent to which there exists a rural disadvantage in welfare outcomes in the United States. The answer to this question has implications for future research on rural poverty. If empirical studies suggest that rural residence is an important determinant of poverty, then a key area for research might be to improve our understanding of the specific structural conditions that foster rural poverty. If, however, accounting for residential selection, there exists no measured effect of nonmetro residence on poverty, then at least two research questions seem particularly important. One, why do people with certain attributes related to human impoverishment choose to live in rural places? And two, what combination of human-capital and community-strengthening policies are most likely to reduce rural poverty and its unfavorable consequences?

Endnotes

[1] The terms "nonmetro" and "rural" are used interchangeably in this paper to refer to counties outside of metropolitan areas.

[2] See Brown, Duncan, and Stafford, and Hill for detailed descriptions of the PSID.

[3] The main national surveys used for poverty research are the PSID, the Current Population Survey (CPS), the Survey of Income and Program Participation (SIPP), the National Longitudinal Survey of Youth (NLSY), and the National Survey of America's Families (NSAF). Of these surveys, only the PSID and the CPS provide public-use access to data on metro/nonmetro residence. Furthermore, in the CPS, a number of observations are suppressed for the area variables in order to protect the anonymity of respondents.

[4] Marginal effects in the Probit model indicate percentage point rather than percentage change. To arrive at this percentage figure I divided the marginal effect by the predicted probability of being poor (0.1032).

[5] One fixed effects approach involves the use of data from multiple siblings of households to difference out fixed family effects (e.g. Aaronson). This helps reduce the bias associated with unobserved family factors that influence both neighborhood choice and other individual behaviors; but the method is data intensive and is not particularly useful for studies measuring contemporaneous neighborhood effects on adults. Panel data regression with individual fixed effects has also been used in the attempt to distinguish causal neighborhood effects from neighborhood choice (e.g. Weinberg, Reagan, and Yankow). This approach can only account for neighborhood selection related to time-invariant individual factors (Dietz).

[6] The Sargan statistic is equal to the number of observations times the R-square from a regression in which the residuals of the IV estimate of the poverty equation are regressed on the instruments (Davidson and MacKinnon).

[7] I also use an omitted-variable regression version of the Hausman test (Spencer and Berk) to test for exogeneity of predicted nonmetro residence. The calculated F-statistic (table 4) is less than the critical value of 3.00, implying the predicted nonmetro residence variable can be treated as exogenous.

References

Aaronson, D. "Using Sibling Data to Estimate the Impact of Neighborhoods on Children's Educational Outcomes." *J. Human Res.* 33, 4(1998): 915-46.

Brown, C., G.J. Duncan, and F.P. Stafford. "Data Watch: The Panel Study of Income Dynamics," *J. Econ. Perspect.* 10, 2(1996), 155-68.

Brown, D.L., and T.A. Hirschl. "Household Poverty in Rural and Metropolitan-Core Areas of the United States." *Rural Sociology* 60, 1(1995): 44-66.

Cotter, D.A. "Poor People in Poor Places: Local Opportunity Structures and Household Poverty." *Rural Sociology* 67, 4(2002): 534-55.

Davidson, R., and J. G. MacKinnon. *Estimation and Inference in Econometrics*. New York: Oxford University Press, 1993.

Dietz, R.D. "The Estimation of Neighborhood Effects in the Social Sciences: An Interdisciplinary Approach." *Soc. Sci. Res.* 31(2002): 539-75.

Duncan, G.J., and S. Raudenbush. "Neighborhoods and Adolescent Development: How Can We Determine the Links?" In *Does it Take a Village? Community Effects on Children, Adolescents, and Families*, eds., A. Booth and A.C. Crouter. State College, PA: Pennsylvania State University Press, 2001.

Evans, W.N., W.E. Oates, and R.M. Schwab. "Measuring Peer Group Effects: A Study of Teenage Behavior." *J. Polit. Economy* 100, 5(1992): 966-91.

Haynie, D.L., B.K. Gorman. "A Gendered Context of Opportunity: Determinants of Poverty Across Urban and Rural Labor Markets." *Sociological Quart.* 40, 2(1999): 177-97.

Hill, M.S. *The Panel Study of Income Dynamics: A User's Guide*. Newbury Park, CA: Sage Publications, 1992.

Newhouse, J.P., and M. McClellan. “Econometrics in Outcomes Research: The Use of Instrumental Variables.” *Annual Rev. Public Health* 19(1998): 17-34.

Rank, M.R., H. Yoon, and T.A. Hirschl. “American Poverty as a Structural Failing: Evidence and Arguments.” *J. Sociology Soc. Welfare* 30, 4(2003): 3-29.

Sargan, J. D. “The Estimation of Economic Relationships using Instrumental Variables.” *Econometrica* 26 (1958): 393-415.

Schiller, B.R. *The Economics of Poverty and Discrimination*. Englewood Cliffs, NJ: Prentice Hall, 1995.

Smith, R.J., and R.W. Blundell. “An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply.” *Econometrica* 54, 3 (1986): 679-86.

Snyder, A.R., and D.K. McLaughlin. “Female-Headed Families in Poverty in Rural America.” *Rural Sociology* 69, 1(2004): 127-49.

Spencer, D.E., and K.N. Berk. “A Limited Information Specification Test.” *Econometrica* 49, 4 (1981): 1079-85.

Thompson, A., and D.R. McDowell. “Determinants of Poverty Among Workers in Metro and Nonmetro Areas of the South.” *Rev. Black Polit. Economy* 22, 4(1994): 159-77.

Weber, B., and L. Jensen. “Poverty and Place: A Critical Review of Rural Poverty Literature.” RPRC Working Paper 04-03 Rural Poverty Research Center, Columbia, MO, June 2004.

Weinberg, B.A., P.B. Reagan, and J.J. Yankow. “Do Neighborhoods Affect Hours Worked: Evidence from Longitudinal Data.” *J. Lab. Econ.*, forthcoming 2004.

Table 1. Descriptive Statistics of Explanatory Variables, Households 1993

	Mean or Frequency [a]	Standard Deviation [a]
Endogenous variables		
Poor (income <= official threshold)	0.116	
County of residence is nonmetropolitan	0.259	
Exogenous variables		
Age householder (years)	47.515	17.204
Householder’s education (years)	12.818	2.768
Householder’s work experience since age 18 (years)	18.754	13.668
Householder is disabled	0.031	
Householder is African American	0.129	
Household type		
Female headed	0.297	
Male headed	0.158	
Married couple [b]	0.545	
Number of adults	1.742	0.710
Number of children	0.649	1.046
Dummy for child < 6 years	0.149	
County unemployment rate	7.221	2.387
Region		
Northeast	0.213	
North central	0.280	
West	0.180	
South	0.326	
Identifying instruments		
Householder grew up in a city	0.334	
Householder is atheist or agnostic	0.246	
Number of observations [c]		6,357

a. Means and standard deviations are weighted by the PSID core sample household weight.
b. Head is married or has a cohabitator with whom he/she has lived for at least one year.
c. Number of observations is less than the sample size of 6,750 because the household weight variable is not available for all observations.

Table 2. Probit Results for Household Poverty [a]

	Coefficient	Robust Std. Err.	Marginal Effect
Constant	* 0.5177	0.2461	
Age (years)	* -0.0220	0.0085	-0.0039
Age squared	* 0.0003	0.0001	0.00005
Education (years)	* -0.1470	0.0098	-0.0263
Work experience (years)	* -0.0088	0.0027	-0.0016
Disabled	* 0.8740	0.0971	0.2385
African American	* 0.5408	0.0490	0.1073
Household type (married couple excluded)			
Female headed	* 0.7464	0.0609	0.1606
Male headed	* 0.6203	0.0762	0.1441
Number of adults	* -0.2105	0.0393	-0.0378
Number of children	* 0.1730	0.0209	0.0311
Child < 6 years	* 0.1987	0.0601	0.0382
County unemployment	* 0.0348	0.0093	0.0063
Region (South excluded)			
Northeast	* -0.2726	0.0762	-0.0432
North central	-0.0479	0.0550	-0.0085
West	-0.0765	0.0722	-0.0133
Observed nonmetro residence	* 0.1035	0.0536	0.0192
Number of observations			6,750
Wald Chi-Square (16) [b]			1,053.26
Pseudo R-squared			0.29

a. * implies significance at the 0.05 probability level or better. Huber/White robust standard errors reported.

b. Wald test for joint significance of the explanatory variables, distributed as a χ^2 with a critical value of 26.30 for 16 degrees of freedom at 0.05 probability.

Table 3. First-Stage Probit Results for Rural/Urban Residential Choice [a]

	Coefficient	Robust Std. Err.	Marginal Effect
Constant	-0.0001	0.2127	
Age (years)	-0.0114	0.0076	-0.0031
Age squared	0.0001	0.0001	0.00003
Education (years)	* -0.0432	0.0073	-0.0116
Work experience (years)	-0.0032	0.0025	-0.0009
Disabled	0.0422	0.0975	0.0115
African American	* -0.5996	0.0463	-0.1483
Household type (married couple excluded)			
Female headed	* -0.3510	0.0575	-0.0882
Male headed	* -0.1444	0.0678	-0.0369
Number of adults	* -0.0832	0.0336	-0.0223
Number of children	* 0.0503	0.0196	0.0135
Child < 6 years	-0.0767	0.0563	-0.0202
County unemployment	* 0.1531	0.0085	0.0411
Region (South excluded)			
Northeast	* -0.9577	0.0703	-0.1821
North central	-0.0707	0.0451	-0.0186
West	* -0.8765	0.0674	-0.1731
Identifying instruments			
Head grew up in a city	* -1.0088	0.0478	-0.2399
Head is atheist or agnostic	* -0.3156	0.0580	-0.0771
Number of observations			6,750
Wald Chi-Square (17) [b]			1,124.49
Wald Chi-Square (2) [c]			485.07
Pseudo R-squared			0.21

a. * implies significance at the 0.05 probability level or better. Huber/White robust standard errors reported.

b. Wald test for joint significance of the explanatory variables, distributed as a χ^2 with a critical value of 27.59 for 17 degrees of freedom at 0.05 probability.

c. Wald test for joint significance of the instruments, distributed as a χ^2 with a critical value of 5.99 for two degrees of freedom at 0.05 probability.

Table 4. Second-Stage Probit Results for Household Poverty [a]

	Coefficient	Robust Std. Err.	Marginal Effect
Constant	* 0.6731	0.2516	
Age (years)	* -0.0231	0.0085	-0.0042
Age squared	* 0.0003	0.0001	0.00005
Education (years)	* -0.1529	0.0100	-0.0275
Work experience (years)	* -0.0090	0.0027	-0.0016
Disabled	* 0.8787	0.0972	0.2407
African American	* 0.4625	0.0578	0.0907
Household type (married couple excluded)			
Female headed	* 0.7073	0.0627	0.1512
Male headed	* 0.5961	0.0766	0.1375
Number of adults	* -0.2161	0.0395	-0.0389
Number of children	* 0.1773	0.0210	0.0319
Child < 6 years	* 0.1890	0.0601	0.0364
County unemployment	* 0.0518	0.0115	0.0093
Region (South excluded)			
Northeast	* -0.3844	0.0882	-0.0580
North central	-0.0754	0.0555	-0.0132
West	* -0.1793	0.0818	-0.0298
Predicted nonmetro residence	-0.2751	0.1647	-0.0495
Number of observations			6,750
Wald Chi-Square (16) [b]			1,054.44
Sargan statistic [c]			0.15
Hausman statistic [d]			1.60
Pseudo R-squared			0.29

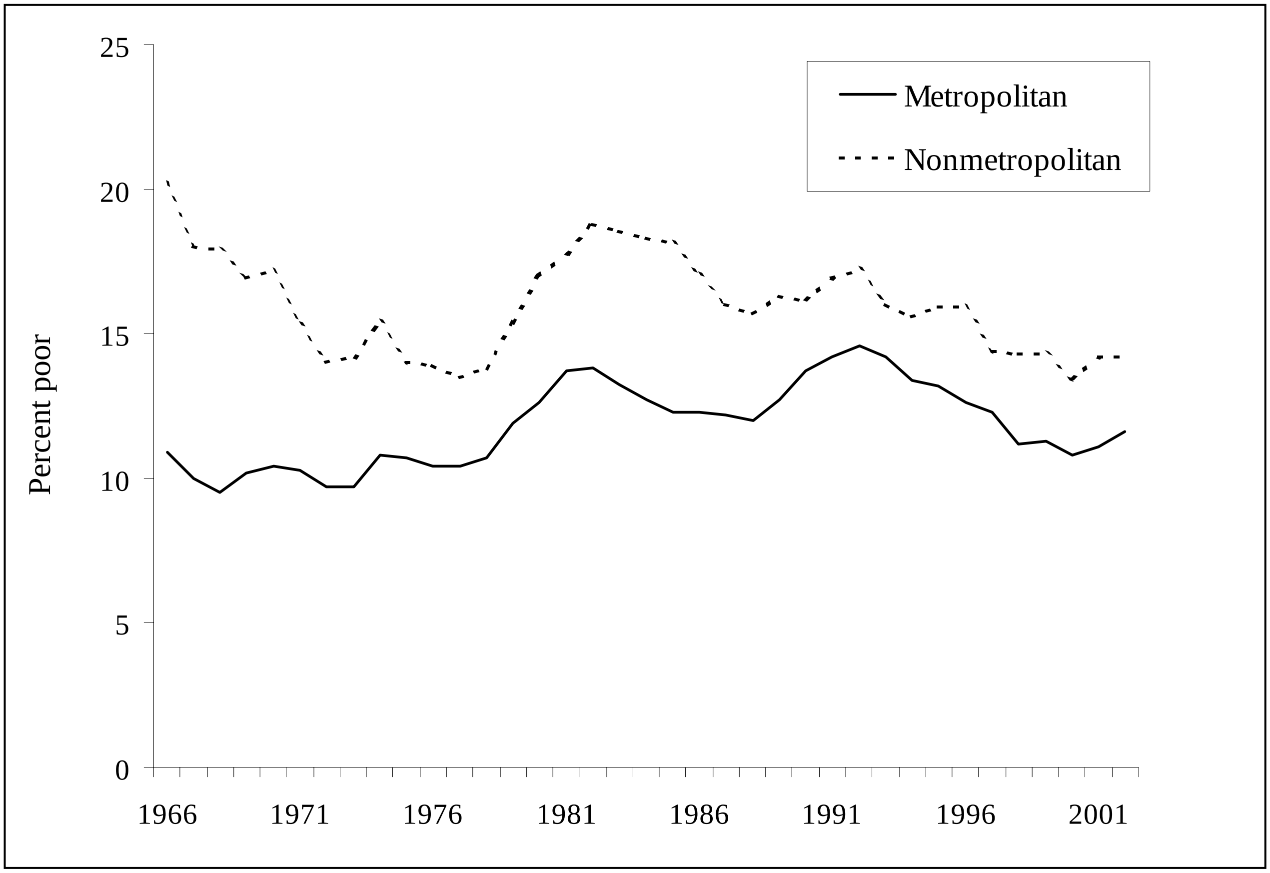
a. * implies significance at the 0.05 probability level or better. Huber/White robust standard errors reported.

b. Wald test for joint significance of the explanatory variables, distributed as a χ^2 with a critical value of 26.30 for 16 degrees of freedom at 0.05 probability.

c. Sargan test of overidentifying restrictions, distributed as a χ^2 with a critical value of 3.84 for 1 degree of freedom at 0.05 probability.

d. Omitted-variable regression version of the Hausman test for exogeneity of the predicted nonmetro residence variable, distributed as an *F*-statistic with a critical value of 3.00 at 0.05 probability.

Figure 1. People in Poverty, by Residence: 1966-2002



Source: U.S. Bureau of the Census, Current Population Survey, annual March Supplement.