DRAFT: PLEASE DO NOT CITE

Measuring the Effect of Rural Residence on Individual Employment Outcomes: Is Rural Residence Endogenous?

Monica Fisher [a+] and Ken Foster [b]

[a] University of Missouri, Truman School of Public Policy and Rural Poverty Research Center, Oregon State University, 213 Ballard Hall, Corvallis, OR 97331; (541) 737-1397 (phone); (541) 737-2563 (fax); fisherm@oregonstate.edu

[b] Department of Agricultural Economics, Purdue University, 403 West State Street, West Lafayette, IN 47907; (765) 494-1116 (voice), (765) 494-9176 (fax); kfoster@purdue.edu

[+] Corresponding author

Prepared for the American Agricultural Economics Association Annual Meeting,

August 1-4 2004, Denver, Colorado

Abstract

This paper examines the relationship between annual hours worked and rural residence for a sample of working-age (18-64 years) householders using 1993 Panel Study of Income Dynamics data. The basic question we address is whether and to what extent failure to account for rural residential choice biases the measured effect of rural residence on labor market outcomes. Results from a single equation model that assumes rural residence is exogenous finds no statistically significant relationship between annual hours worked and living in a rural area. By contrast, a simultaneous model that accounts for the possibility that rural residence is a choice indicates that rural people worked 307 hours more than urban people, all else being equal. A Smith-Blundell test for exogeneity of rural residence suggests that rural location should be treated as a choice variable. Study findings highlight the importance of testing, and if necessary, correcting for endogeneity of rural residence if we are to obtain accurate measures of the effect of living in a rural location on individual behavior and well-being.

Introduction

Recent research suggests that neighborhood characteristics such as the poverty rate and the extent of racial segregation affect individual employment outcomes (e.g. Cutler and Glaeser 1997; Weinberg, Reagan, and Yankow forthcoming). In this paper we ask if labor market outcomes vary across another community attribute: nonmetropolitan (nonmetro) versus metropolitan (metro) location. Why might there exist rural-urban differences in work behavior?[1] Social interaction models posit that community norms influence people's work aspirations (Akerlof 1997), and informal job contacts substantially improve an individual's chances of obtaining work either through improved information access or better hiring prospects (Granovetter 1995). Qualitative research indicates that economic self-sufficiency remains a core value in many rural areas (e.g. Pindus 2001; Wells 2002), and that rural people often have strong personal networks (e.g. Beggs, Haines, and Hurlbert 1996; Hofferth and Iceland 1998). If norms about work and the strength of social networks vary across metro and nonmetro areas, there may be rural-urban differences in work behavior.

The spatial mismatch hypothesis posits that people residing in communities spatially isolated from jobs face important barriers to work: long commutes and reduced access to employment information.[2] Rural labor markets tend to be less favorable (e.g. higher unemployment rates and lower job growth) (Gibbs 2002), and nonmetro residents often live at a great distance from job sites (Dewees 2000). For instance, in rural Kentucky, Mississippi, South Dakota, and Texas, welfare participants live 20-40 miles from the nearest

[1] The terms "nonmetro" and "rural" are used interchangeably in this paper to refer to counties outside of metropolitan areas.
[2] This hypothesis, focused primarily on urban areas, was initially used to explain unfavorable employment outcomes among central city African Americans (Kain 1968). See Ihlanfeldt and Sjoquist (1989) for a review.

place of employment (Harvey et al. 2002). Low-income central city residents also face long commutes to work, because entry-level jobs are increasingly located in the suburbs. Access to reliable transportation is therefore key to labor force participation for many central city and remote-rural residents. Public transit systems are better developed in metro than nonmetro areas, with nearly 40 percent of rural residents living in areas where there is no form of public transportation (Rucker 1994). However, vehicle ownership is higher in rural areas; 95 percent of nonmetro households own a vehicle compared with 80 percent of central city households (Nightingale 1997). In sum, theory and empirical evidence suggest possible rural-urban differences in labor market outcomes, although it is difficult to predict ex ante the direction of such a relationship.

There is some empirical research on the relationship between rural residence and labor market behavior. Davis, Connolly, and Weber (2003) find a negative but statistically weak association between living in a rural labor market area and employment probability among their sample of jobless poor Oregonians. Kilkenny and Huffman (2003) similarly find no statistically significant behavioral differences with respect to labor force participation between Midwestern householders residing in urban versus rural places. Phimister, Vera-Toscano, and Weersink (2002) show that observed rural-urban differences in female labor force participation in Canada are mainly attributable to observed socioeconomic characteristics. In sum, extant research indicates few rural-urban differences in labor force participation.

The present study contributes to existing work by relaxing the usual assumption that rural residence is exogenous to work behavior. We question the validity of this assumption, because people have some degree of freedom to choose where they live. If important factors influence both employment outcomes and peoples' residential mobility decisions, and these factors are not controlled for in the empirical model, estimates of a rural effect can be biased. A plausible scenario is that people who are less geographically mobile are more likely to choose rural residence (e.g. Feridhanusetyawan and Kilkenny 1996) and also have a lower probability of being employed. In such case, the effect on employment of living in a rural area could be understated, if the empirical model does not include a proxy for mobility. Another conceivable scenario is as follows. Studies show that work supports such as job training programs, child care facilities, and public transport are less available in rural areas (e.g. Colker and Dewees 2000; Fletcher 2002; Rucker 1994). It is plausible that people who rely on these work supports to find and retain employment self-select into urban areas. Failure to account for such neighborhood selection may result in biased estimates of the effect of rural residence on labor force participation. The suggestion that rural residence is to some extent "driven" by other factors leads to the implication that this variable is in all likelihood stochastic rather than deterministic as has been assumed in previous research.

In this paper we examine the association between rural residence and the annual labor hours of working-age (18-64 years) householders using Panel Study of Income Dynamics (PSID) data. While analysts have examined the relationship between rural residence and labor force participation, an unexplored hypothesis is that there exist rural-urban differences in annual hours worked. Our empirical approach is to estimate and compare results from two models: a single equation model that assumes living in a rural area is exogenous to annual hours worked and a simultaneous equation model that accounts for potential endogeneity of rural residence (Evans, Oates, and Schwab 1992 use a similar approach to study peer group effects). The basic question addressed is whether and to what extent failure to account for rural residential choice biases the measured rural effect. While the paper focuses on links between rural residence and labor market outcomes, findings are relevant to a general body of research that measures place effects on individual behavior and well-being.

Data Description

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 (see Brown, Duncan, and Stafford 1996 and Hill 1992 for detailed descriptions of the PSID). 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 our analyses because it is one of only two nationally-representative datasets that provides, for public use, information on rural/urban residence for certain years.[3]

We focus on a sample of 4,917 non-retired, working-age (18-64 years) individuals who were heads of household in 1993. We choose 1993 as the analysis year because it is the most recent year for which all of the required data for our analyses are available. In selecting an analysis year, two main factors are relevant. First, we require two consecutive years of data because some variables in a given year's data file concern the present year (e.g.

[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 respondent anonymity.

age and education) while others refer to the previous year (e.g. hours worked and nonlabor income). For this reason we cannot use the most recent data files (2001 and 1999), because in 1997 the PSID moved from annual to biennial interviewing. A second factor is that a variable for rural/urban residence is not available for all years; such a variable exists for 1968-1993, 1999, and 2001.

Several criteria were used to arrive at our final sample of 4,917 individuals. First, we focus on non-retired, working-age individuals given our interest in labor market outcomes. Second, our sample consists only of household heads because the PSID data files provide more extensive information for these individuals compared with other family members. A third sampling criterion is that we consider only those householders 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. Fourth, we include only those respondents that resided in the United States during the survey year. Finally, we drop all observations with incomplete data for the variables used in the analyses. For example, we exclude about 200 individuals who reported labor hours for 1993 but did not report a wage for the same year; this is to avoid an explicit errors in variable problem. Table 1 provides definitions and descriptive statistics for key variables used in this study.

Empirical Analysis

Single Equation Model

We begin with an empirical model in which rural residence is assumed to be an exogenous variable, paralleling the common practice of existing work. The model is a single equation Tobit model of the form

(1) $$h = \alpha_0 + \alpha_1 x + \alpha_2 \hat{w} + \alpha_3 u + \alpha_4 r + \varepsilon_1,$$

where dependent variable h is annual hours worked.[4] Explanatory variables are individual and household characteristics $\boldsymbol{x}$ (including the age of the youngest child and number of children in the household as well as the householder's age, race, gender, marital status, education, work experience, and disability status), the householder's predicted wage $\hat{w}$, the county unemployment rate u, and a binary variable r indicating whether the county of residence is rural. Our selection of explanatory variables is consistent with other empirical models in the rural poverty literature (e.g. Brown and Hirschl 1995; Kilkenny and Huffman 2003; Mills and Hazarika 2003).

Since observed wages are endogenous to the choice to work, we control for possible sample-selection bias by jointly estimating a participation and shadow wage equation using maximum likelihood (ML) methods. ML is preferred over the Heckman two-step approach because it is consistent and efficient, whereas the two-step method is not fully efficient (Nawata 1994). Estimated parameters from the shadow wage equation are then used to predict a wage for each observation. Using predicted wages in labor hours equation (1) for every observation, rather than a combination of imputed and observed wages, helps purge the variable of possible measurement error or endogeneity. The participation and wage equations are

(2) $$y = \gamma_0 + \gamma_1 x + \gamma_2 u + \gamma_3 r + \varepsilon_2$$

(3) $$w = \delta_0 + \delta_1 x^* + \delta_2 u + \delta_3 r + \delta_4 s + \varepsilon_3,$$

[4] We employ a Tobit model because 495 householders in the sample did not work. The Tobit technique accounts for this truncation in the dependent variable.

where $\boldsymbol{x}$, u, and r are defined as above, y is the probability that the householder works, w is the natural log of observed wages, $\boldsymbol{x}^*$ is a subset of the variables in $\boldsymbol{x}$, and $\boldsymbol{s}$ is a set of state-of-residence binary variables.

To identify the shadow wage equation, we include in the participation equation a set of variables which we hypothesize affect participation in the labor force by altering the householder's shadow wage, but which do not directly affect the observed market wage. Nonlabor income, the number of children, and the age of the youngest child should affect an individual's reservation wage and therefore her/his decision to work, without altering the offered wage. Following Kilkenny and Huffman (2003), we identify the labor hours equation with use of a set of state binary variables. We expect state-of-residence's effect on labor supply operates through its effect on the shadow wage rather than directly. Results for participation and shadow wage are available upon request.

Table 2 presents Tobit results for the labor hours equation. The calculated Wald statistic of 1509.76 is significant at the 95% confidence level, providing support for the hypothesis of joint significance of the explanatory variables. In addition, most of the point estimates are individually different from zero at the 95% confidence level and have the expected signs. Findings indicate that education and work experience of the householder have a positive influence on annual hours worked. We find that hours worked increases with age until the householder reaches 25 years, and then decreases thereafter. Results show that householders work fewer hours if they are single, female, African American, disabled, have young children, and live in a county with a high unemployment rate.

Turning to the result of primary interest, our findings do not provide statistical support for the hypothesis that residents of rural places work more than urban inhabitants, controlling for other key factors. Such a finding is consistent with existing work that finds few rural-urban differences in labor force participation (e.g. Davis, Connolly, and Weber 2003; Kilkenny and Huffman 2003; Phimister, Vera-Toscano, and Weersink 2002). A maintained hypothesis is that current estimates of a rural effect on employment are biased because residence in a rural 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 neighborhood choice; the second, fixed effects strategies, involves introducing controls for individual heterogeneity (Weinberg, Reagan, and Yankow forthcoming). Each approach has advantages and drawbacks, but for our purposes instrumental variables is more appropriate.[5] Below we turn to a simultaneous equation model of annual labor hours, asking whether and to what extent failure to account for rural residential choice biases the measured rural effect

Simultaneous Equation Model

In this section we treat rural residence as an endogenous variable in the context of a simultaneous equation model. We assume the probability an individual resides in a rural location r is a function of individual-level variables that determine labor supply $\boldsymbol{x}$, a set of identifying instruments $\boldsymbol{z}$ assumed to affect residential choice but not labor market decisions

[5] One fixed effects approach involves the use of data from multiple siblings of households to difference out fixed family effects (e.g. Aaronson 1998). 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 forthcoming). This approach can only account for neighborhood selection related to time-invariant individual factors (Dietz 2002).

(binary variables indicating whether the householder grew up in a rural area, a small town, or a city), and a random error term ε_4. The rural residential choice model is

(4) $$r = \beta_0 + \beta_1 x + \beta_2 z + \varepsilon_4,$$

and the simultaneous model consists of equations (1) and (4).

Table 3 presents results from the first-stage Probit regression. Results show that a householder's choice of living in a rural area is positively correlated with being married and having children. A plausible explanation is that married people with children prefer living in rural areas, because nonmetro locations have relatively low housing costs and provide a generally favorable environment for child rearing (e.g. lower crime and population density) (Feridhanusetyawan and Kilkenny 1996). Findings in table 3 also suggest that householders who are African American and more educated are less likely to live in rural areas. The latter finding is consistent with a hypothesis that more-educated individuals prefer living in metro areas where, compared to nonmetro locations, there are higher proportions of jobs that require advanced education (Feridhanusetyawan and Kilkenny 1996). We find that rural residence is positively correlated with the county unemployment rate.

The instruments used to identify the annual hours worked equation are binary variables indicating whether the household head grew up in a rural area, a small town, or a city (the reference category is a combination of places). It is important to examine the validity of the instruments, because the two-stage instrumental variables strategy only corrects for endogeneity to the extent that good instruments are employed (Dietz 2002). As shown in table 3, two of the instruments are individually statistically significant in the rural residence equation. The χ^2 statistic for the joint significance of the instruments is 561.05; at 0.05 probability, the instruments are jointly significant. Below we report results from an omitted-variable regression version of the Hausman test (Spencer and Berk 1981) which provides additional support for the validity of the instruments.

Findings for the simultaneous model of hours worked are provided in table 4.[6] The calculated Wald statistic of 1524.98 is significant at the 95% confidence level, indicating that the explanatory variables are jointly significant. Focusing first on all explanatory variables other than the rural residence binary variable, we see that coefficient estimates are similar for the single equation and simultaneous equation models (compare table 2 and table 4). The sign of the coefficient estimate for each of the regressors is the same across equations and differences in magnitude are not large. In addition, the set of statistically significant explanatory variables is the same for single equation and simultaneous models.

Turning to the parameter estimates for the rural residence variable, findings are quite striking. Whereas the rural effect is statistically weak in the single equation model (table 2), in the simultaneous model (table 4) a rural effect appears. Accounting for the possibility that living in a rural location is a choice, we find a positive statistically significant relationship between annual hours worked and rural residence. All else being equal, rural residents worked 307 hours more than their urban counterparts in 1993. Given the large difference in the measured effect of rural residence between single equation and simultaneous equation models, it seems imperative to examine the validity of the chosen instruments. Earlier we reported results from the Probit model of rural residential choice which show that the instruments are correlated with rural residence, one measure of instrument validity. Here we use an omitted-variable regression version of the Hausman test

---

[6] The predicted wage in table 4 comes from estimation of a Heckman model of labor force participation and shadow wage, estimated with predicted rather than observed rural residence as a right-hand-side variable. Results are available upon request.

(Spencer and Berk 1981) to investigate another aspect of the validity of the instruments: identifying instruments should be uncorrelated with the error term of the annual hours worked equation. The calculated $\chi^2(3)$ statistic of 5.89 is less than the ($p = 0.05$) critical value of 7.89, suggesting that the instruments are not jointly significant in the annual hours equation. This provides additional support for the validity of the instruments.

As a final measure of the soundness of our results, we test whether rural residence is exogenous to annual hours worked, using the approach proposed by Smith and Blundell (1986) for simultaneous limited dependent variable models. This test is essentially one for exclusion of residuals from an auxiliary regression of rural residence on all explanatory variables and instruments. As above, the instruments are binary variables indicating whether the householder grew up in a rural area, small town, or city. We employ a Stata program "probexog-tobexog" to implement the exogeneity test.[7] The Smith-Blundell test yields a $\chi^2(1)$ statistic of 4.45 ($p = 0.03$). Thus, we reject the null hypothesis of exogeneity of rural residence in the annual hours worked equation. Findings of the Smith-Blundell test suggest that endogeneity of the rural residence variable should be accounted for in the econometric model of labor hours if we are to obtain an unbiased estimate of the effect of rural residence on labor supply.

Conclusion

We have examined the relationship between annual hours worked and rural residence for a sample of working-age (18-64 years) householders using 1993 data from the PSID.

[7] For the Smith-Blundell test, we do not include predicted wages among the set of regressors.

Two estimation strategies were employed. One, we estimated a single equation Tobit model of annual hours worked as a function of shadow wages, characteristics of the householder, county unemployment rate, and rural residence. This model assumes that rural residence is an exogenous variable, paralleling the common practice of previous research. Our second approach involved estimating a two-stage model that relaxes the assumption of rural residence exogeneity. In the first stage, rural residence was modeled as a function of householder characteristics, county unemployment rate, and a set of instruments. In the second stage, predicted rural residence from the first-stage regression replaced observed rural residence in our model of annual hours worked.

Results of the study provide both methodological and empirical insights. From a methodological standpoint, we respond to Jensen and Weber's (2004) 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 rural 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. Our study underscores the importance of testing and, if necessary, correcting for endogeneity in the econometric measurement of the effects of rural residence on individual outcomes. We find a statistically insignificant rural coefficient in the single equation model, but the simultaneous model shows a positive correlation between annual hours worked and rural residence. Tests for the validity of instruments used to identify hours worked provide some support for our choice of instruments. A Smith-Blundell test for exogeneity of rural residence indicates that living in a rural area is a choice. In tandem, our empirical findings seem to suggest that failure to account for residential endogeneity in empirical modeling leads to an under-estimation of the effect of rural residence on annual hours worked.

From an empirical perspective we present new findings on the relationship between employment and rural residence. Existing research suggests few rural-urban differences in labor force participation (e.g. Davis, Connolly, and Weber 2003; Kilkenny and Huffman 2003; Phimister, Vera-Toscano, and Weersink 2002). Our study results indicate that rural people work more hours during the year than do urban residents, all else being equal. Such a finding may indicate that informal job contacts are more plentiful and the work ethic stronger in nonmetro areas. Or, study findings may suggest that the spatial mismatch between where people live and where jobs are located is more problematic in metro areas. There is a need for further study of the specific factors that give rise to observed rural-urban differences in labor supply. An improved understanding of these factors will enable improved targeting of rural- and urban-specific employment assistance programs.

References

Aaronson, D. “Using Sibling Data to Estimate the Impact of Neighborhoods on Children’s Educational Outcomes.” *Journal of Human Resources* 33 (4): 915-946.

Akerlof, G. 1997. “Social Distance and Social Decisions.” *Econometrica* 65 (5): 1005-1027.

Brown, C., G.J. Duncan, and F.P. Stafford, 1996. “Data Watch: The Panel Study of Income Dynamics,” *The Journal of Economic Perspectives* 10 (2), 155-168.

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

Cutler, D.M. and E.L. Glaeser. 1997. “Are Ghettos Good or Bad?” *Quarterly Journal of Economics* 112 (3): 827-872.

Davis, E.E., L.S. Connolly, B.A. Weber. 2003. “Local Labor Market Conditions and the Jobless Poor: How Much Does Local Job Growth Help in Rural Areas?” *Journal of Agricultural and Resource Economics* 28(3): 503-518.

Dewees, S. 2000. “Transportation in Rural Communities: Strategies for Serving Welfare Participants and Low-Income Individuals.” Rural Welfare Issue Brief, Administration for Children and Families, U.S. Dept. of Health and Human Services.

Dietz, R.D. 2002. “The Estimation of Neighborhood Effects in the Social Sciences: An Interdisciplinary Approach.” *Social Science Research* 31: 539-575.

Evans, W.N., W.E. Oates, and R.M. Schwab. 1992. “Measuring Peer Group Effects: A Study of Teenage Behavior.” *Journal of Political Economy* 100(5): 966-991.

Feridhanusetyawan, T. and M. Kilkenny. 1996. “Rural/Urban Residential Location Choice.” Center for Agricultural and Rural Development Working Paper 96-WP 157. Ames: Iowa State University.

Gibbs, R.M. 2002. “Rural Labor Markets in an Era of Welfare Reform.” In B.A. Weber, G.J. Duncan, and L.A. Whitener (eds.) *Rural Dimensions of Welfare Reform*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Granovetter, M. *Getting a Job*, second edition. Chicago, IL: University of Chicago Press, 1995.

Harvey, M., G.F. Summers, K. Pickering, and P. Richards. 2002. “The Short-Term Impacts of Welfare Reform in Persistently Poor Rural Areas.” In B.A. Weber, G.J. Duncan, and L.A. Whitener (eds.) *Rural Dimensions of Welfare Reform*. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

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

Hofferth, S.L. and J. Iceland. 1998. "Social Capital in Rural and Urban Communities." *Rural Sociology* 63 (4): 574-598.

Ihlanfeldt, K.R. and D.L Sjoquist. 1989. "The Impact of Decentralization on the Economic Welfare of Central City Blacks." *Journal of Urban Economics* 26: 110-130.

Jensen, L. and B. Weber. 2004. "Poverty and Place: A Critical Review of Rural Poverty Literature." Paper prepared for the Importance of Place in Poverty Research and Policy conference, March 3-4 2004, Washington, D.C.

Kilkenny, Maureen, and Sonya Kostova Huffman. 2003. "Rural/Urban Welfare Program and Labor Force Participation." *American Journal of Agricultural Economics* 85(4): 914-27.

Mills, B F. and G. Hazarika. 2003. "Do Single Mothers Face Greater Constraints to Workforce Participation in Non-Metropolitan Areas?" *American Journal of Agricultural Economics* 85 (1): 143-61.

Nawata, K. 1994. "Estimation of Sample Selection Bias Models by the Maximum Likelihood Estimator and Heckman's Two Step Estimator." *Economic Letters* 45 ( ): 33 - 40.

Phimister, E., E. Vera-Toscano, and A. Weersink. 2002. "Female Participation and Labor Market Attachment in Rural Canada." *American Journal of Agricultural Economics* 84 (1): 210-221.

Pindus, N.M. 2001. "Implementing Welfare Reform in Rural Communities." Washington, D.C.: The Urban Institute. http://www.urban.org/pdfs/rural-welfarereform.pdf.

Rucker, G. 1994. "Status Report on Public Transportation in Rural America." Rural Transit Assistance Program, Federal Transit Administration. http://www.fta.dot.gov/library/program/rurlstat/rurlstat.html.

Smith, R.J. and R.W. Blundell. 1986. "An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply." *Econometrica* 54(3): 679-686.

Wells, B. 2002. "Women's Voices: Explaining Poverty and Plenty in a Rural Community." *Rural Sociology* 67 (2): 234-254.

Table 1. Descriptive Statistics of Explanatory Variables, Householders 1993

	Mean or Frequency [a]	Standard Deviation [a]
Endogenous variables		
employed	0.93	
wage (1993 US$) [b]	17.32	19.51
rural residence	0.24	
Exogenous variables		
age (years)	39.03	10.82
female	0.28	
African American	0.15	
married [c]	0.52	
education (years)	13.20	2.38
work experience (years)	13.00	9.10
disabled	0.03	
nonlabor income/10,000 (1993 US$) [d]		
number of children	0.88	1.16
dummy for child < 6 years	0.20	
county unemployment rate	7.19	2.30
where grew up (combination places excluded)		
rural area	0.16	
small town	0.46	
city	0.36	
Number of observations [e]		3,387

a. Means and standard deviations are weighted by the PSID core sample individual weight.
b. Summary statistics are only for those individuals who were employed in 1993.
c. Head is married or has a cohabitator with whom he/she has lived for at least one year.
d. Nonlabor income is the sum of income from dividends, trust funds, interest, and rental property.
e. Number of observations is less than the sample size of 4,917 because the individual weight variable is not available for all observations.

Table 2. Tobit Results for Annual Hours Worked, Single Equation Model [a]

	Coefficient	Robust Stand. Err.	Marginal Effect [b]
Constant	* 514.75	203.10	
Predicted wage	30.21	17.93	27.72
Predicted wage squared	* -1.31	0.42	-1.20
Age	* 35.82	10.24	32.87
Age squared	* -0.71	0.12	-0.66
Female	* -318.9	44.61	-287.65
African American	* -223.35	35.99	-203.70
Married	* 256.96	45.88	234.64
Education	* 66.5	10.15	61.01
Work experience	* 28.36	3.36	26.02
Disabled	* -2319.17	125.92	-1380.47
Number of children	* -28.52	13.70	-26.17
Child < 6 years	* -106.39	32.65	-97.08
County unemployment rate	* -24.33	5.54	-22.32
Rural residence	57.53	37.81	52.94
Number of observations			4,917
Wald statistic (14) [c]			1509.76

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

b. In the Tobit framework, a change in the independent variable is decomposed into two separate effects: the effect on the conditional mean of the dependent variable in the positive portion of the distribution, and the impact on the probability that the observation falls in that part of the distribution. See Greene (2000).

c. Wald test for joint significance of the explanatory variables, distributed as a χ^2 with a critical value of 23.69 for 14 degrees of freedom at 0.05 probability.

Table 3. Probit Results for Rural Residential Choice [a]

	Coefficient	Robust Stand. Err.	Marginal Effect [b]
Constant	-0.3301	0.3395	
Age	-0.0260	0.0155	-0.0069
Age squared	0.0001	0.0002	0.00004
Female	-0.0520	0.0754	-0.0136
African American	* -0.3069	0.0514	-0.0784
Married	* 0.1517	0.0705	0.0397
Education	* -0.0555	0.0098	-0.0147
Work experience	0.0017	0.0049	0.0004
Disabled	0.0801	0.1190	0.0219
Number of children	* 0.0472	0.0220	0.0125
Child < 6 years	-0.0610	0.0598	-0.0159
County unemployment rate	* 0.1282	0.0101	0.0339
Where grew up (combination places)			
Rural area or small town	* 0.8257	0.1573	0.2646
Small town	0.2600	0.1528	0.0702
City	* -0.7299	0.1576	-0.1807
Number of observations			4,917
Wald statistic (13) [c]			777.90
Pseudo R-squared			0.193

a. * implies significance at the 0.05 probability level or better. Huber/White robust standard errors reported.
b. For binary variables, marginal effects are interpreted as the change in the probability of rural residence associated with a discrete change in the explanatory variable.
c. Wald test for joint significance of the explanatory variables, distributed as a χ^2 with a critical value of 22.36 for 13 degrees of freedom at 0.05 probability.

Table 4. Tobit Results for Annual Hours Worked, Simultaneous Equation Model [a]

	Coefficient	Robust Stand. Err.	Marginal Effect [b]
Constant	* 403.54	199.84	
Predicted wage [c]	29.44	16.56	27.03
Predicted wage squared [c]	* -1.26	0.40	-1.16
Age	* 37.78	9.99	34.68
Age squared	* -0.73	0.12	-0.67
Female	* -315.90	44.52	-285.22
African American	* -188.27	36.05	-172.01
Married	* 238.26	45.17	217.81
Education	* 71.86	9.38	65.98
Work experience	* 28.13	3.35	25.83
Disabled	* -2330.47	123.12	-1384.71
Number of children	* -32.12	13.67	-29.49
Child < 6 years	* -96.57	32.63	-88.22
County unemployment rate	* -33.69	6.02	-30.93
Predicted rural residence	* 333.90	87.96	306.55
Number of observations			4,917
Wald statistic (14) [d]			1532.00

a. * implies significance at the 0.05 probability level or better. Huber/White robust standard errors reported.
b. In the Tobit framework, a change in the independent variable is decomposed into two separate effects: the effect on the conditional mean of the dependent variable in the positive portion of the distribution, and the impact on the probability that the observation falls in that part of the distribution. See Greene (2000).
c. The predicted wage comes from estimation of a Heckman model of labor force participation and shadow wage, estimated with predicted rather than observed rural residence as a right-hand-side variable.
d. Wald test for joint significance of the explanatory variables, distributed as a χ^2 with a critical value of 23.69 for 14 degrees of freedom at 0.05 probability.