SCHOOL QUALITY AND RURAL IN-MIGRATION: CAN IMPROVING THE QUALITY OF RURAL SCHOOLS ATTRACT NEW RESIDENTS?

Alexander W. Marré[1]
amarre@ers.usda.gov

Anil Rupasingha
anil.rupasingha@ers.usda.gov

Selected Paper prepared for presentation at the 2017 Agricultural & Applied Economics Association Annual Meeting, Chicago, Illinois, July 30-August 1

The views expressed here are the authors' alone and should not be attributed to the U.S. Department of Agriculture or the Economic Research Service.

[1] The authors are Research Economists at the U.S. Department of Agriculture's Economic Research Service, 1400 Independence Ave., SW, Mail Stop 1800, Washington, DC 20250-0002.

1. INTRODUCTION

Population loss is a critical issue for many nonmetropolitan (nonmetro) areas. As natural change – the number of births minus deaths – declines in nonmetro areas, attracting more new residents than those who leave is becoming more important in stemming or reversing population loss. Yet net migration to nonmetro areas has been negative since 2010 and nonmetro areas lost population overall between 2010 and 2015. Population loss has consequences for the prosperity of nonmetro areas. From an economic standpoint, it reduces the customer and tax base for businesses and local government. For local governments, the loss implies a higher price to providing public services to the remaining population. Kilkenny (2010) notes that as population declines in rural communities, there is a potential for them to fall below the minimum efficient scale needed to survive. For example, one study of rural communities in Oregon found that the smallest rural communities, with less than 1,250 people, were most at risk of losing population (Chen, Etuk & Weber, 2013).

The purpose of this study is to examine whether improving the quality of local schools in nonmetropolitan counties could attract in-migrants. Our hypothesis is that improving the quality of public schools is one public good that could make a county more attractive to in-migration, all else equal. Movers with a preference for school quality would seek it out and "vote with their feet" (Tiebout, 1956). This work is situated in two sets of literature. First, there is an extensive literature estimating the willingness to pay for indicators of school quality, with most showing a positive capitalization of school quality into housing prices (Black, 1999; Nguyen-Hoang & Yinger, 2011). This literature indicates that the desirability of good local schools may be reflected in home prices.

Second, there is a body of literature in the rural development field on factors associated with the movement of people and firms to nonmetrofrom metropolitan (metro) counties. There was interest in the 1990s, in particular, to explain a change in migration patterns towards rather than away from nonmetro counties. The deconcentration hypothesis considered the role of retiree migration, the availability of telecommuting in metro-adjacent nonmetro counties, and amenity-based migration as factors (Frey and Johnson 1996; Nelson and Nelson 2011; Cromartie 1998; Deller, S.C., T.-H. Tsai, D.W. Marcouiller & D.B.K. English, 2001; Partridge, Ali, and Olfert 2010a; Nelson, Oberg, and Nelson 2010). Most recently, Rupasingha, Liu and Partridge (2015) estimate the determinants of migration from metropolitan to nonmetropolitan counties over two time periods to test the deconcentration hypothesis and an economic restructuring hypotheses – focusing on economic factors such as unemployment and wages – on metro-nonmetro population flows. Their findings show a slight diminution of amenity-driven migration and a slight increase the role of economic factors influencing nonmetro-metro migration patterns.

We follow Rupasingha, Liu and Partridge (2015), using the same model and variables, except for the inclusion of school quality variables. Our empirical model uses county-to-county migration flows from the U.S. Census Bureau between 2005 and 2009 and data on dropout rates from the 2000 Census and test scores form the Global Report Card. A Poisson regression model is estimated with fixed effects controlling for characteristics in the origin counties. Independent variables control for observed characteristics in the destination counties, including measures of school quality. To the best of our knowledge, this is the first study to examine the role that local school quality plays in attracting in-migrants to nonmetro counties from metro counties.

2. MODEL AND ECONOMETRIC APPROACH

Our conceptual model is based on a utility maximization approach and largely follows the setup presented in Rupasingha, Liu and Partridge (2015). Prospective migrants maximize utility across a set of alternative locations, $j = 1,\ldots,J$, compared with the utility derived at their current location, i. Utility is defined as a function of potential earnings in each location, w_i or w_j, the costs associated with moving from i to j, c_{ij}, and amenities in each location, a_i or a_j. The quality of local schools can be thought of as a location-specific amenity. A person evaluates d_{ij}, the difference in utility between moving to location j and staying in the current location i, as shown in equation (1), and moves if $d_{ij} > 0$ for some j.

(1) $d_{ij} = U\left(w_j - c_{ij}, a_j\right) - U(w_i, a_i)$

Following the random utility model by McFadden (1974), the choice between any two locations, i and j, can be expressed as a vector of characteristics in both places (X_{ij}).

(2) $d_{ij} = \beta' X_{ij} + \varepsilon_{ij}$.

The individual chooses to migrate to the prospective location j that maximizes their utility. Therefore, the probability of a move to location j is defined as:

(3) $P_{ij} = P\left(U_{ij} > U_{ik}\right)$ for all $k \neq j$.

As shown by McFadden (1974), if the ε_{ij} are independent and identically distributed following an extreme value type-I distribution, then (3) can be written as:

(4) $$P_{ij} = \frac{\exp(\beta' X_{ij})}{\sum_j \exp(\beta' X_{ij})}$$

Equation (4) is typically used to estimate a conditional logit model that is used in firm location studies. Guimarães et al. (2004) show that a Poisson regression model can be used instead of a conditional logit model if certain conditions are met. The benefits of using a Poisson regression model include that it makes estimating very large alternative location choices computationally-feasible, a weakness of the conditional logit model. The Poisson model also allows one to avoid a potential violation of the independence of irrelevant alternatives assumption underlying the conditional logit model framework. Yet another advantage of the model used for this analysis is that it mitigates endogeneity concerns that could arise using other methods. In addition to lagging the school quality variables to the extent that we are able – the dropout rate is at least five years before in-migration is observed – the migration flows from any one particular county to a destination county is unlikely to be large enough to influence other destination county characteristics. Finally, the model used here controls for both “push” and “pull” factors associated with migration flows. “Push” factors from the origin county are controlled by county-level fixed effects and “pull” factors to the destination county are controlled for by observed county-level characteristics (including school quality) and state-level fixed effects.

3. DATA

Table 1 shows the variable names used in this study, a brief description, and the mean and standard deviation. Except for the school quality variables and the industry mix variable, all of the variables are identical to those used in Rupasingha, Liu and Partridge (2015). The dependent variable is the number of people moving from a particular metro county to a particular nonmetro county between 2005 and 2009 and is taken from the Census Bureau’s 2005-2009 American

Community Survey. A more detailed description of the other non-school quality variables may be found in Rupasingha, Liu and Partridge (2015). The industry mix variable measuring exogenous local demand conditions was calculated using data from the National Establishment Time-Series Database at the 1-digit Standard Industrial Classification level following Bartik (1991). It is calculated as the sum of the products of initial county-level industry employment shares in 2000 by the 2000-2007 national employment growth rates for each industry. As in Rupasingha, Liu and Partridge (2015), the coefficient of variation is used for unemployment rates given the possibility that migration and unemployment rates could be simultaneously determined (Etzo 2010). State fixed effects are included to capture – in part – unobserved state education policies.

One of the school quality variables used is the dropout rate (*dropoutrt*), measured as the share of the civilian population between the ages of 16 and 19 that are not enrolled in high school, and is taken from the 2000 Population Census. The advantage of using this measure of high school enrollment versus district or school-level graduation rates from the Department of Education's Common Core of Data is that it is available at the county-level rather than the district-level. The other school quality variables are based on student test scores in reading (*wpctreading*) and math (*wpctmath*) and are available at the school-district-level. Data on test scores are from the George W. Bush Institute Global Report Card, which compiles district-level data on state exams. The weighted-average[2] percentage of students deemed "proficient" in reading and math in grades 3 through 8 is aggregated to a single proficiency rate for each district. These proficiency rates are then standardized to the state-level, which are in turn standardized to

[2] Weighted by students in grade levels 3-8.

the national-level using the National Assessment of Educational Progress (NAEP). [3] The test scores from the Global Report Card comparing districts has been used in other research. For example, Chetty, Hendren, Kline and Saez (2014) found that school quality was correlated with high economic mobility.

Test scores at the school-district-level are translated into counties by using the county identifier provided in the U.S. Department of Education's Common Core of Data. It indicates the county in which each school district's headquarters is located. For the cases where more than one school district is associated with a county, we calculate an enrollment-weighted average of the reading and math test scores. Since school districts cross county boundaries, this approach can introduce measurement error. The enrollment-weighted average test scores for math and reading may not capture all of the schools in a particular county. Future research will attempt to address this issue by geocoding schools to counties, thus creating a more precise measure of the relative enrollment size of districts in each county.

4. RESULTS

For the purposes of this paper, we focus our discussion solely on the preliminary results for the school quality variables. Results for other variables in the model are similar to those presented in Rupasingha, Liu and Partridge (2015). Results from estimation of the fixed effect Poisson

[3] The technical appendix to the Global Report Card provides an example: "For example, in 2007 the average student in the Scarsdale School District in Westchester County, New York scored nearly one standard deviation above the mean for New York on the state's math exam. The average student in New York scored six hundredths of a standard deviation above the national average on the NAEP exam given in the same year, and the average student in the United States scored about as far in the negative direction (-0.055) from the international mean on PISA. Our final Global Report Card score for Scarsdale in 2007 is equal to the sum of the district, state, and national estimates (1 + 0.06 + -0.055 = 1.005). Since the final Global Report Card score is expressed in standard deviation units, it can easily be converted to a percentile for easy interpretation. In our example Scarsdale would rank at the seventy seventh percentile internationally in math" (Greene and McGee, 2011).

models are presented in Table 2. Two separate regressions were estimated because the reading and math test score variables are highly correlated (correlation of 0.86). The first two columns (Model 1) report estimated coefficients, robust standard errors and marginal effects for the model with reading test scores and the last two columns (Model 2) report these for the model with math test scores. Factor changes are used to calculate the marginal effect of each independent variable, k, as $\left(e^{\beta_k} - 1\right)$ multiplied by 100 (Long 1997, p. 223-6). These are interpreted as the percentage change in the expected count for a unit change in the explanatory variable, all else constant.

All of the school quality variables in both models are statistically significant and have expected signs. The coefficient of the dropout rate is negative, suggesting that a nonmetro county with a higher high school dropout rate will attract fewer people from metro counties. Specifically, in the model with reading test scores, a one percentage point increase in the share of high school dropouts is estimated to reduce in-migration from metro counties by 1.3 percent. The estimated marginal effect is a reduction of 1.5 percent in the model with math test scores. Test scores have positive impacts in both models, indicating that nonmetro counties with higher test scores attract more migrants from metro counties, holding all other observed characteristics constant. A one percentage point increase in the percentage of students proficient in reading is estimated to increase in-migration by 1.3 percent from metro counties and a one percentage point increase in math scores is estimated to increase in-migration by 0.9 percent from metro counties.

5. CONCLUSION

To the best of our knowledge, this is the first paper to examine the effect of school quality on migration patterns from metro to nonmetro counties at the national scale. Anecdotal evidence from interviews of return migrants to rural communities suggested that the quality of local schools is one factor drawing some in-migrants to particular rural areas (Cromartie, von Reichert and Arthun, 2015). Our results suggest that for the 2005-2009 time period, the quality of schools, as measured by the share of high school dropouts and the share of students proficient in math and reading, had a positive "pull" effect on migration from urban to rural areas. Our use of the Poisson regression technique, following Rupasingha, Liu and Partridge (2015) is relatively new in the population migration literature and helps to ameliorate potential concerns of endogeneity.

Results are preliminary and the research presented here needs refinement in a few key areas:

- Addressing measurement error in test scores – Test score data is at the district-level and many school districts cross county boundaries. A more precise measurement of test scores at the county-level will need to weight scores by the shares of overlapping school districts that are in a particular county.
- Refining in-migration flows by age – School quality is likely to be more important to migrants with school-age children. Therefore, the effects estimated here may be smaller than for particular age groups.
- Expanding dependent variable – The analysis should be extended to the effects of school quality on drawing in residents from other nonmetro counties as well as any "push" effects that poor quality schools may have on residents.
- Comparison across time – As migration declines in the United States, it would be helpful to compare the effects of school quality over time on migration decisions. Are school quality and natural amenities becoming less important over time?

REFERENCES

Barkley, D.L., M.S. Henry and S. Bao. 1998. "The role of local school quality in rural employment and population growth." *Review of Regional Studies* 28(1):81-102.

Bartik, T.J. 1991. *Who Benefits from State and Local Economic Development Policies?* Upjohn Institute, Kalamazoo, MI.

Betts, J.R. 1995. "Does school quality matter? Evidence from the National Longitudinal Survey of Youth." *Review of Economics and Statistics* 77(2):231-250.

Black, S.E. 1999. "Do better schools matter? Parental valuation of elementary education." *Quarterly Journal of Economics* 114(2):577-599.

Chen, Y., L. Etuk & B. Weber. 2013. "Are small communities at risk of population loss?" *Annals of Regional Science* 51:343-353.

Cromartie, J.B. 1998. "Net migration in the Great Plains Increasingly Linked to Natural Amenities and Suburbanization." *Rural Development Perspectives* 13: 27–34.

Cromartie, J., C. von Reichert, and R. Arthun. 2015. Factors Affecting Former Residents' Returning to Rural Communities, ERR-185, U.S. Department of Agriculture, Economic Research Service.

Deller, S.C., T.-H. Tsai, D.W. Marcouiller & D.B.K. English. 2001. "The role of amenities and quality of life in rural economic growth." *American Journal of Agricultural Economics* 83(2):352-365.

Etzo, I. 2010. "The determinants of the recent interregional migration flows in Italy: A panel data analysis." *Journal of Regional Science* 51(5):948-66.

Figlio, D.N. and M.E. Lucas. "What's in a grade? School report cards and the housing market." *American Economic Review* 94(3):591-604.

Frey, W.H. and K.M. Johnson. 1996. "Concentrated Immigration, Restructuring and the Selective Deconcentration of the United States Population." Population Studies Center Research Report Series No. 96–371, The Population Studies Center at the University of Michigan.

Greene, J.P. and J.B. McGee. 2011. "Technical Appendix." George W. Bush Institute Global Report Card. Available online at: http://www.globalreportcard.org/docs/AboutTheIndex/Global-Report-Card-Technical-Appendix-9-28-11.pdf

Kilkenny, M. 2010. "Urban/Regional economics and rural development.' *Journal of Regional Science* 50(1):449-470.

McFadden, D.L. 1974. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, ed. Zrembka P., 105-42. New York: Academic Press.

Nelson, L. and P.B. Nelson. 2011. "The Global Rural: Gentrification and Linked Migration in the Rural USA." *Progress in Human Geography* 35(4): 441–59.

Nelson, P.B., A. Oberg and L. Nelson. 2010. "Rural Gentrification and Linked Migration in the United States." *Journal of Rural Studies* 26(4): 343–52.

Partridge, M.D., D.S. Rickman, M.R. Olfert and K. Ali. 2012. "Dwindling US Internal Migration: Evidence of Spatial Equilibrium or Structural Shifts in Local Labor Markets?" *Regional Science and Urban Economics* 42:375–88.

Rupasingha, A., Y. Liu & M. Partridge. 2015. "Rural bound: Determinants of metro to non-metro migration in the United States." *American Journal of Agricultural Economics* 97(3):680-700.

Tiebout, C.M. 1956. "A pure theory of local expenditures." *Journal of Political Economy* 64:416-424.

Table 1. Variable Descriptions and Descriptive Statistics

Variable	*Description*	*Mean*	*Std. Dev.*
depend2	Migration from nonmetro to metro counties	0.905	13.82
wage	Wage and salary per job, in thousands of dollars	23.120	3.80
popden00	Population per square mile, in thousands	0.045	0.11
distance	Actual distance between an origin county and destination county on average, in hundreds of miles	9.103	5.96
distance_sq	Distance squared	118.408	181.64
indusmix1	Industry mix national employment growth	107.651	3.050
amnscale	Natural amenities index	-0.051	2.25
elder	Percentage of population over 64 years old	16.066	3.907
metroadj	Nonmetro counties that are adjacent to a metro area (0,1)	0.518	0.500
pctax02	Per capita local taxes, in thousands	0.962	0.813
pcgexp02	Per capita local government expenditure, in thousands	2.960	1.222
mhv	Median housing value, in thousands	70.552	36.165
cvurate00_09	Coefficient of variation of unemployment rate	0.242	0.090
dropoutrt	Percent of the civilian population ages 16-19 not enrolled in school and not a high school graduate	9.661	5.658
wpctreading	Percent proficient in reading, 2004-2009	50.937	10.290
wpctmath	Percent proficient in math, 2004-2009	50.688	12.043

Table 2. Fixed Effect Poisson Estimation Results of Metro-to-Nonmetro County In-Migration between 2005 and 2009

	Model 1 *Reading Test Scores*		Model 2 *Math Test Scores*	
	Est. Coeff.[1] *(S.E.)*	*M.E.*	*Est. Coeff.*[1] *(S.E.)*	*M.E.*
wage	0.037*** (0.003)	3.811	0.039***(0.003)	3.951
popden00	0.630*** (0.042)	87.696	0.637***(0.042)	89.089
distance	-0.887***(0.053)	-58.816	-0.887***(0.053)	-58.804
distance_sq	0.028*** (0.003)	2.860	0.028***(0.003)	2.858
indusmix9500	0.036*** (0.004)	3.670	0.037***(0.004)	3.803
amnscale	0.057*** (0.008)	5.836	0.057***(0.008)	5.841
elder	-0.095***(0.005)	-9.103	-0.093***(0.005)	-8.895
metroadj	0.209***(0.024)	23.238	0.207***(0.024)	22.989
pctax02	-0.438***(0.034)	-35.482	-0.451***(0.034)	-36.311
pcgexp02	-0.167*** (0.015)	-15.398	-0.172***(0.016)	-15.765
mhv	0.006***(0.000)	0.607	0.006***(0.000)	0.635
cvurate00_09	-0.215 (0.159)	-19.353	-0.158(0.158)	-14.619
dropoutrt	-0.013***(0.002)	-1.301	-0.015***(0.002)	-1.507
wpctreading	0.013*** (0.001)	1.259	---	---
wpctmath	---	---	0.009***(0.001)	0.898
Log L.	-4,801,690.6		-4,803,696.7	
Wald	15,825.50		15,532.56	
State Fixed Effect[2]	Yes		Yes	
County Fixed Effect[3]	Yes		Yes	
Obs.	2,156,608		2,156,608	

Notes:
[1] *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$
[2] State fixed effects are for the destination county.
[3] County fixed effects are for the origin county.