Human Capital Formation in Rural America: Differences between Rural and Urban High School Dropouts

Elton Mykerezi
Department of Applied Economics
University of Minnesota
218F Classroom Office Building
1994 Buford Ave.
St. Paul, MN 55108
Tel.: 612-625-2749, Fax: 612-625-2768
E-mail: myker001@umn.edu

Jeff Jordan
Department of Agricultural and Applied Economics
1109 Experiment St.
Room 206, Stuckey Building
University of Georgia Griffin Campus
Griffin, GA 30223-1797
Ph: (770) 412-4787, Fax: (770) 412-4789
E-mail: jjordan@uga.edu

Genti Kostandini
Department of Agricultural and Applied Economics
University of Georgia
1109 Experiment Street
223 Stuckey Building
Griffin, GA 30223-1797
Tel.: 770-228-7231 x 121, Fax: 770-228-7208
E-mail: gentik@uga.edu

Human Capital Formation in Rural America: Differences between Rural and Urban High School Dropouts

Introduction

It is well-known that remaining in school at least through high school graduation is vital to staying out of low-wage America. In addition to lower wages, students who don't finish high school are more likely to be unemployed, to end up in prison, to need public assistance and to die at a younger age. High dropout also has social costs reflected in lost tax revenue and increased expenditures for health care, corrections, food stamps, subsidized housing and public assistance, making drop-out prevention a priority for policy. Yet, many continue to drop out of school early. A better understanding of the causes leading to high dropout rates is crucial in the design of effective policy.

In studies conducted since the 1970s scholars have isolated dozens of predictors for students who are likely to drop out. Test scores and poor grades while important are not the only determinants of dropouts. Early studies suggest that dropouts have low self-esteem and find school quality and the decision to work to affect the decision to drop out. More recent studies have also examined different factors that influence the decision to drop out of high school, finding that youths who drop out of high school have lower ability and/or motivation, lower expectations after graduation and put higher value on leisure compared to those who graduate.

Youths living in rural areas may face different socioeconomic characteristics as well as a unique educational environment compared to their urban counterparts. Thus it is also important to investigate the determinants of high school dropouts on rural and urban areas separately. While urban students drop out more frequently recent studies suggest

that rural dropouts were more likely to report getting a job or not getting along with the teacher as causes for drop out and they gave lower ratings to school effectiveness and discipline. In addition, rural dropouts are more likely to be American Indian or White while urban dropouts were more likely to be Black or Hispanic. These analyses are however, largely descriptive in nature. Potential fundamental differences in the mechanisms that lead to dropping out across rural versus urban areas have not been investigated in a multivariate framework with recent data.

In addition, scholars have even been in stark disagreement on such a basic statistic as the U.S. high school graduation rate. For instance, in various studies that used a variety of data sources and definitions, the U.S. graduation rate has been estimated to be anywhere from 66 to 88 percent in recent years. The range of estimated minority dropout rates is particularly high; from 50 to 85 percent (Heckman and LaFontaine, 2007). Because of these difficulties debates regarding the dropout rates, their distribution across race and ethnicity lines, and time trends persist (Mishel and Roy, 2006; Mathews, 2006; Chaddock, 2006; Heckman and LaFontaine, 2007). There has been virtually no discussion of issues surrounding rural graduation rates and differences with urban graduation rates. Youths living in rural areas may face different socioeconomic characteristics as well as a unique educational environment compared to their urban counterparts. Given differences in urban and rural life, studies that investigate high school dropout rates and their causes for rural and urban areas separately may provide important insights to policy makers.

In a 2007 study, Heckman and LaFontaine systematically considered the sources of bias across a number of nationally representative datasets and documented multiple sources of bias in data and methods for estimating the US high school drop out rate Dropout rates in rural areas are rarely estimated, and to our knowledge no recent effort has been made to estimate bias-free dropout rates for rural areas. Lack of attention regarding how dropout rates are measured may, in part, be responsible for conflicting findings on the question of whether rural high school students are significantly different that their urban counterparts (e.g. McCaul,1988; Fan and Chen, 1999; Roscigno and Crowley, 2001). There have been conflicting findings on this question even among studies that use the same dataset. Reaves and Bylund (2005) state that some of the conflicts in findings can be explained by differences in what is considered "rural" and on empirical methods.

The only other indicator of the dropout rate for rural areas at a national level is provided by the U.S. Department of Agriculture's Economic Research Service (ERS). However, they only publish completion rates among adults 25 years of age or older who live in non metropolitan areas using census data. This does not reflect current dropout conditions as they may partly reflect past drop out rates and migration.

In this study we use recent and past national representative data sets to provide an in depth analysis of high school dropout rates in the U.S. correcting for any potential bias and paying particular attention to urban-rural differences. More specifically, in this study we: (1) Use recent, geo-coded nationally representative data from the National Longitudinal Survey of Youth in 1997 (NLSY97) and adopt the recommendations of Heckman and LaFontaine (2007) to reduce data bias in several categories of locations of

various degrees or "rurality"; (2) base our discussion on four broad categories of locations based on Beale Codes, a classification system that is widely used by the ERS to categorize US counties in terms of "rurality"; (3) Use rich, household level data that allows us to address questions of whether the causes of rural dropout are different from urban; (4) use geo-coded household level data from a similar cohort of youth as the NLSY97 but who attended high school in the late 70's and early 80's to examine whether the rural-urban differences in graduation and its correlates have changed over the last 30 years. Of particular note is the fact that using geo-coded data allows us to examine whether the nature of rural labor markets is now, or has been in the past, particularly important in the decision by rural youths to leave school before graduation.

Background on High School Dropout Rates

In studies conducted since the 1970s scholars have isolated dozens of predictors for students who are likely to drop out. According to researchers, there are a number of factors that influence one's decision to drop out of highs school. Test scores and poor grades while important are not the only determinants of dropouts (McCaul 1988: Rumberger 1983). Early studies suggest that dropouts have low self-esteem (e.g. Rumberger 1983) and find school quality and the decision to work (e.g. McCaul 1988) to affect the decision to drop out. More recent studies have also examined different factors that influence the decision to drop out of high school. For example, Ekstein and Wolpin (1999) use the 1979 National Longitudinal Survey of Youth (NLSY) and, among other factors, find that youths who drop out of high school have lower ability/and or motivation, lower expectation after graduation, and put higher value on leisure compared

5

to those who graduate. Montmarquette et al. (2007) using a dataset from Canada indicate that parent education and attending a private school reinforce the decision to favor schooling over labor. Their study also points out that the legal age to access the labor market, high minimum wages and low unemployment rates influence the decision to drop out.

Studies have found rural urban differences on both high school dropout rates and the likely causes of dropping out. For example, Pallas (1987) finds that urban students drop out more frequently. McCaul (1988) examines the differences in drop out rates between rural and urban dropouts. His study suggests that rural dropouts were more likely to report getting a job or not getting along with the teacher as causes for dropping out and they gave lower ratings to school effectiveness and discipline. In addition, rural dropouts were more likely to be American Indian or White while urban dropouts were more likely to be Black or Hispanic.

The question of whether rural schools are inferior to urban and suburban schools has produced conflicting results sometimes even among studies that used the same dataset. For instance, using data from the National Education Longitudinal Survey (NELS) Fan and Chen (1999) concluded that "Rural Schools do as well as non-metro schools; sometimes better" (pg. 42). Roscigno and Crowley (2001) also employed NELS data and found that students living in rural areas of the U.S. exhibit lower levels of educational achievement and a higher likelihood of dropping out of high school than do their non-rural counterparts. Reaves and Bylund (2005) suggest that the conflicting findings are mostly due to the use of different methods and definitions for what is rural. Specifically, they state: "Divergent findings were obtained by the three teams of

6

researchers—Fan and Chen (1999), Roscigno and Crowley (2001), and Israel, Beaulieu, and Hartless (2001)—all of whom relied on what was substantially the same NELS data. However, as one reads the reports of these separate investigations, it becomes clear that varying problem orientations, research designs, and definitions of variables can lead to divergent results and conclusions."

ERS's published matriculations of educational attainment in non-metro areas for adult's aged 25 years or older show that 31.3 percent of people had less than high school in nonmetro areas in 1990 vs. 23.1 percent in metro areas and in 2000 these rates were 23.2 percent and 18.7 percent for metro and nonmetro areas, respectively (ERS website). These may provide an indicator of whether rural dropout is higher or lower, but they are very noisy indicators of recent dropout rates. First, the fact that only rates for all adults 25 years of age or older are published makes it difficult to observe whether the rate is attributable to recent dropouts or whether they reflect conditions decades ago. Second, these estimates reflect migration decisions as well as graduation rates. For instance, even if graduation rates in rural areas are similar to those in urban areas, but there is a net migration of educated adults from rural to urban areas, the share of adult completers would be lower in non-metro areas. Also, time trends in the share of adults with degrees are rather uninformative as they encompass graduation rates over time, changes in net out migration and in return migration (usually of older households) (references).

Another indicator of dropout is provided by Balfanz and Legters (2004), in a study identifying schools with particularly high rates of non completion (drop-out factories). The high dropout in these southern dropout factories appears to be, in part, because of rural poverty. This approach is useful for policy, but it does not shed light on

whether the reasons for dropping out are different in rural than in urban areas, nor does it provide a clear comparison of how rural schools do relative to urban schools.

Methodology and Empirical Strategy

In this study we follow Heckman and LaFontaine (2007) in define who is a 'high school graduate' and how to isolate graduates with 1979 and 1997 cohorts of the NLSY. They argue that there are several sources of potential bias in many studies and datasets and they recommend several corrective steps. First, they argue strongly against the inclusion of individuals who have a GED since it appears to bias the estimates especially when it comes to minorities as they obtain disproportional shares of GED. Second, they argue that the prison population should be included in these calculations. Third, the inclusion of immigrants who have come to the U.S. after they completed high school causes downward bias on the estimates and they should be excluded. Fourth, bias in the coverage of the dataset that is used should be determined. Finally, they recommend the use of the eighth-grade enrollment as the base for dropout estimation.

We designate dropouts consistently. We then estimate probit models to address two questions: Are rural youth at more or less risk of dropping out relative to their urban counterparts? Are the determinants of dropping out different for rural than for urban students? To address the first question we estimate probit regressions with the whole population and include indicators of rurality. To address the second question, we estimate separate models by rural/urban zones. The above models are estimated for two cohorts (1997 and 1979) to examine possible changes in rural drop out rates and the determinants of dropping out over the last 20 years. We only control for a limited set of variables that

8

reflect conditions at or prior to 8th grade. Specifically we control for familial composition in 1997 for the NLS 97 and at age 14 for NLS79 cohort. While information on school attributes and individual performance in school is available we do not explicitly control for these variables in the reduced form models employed in this paper.

Data

This study employs data from the NLSY79 and the NLSY97 survey from the Bureau of Labor Statistics. Both samples are nationally representative. The NLSY79 consists of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979. They were interviewed annually through 1994 and currently are being interviewed biannually. The NLSY97 consists approximately 9,000 youths who were 12 to 16 years old as of December 31, 1996. The first round of the survey started in 1997 with both the parents and the youths interviewed and youths continue to be interviewed on an annual basis.

Based on the Beal Codes (Table 1), we grouped individuals in four categories: those living in cities (zone 1) if the Beale code is equal to 0 or 1, suburban/metro area (Zone 2) if code is 2 or 3, adjacent non-metro (Zone 3) if code is 4 5 or 6, remote non-metro (Zone 4) if code is 7, 8 or 9.

5. Results

We focus our discussion on two sets of results. First we just look at drop out rates using NLSY79 and NLSY97 for the entire sample and then for each rural 'zone' separately

(table 2). Second we look at the determinants of dropping out using NLSY97 (table 3) and NLSY79 (table 4).

Dropout estimates in table 2 suggest that rural zones have very similar graduation rates compared to the whole sample. Results for the nationally representative samples are very similar to the findings of Heckman and LaFontaine (2007) suggesting a decline in graduation the rate. Specifically, estimates show a 3 percentage point decline in the overall graduation rate. Descriptive statistics also indicate that the rates are very similar across zones in both years (23.1%, 22.7%, 22.9% and 22.9% in zones 1 through 4 respectively in 1997). In both years, however, suburban areas (zone 2) show the lowest drop out rates. In 1979 the highest dropout is in remote rural areas (zone 4), but by 1997 urban areas (zone 1) show the highest drop out. The increase in dropout over time in urban areas is exactly equal to the national average (3.2 points), while the increase in zone 2 is 3.8 points and in zone 4 it is only 1.4 points.

Tables 3 and 4 present marginal effects associated with probit models that relate the probability of graduation to indicators of rural zone for the 1997 and 1979 cohorts respectively. Estimates indicate that there is no evidence that the dropout rate is higher in rural areas for the 1997 cohort as both estimates that do and do not control for gender, race, ethnicity, family size, composition and economic hardship show no statistically significant differences in drop out by location. In 1979, uncontrolled estimates show statistically similar rates of drop out across regions, but once familial background race ethnicity and gender are controlled for, remote areas appear to have slightly smaller dropout ($p<0.1$). Thus, we find suggestive evidence that schools in remote areas had an advantage in the 80s, but that is no longer the case. Similar to the findings of McCaul

(1988) our results suggest that Blacks and Hispanics have a lower probability of graduation in both the 70s and the 90s relative to whites. In addition, all familial arrangements have a strong negative impact on graduation relative to living with both biological parents. Males are also at a disadvantage relative to females.

In 1997, males show the lowest disadvantage to females in zone 1 and the highest in zone 4, with urban male youth being 4.1 percentage points less likely to graduate than their female counterparts but males in remote areas are 6.9 percentage points less likely to graduate than rural females. The only other notable difference is that in remote areas (Zone 1) Blacks aren't less likely to graduate then their white counterparts. Also, Hispanics are less likely to graduate than their non Hispanic counterparts in zones 1 and 2 but not 3 and 4.

6. Concluding remarks

This study estimates the high school dropout rate in rural and urban areas using recently developed methods that attempt to reconcile apparent discrepancies in the U.S. 'real' dropout rates. Our findings suggest that high school dropout rates are higher by 3 percentage points the 90s compared to the 70s.Blacks and Hispanics seems to have a disadvantage compared in both rural and urban areas. In addition males show a lower rate of high school graduation when compared to females.

The main objective of this study was to have a closer look at the high school differences in dropout rates between urban and rural areas. However, when we consider different types of rural areas it appears that their graduation rates are very similar to those in the urban areas. Thus rural areas do not seem to be in a disadvantage in the 90s.

References

Balfanz, R., and Legters, N. 2004. Locating the Dropout Crisis. Report by the Center for Research on the Education of Students Placed at Risk, John Hopkins University, Baltimore, MD.

Chaddock, G.R. 2006. 'US high school dropout rate: high, but how high?' The Christian Science Monitor, June 21. Accessed April 2010 at: http://www.csmonitor.com/2006/0621/p03s02-ussc.html.

Eckstein, Z., and Wolpin, K.I. 1999. Why Youths Drop Out of High School: The Impact of Preferences, Opportunities and Abilities. *Econometrica* 67(6): 1295-1340.

Fan, X., and Chen, M. J. 1999. Academic Achievement of Rural School Students; A Multi-year Comparison with their Peers in Suburban and Urban schools. *Journal of Research in Rural Education* 15:31-46.

Greene, J.P., and Winters, M.A. 2006. Counting Diplomas and the 9-th Graders. Washington Post, May 23. Accessed April 2010 at: http://www.washingtonpost.com/wp-dyn/content/article/2006/05/22/AR2006052201189.html.

Heckmam, J.J., and LaFontaine, P.A. 2007. The American High School Graduation Rate: Trends and Levels. Discussion paper of the Institute for the Study of Labor, Bonn, Germany.

Israel, G. D., Beaulieu L.J., and Hartless, G. 2001. The Influence of Family and Community Social Capital on Educational Achievement. *Rural Sociology*. 66(1):43-68.

Mathews, J. 2006. Dropout data raise questions in two fronts. Washington Post, May 23. Accessed April 2010 at: http://www.washingtonpost.com/wpdyn/content/article/2006/05/22/AR2006052201187.html

McCaul, E. 1988. Rural Public High School Dropouts: Data from High School and Beyond. Paper presented at the Annual Conference of the New England Education Research Organization, Rockland, ME, April 27.

Miller, S. 2010. 'Obama Tackles High Dropout Rates, Targets 'Chronically Troubled' Schools', ABC News, March 1. Accessed April 2010 at: http://blogs.abcnews.com/politicalpunch/2010/03/obama-tackles-high-dropout-rates-targets-chronically-troubled-schools-.html.

Michel, L., and Roy, J. 2006. 'Tracking Students Over Time Using Surveys', Washington Post, May 23. Accessed April 2010 at: http://www.washingtonpost.com/wp-dyn/content/article/2006/05/22/AR2006052201197.html

Montmarquette, C., Viennot-Briot, N., and Dagenais, M. 2007. Dropout, School Performance and Working while in School. *The Review of Economics and Statistics* 89(4): 752–760.

Pallas, A.M. 1987. School Dropouts in the United States. In W. T. Denton (Ed.), *Dropouts, pushouts and other casualties* (pp. 23-39). Bloomington, IN.

Reeves, E.B., and Bylund, R.A. 2005. Are Rural Schools Inferior to Urban Schools? A Multilevel Analysis of School Accountability Trends in Kentucky. *Rural Sociology* 70:360-386.

Roscigno, V.J. and Crowley, M. 2001. Rurality, Institutional Disadvantage, and Achievement/Attainment. *Rural Sociology* 66:268-292.

Rumberger, R.W. 1983. Dropping out of High School: The Influence of Race, Sex, and Family Background. *American Educational Research Journal* 20: 199-220.

Table 1. Description of the Rural-Urban Continuum Codes (1983-2003)	
Code	Description
Metro:	
0	Central counties of metro areas of 1 million population or more.
1	Fringe counties of metro areas of 1 million population or more.
2	Counties in metro areas of 250,000 to 1 million population.
3	Counties in metro areas of fewer than 250,000 population.
4	Urban population of 20,000 or more, adjacent to a metro area.
5	Urban population of 20,000 or more, not adjacent to a metro area.
6	Urban population of 2,500 to 19,999, adjacent to a metro area.
7	Urban population of 2,500 to 19,999, not adjacent to a metro area.
8	Completely rural or less than 2,500 urban population, adjacent to a metro area.
9	Completely rural or less than 2,500 urban population, not adjacent to a metro area.

Table 2. High school graduation rate

	NSLY97				
	All	Zone 1	Zone 2	Zone 3	Zone 4
High school dropout rate	0.229	0.231	0.227	0.229	0.229
High school graduation rate	0.771	0.769	0.773	0.771	0.771
Number of observations	8984	4238	3025	981	740
	NLSY79				
High school dropout rate	0.197	0.199	0.189	0.199	0.215
High school graduation rate	0.803	0.801	0.811	0.801	0.786
Number of observations	10755	4684	3497	1489	1085

Table 3. Determinants of dropping out (NLSY 97)

	All Dy/dx	S.E.	All Dy/dx	S.E.		Zone 1 Dy/dx	S.E.		Zone 2 Dy/dx	S.E.		Zone 3 Dy/dx	S.E.		Zone 4 Dy/dx	S.E.	
Zone 2	0.004	0.011	0.000	0.011													
Zone 3	0.001	0.016	-0.012	0.016													
Zone 4	0.001	0.017	-0.016	0.018													
Black			-0.043	0.012	***	-0.053	0.018	***	-0.038	0.021	**	-0.100	0.038	***	0.078	0.040	*
Hispanic			-0.065	0.013	***	-0.078	0.018	***	-0.071	0.023	***	-0.015	0.060		0.072	0.058	
# Kids			-0.012	0.007		-0.014	0.010		-0.003	0.013		-0.037	0.023		0.007	0.029	
# Kids younger than 6			-0.006	0.010		0.013	0.015		-0.018	0.016		-0.016	0.030		-0.052	0.037	
Family Size			-0.010	0.006		-0.011	0.009		-0.011	0.011		0.016	0.020		-0.039	0.027	
Two adults one biological			-0.172	0.017	***	-0.173	0.027	***	-0.159	0.029	***	-0.184	0.051	***	-0.213	0.056	***
Single biological parent			-0.191	0.015	***	-0.180	0.021	***	-0.196	0.025	***	-0.182	0.045	***	-0.237	0.056	***
Other family arrangement			-0.240	0.027	***	-0.226	0.040	***	-0.258	0.047	***	-0.261	0.077	***	-0.213	0.088	**
Family has been through hard times			-0.074	0.025	***	-0.075	0.038	**	-0.057	0.040		-0.109	0.070		-0.070	0.080	
Male			-0.054	0.009	***	-0.041	0.014	***	-0.065	0.016	***	-0.062	0.028	**	-0.069	0.032	**
N	8994		8994			4238			3025			981			740		

Note: *, ** and *** indicate statistical significance at the 10, 5 and 1 percent level, respectively.

Table 4. Determinants of dropping out (NLSY79)

	All		All			Zone 1			Zone 2			Zone 3			Zone 4		
	Dy/dx	S.E.	Dy/dx	S.E.		Dy/dx	S.E.		Dy/dx	S.E.		Dy/dx	S.E.		Dy/dx	S.E.	
Zone 2	0.010	0.010	0.005	0.010													
Zone 3	0.000	0.014	-0.014	0.014													
Zone 4	-0.015	0.016	-0.029	0.017	*												
Black			-0.144	0.014	***	-0.176	0.020	***	-0.110	0.022	***	-0.081	0.045	*	-0.074	0.074	***
Hispanic			-0.049	0.011	***	-0.080	0.016	***	0.010	0.017		-0.010	0.029		-0.109	0.036	
Two adults one biological			-0.190	0.019	***	-0.211	0.030	***	-0.152	0.033	***	-0.253	0.050	***	-0.092	0.061	***
Single biological parent			-0.152	0.015	***	-0.164	0.022	***	-0.124	0.027	***	-0.134	0.045	***	-0.177	0.052	***
Other family arrangement			-0.265	0.031	***	-0.241	0.048	***	-0.276	0.052	***	-0.372	0.079	***	-0.200	0.098	**
Male			-0.060	0.009	***	-0.072	0.013	***	-0.052	0.016	***	-0.065	0.024	***	-0.018	0.030	
N	10755		10755			4684			3497			1489			1085		

Note: *, ** and *** indicate statistical significance at the 10, 5 and 1 percent level, respectively.