The Relationship between Educational Opportunity and Upward Mobility in Rural America

Gabrielle Torrance
Stanford University
torran20@stanford.edu

Presented at the 45th Annual Conference of the Association for Education Finance and Policy
March 20, 2020

PRELIMINARY DRAFT– PLEASE DO NOT CIRCULATE WITHOUT PERMISSION

Abstract

Using data from the Opportunity Insights project and the Stanford Education Data Archive, I examine the relationship between educational opportunity and upward mobility in the rural United States. I consider both third grade achievement scores and “growth rates” as measures of educational opportunity. After controlling for a series of demographic-, economic-, and institution-related factors, I observe a significant correlation between growth rates and mobility. In a fixed effects model, controlling for educational opportunity, race, and socioeconomic status, I find that third grade achievement scores significantly predict mobility. However, in both models, the effect of education on rural upward mobility is substantially weaker than that of demographic factors such as race and socioeconomic status. I also examine effects of educational opportunity on college outcomes in rural areas, finding significant correlations between college outcomes and both measures of opportunity.
Introduction

Upward mobility rests at the center of American conceptions of opportunity, symbolizing the notion that with hard work and dedication, individuals can build a life more prosperous than that of their parents. Recent findings from the Opportunity Insights project, led by Harvard economist Raj Chetty, have renewed interest in the concept of intergenerational mobility and revealed striking levels of geographic variation in opportunity across the United States. One of the project’s more striking revelations is that average levels of upward mobility are higher in rural areas than urban ones. In recent years, the growing threat of automation, a bitterly divisive national politics, and the continued flight of young workers to coastal metropolises have drawn attention to the challenges of growing up in rural America and bolstered the narrative that rural communities are being left behind as urban ones prosper, rendering Chetty’s results somewhat counterintuitive. Despite these findings, however, little research has sought to understand whether the unique socioeconomic characteristics of rural regions may impact those factors most strongly correlated with mobility. Schools, in particular, are known to play a more critical role in the social and economic activity of rural regions, implying that educational opportunity and school quality could play a greater role in the upward mobility of rural students than of urban ones. With this project, I aim to understand the factors most strongly correlated with rural upward mobility and the extent to which the effect of those factors differ across rural and nonrural areas. Given the unique role of schools in rural regions, I pay particular notice to the relationship between upward mobility and educational opportunity, with the goal of understanding how education contributes to long-term outcomes in areas which lack many of the economic and cultural resources of a metropolis.
This project will be the first to utilize the latest editions of both the Opportunity Insights and Stanford Education Data Archive (SEDA) databases to analyze upward mobility in the rural context. This data is fairly comprehensive, covering over 85 percent of counties in the rural United States. Using this data, I am able to comment on mobility differences by subgroup, unlike previous rural mobility literature. In addition to describing mean mobility by rurality and subgroup, I analyze a series of demographic-, education-, economics-, and institution-related covariates, including third grade achievement, eighth grade achievement, and growth rate variables aimed at capturing local educational opportunity. Further, though not the primary focus of this work, I examine the effects of these covariates on four-year college degree completion, to develop a preliminary understanding of whether these covariates affect life outcomes broadly or are most relevant to mobility in particular. I also briefly analyze individual mobility outcomes, in addition to traditional household mobility, though this will largely remain an avenue for exploration in future analysis. This paper is purely descriptive and will not seek to comment on the causality between educational opportunity or other covariates and upward mobility.

The future of rural America remains uncertain, with factors such as automation and continued out-migration threatening existing ways of life and potential for future prosperity. While this research will not aim to address every factor with potential to disrupt rural life outcomes, it will shed light on the link between educational opportunity and long-term economic outcomes in a region of the country which has received little academic notice, with the intention of informing future policy development. State and federal lawmakers are often unfamiliar with the unique settings of rural communities, but with additional insight, they may
develop a greater understanding of the factors most correlated to the long-term success of rural children. My hope is that this research will provide the context needed to inform the development and prioritization of local and national reforms that will promote equality of opportunity across the rural-urban divide.

Background

Upward mobility

The topic of upward mobility in the United States has been the subject of substantial scrutiny in recent literature, particularly in the wake of novel findings from the Opportunity Insights project. In a 2014 study, Chetty and coauthors analyze income tax records from a sample of 10 million children, along with their parents, to examine the ability of individuals born into lower-income families to reach higher income percentiles later in life, thereby offering one of the first large-scale, data-driven windows into intergenerational mobility (Chetty, Hendren, Kline, & Saez, 2014). Strikingly, the authors find that upward mobility varies extensively by geography; while mobility is lowest across the Southern states, many high-opportunity regions are clustered in the Upper Midwest. This research also finds that outcomes are often more highly correlated to community settings than to an individual’s own family situation; for example, even if a child lives in a two-parent family, if she resides in a neighborhood with a high number of single mothers, her mobility outcomes are likely to resemble those of her peers in one-parent households. By noting that such outcomes vary substantially nationwide, and that moving to a better neighborhood in childhood does improve a child’s life outcomes, this work also
underscores the potential to impact upward mobility through place-based policies (Wellisz, 2018).

**Rurality and outcomes**

In both the 2014 paper and a 2017 study of county-level effects on mobility, Opportunity Insights notes that on average, rural areas exhibit higher levels of intergenerational mobility than urban ones (Chetty, Hendren, Kline, & Saez, 2014; Chetty and Hendren, 2017). Chetty and Hendren also note that counties with higher levels of mobility tend to have lower levels of income inequality, poverty concentration, and crime, lower high school dropout rates, and a lower share of single-parent households (Chetty and Hendren, 2017). While they do not conduct a substantive analysis of whether these factors vary in significance between rural and urban areas, subsequent research has utilized the project’s data to do so. Oregon State economist Bruce Weber has pioneered much of this spatial research, first analyzing mobility across metropolitan, micropolitan, and non-core (rural) counties in 2017. Weber finds that race and single-parent households are most strongly correlated with upward mobility across all geographies, and that rural upward mobility is more closely tied to higher rates of out-migration than urban mobility (Weber, Fannin, Cordes, & Johnson, 2017).

**Rurality and education**

One of the key sociological factors which distinguishes remote communities from their metropolitan counterparts is the importance of community institutions, particularly schools. While schools serve central roles in many communities, research suggests that they often operate as more prominent hubs of social, cultural, and economic activity in rural regions than urban ones. In the more intimate rural context, educators are often held more accountable for
students’ success; because of a more robust community buy-in, schools become catalysts for local development, providing a locus of community investment and integration (Schafft, 2016). In recent years, for example, several rural communities have sought to capitalize on this link between education and economic development through the creation of school-community partnerships and "grow your own" teacher initiatives. Given the strength of this connection between rural schools and their communities, I consider education not only as a correlate with upward mobility broadly, but as a factor potentially driving a differentiated path to upward mobility in rural regions.

[to be continued]

Research Questions

The overarching aim of this project is to identify factors strongly associated with upward mobility in rural counties, with an eye toward educational opportunity covariates; the secondary focus will be analyzing college completion and individual mobility outcomes. I conduct an initial descriptive analysis of upward mobility by rurality and by subgroup, examining gender, race, and economic status, to understand whether rural or urban areas offer greater opportunities for youth. Next, I comment on correlations between mobility outcomes and demographic, educational, economic, and institutional covariates, before considering whether the type of rural county in question significantly correlates to mobility outcomes, utilizing U.S. census divisions, rural-urban continuum codes, and economic subtypes to do so. Finally, I employ a fixed effects analysis to comment on the extent of the relationship between educational opportunity and mobility outcomes within counties. Again, I do not attempt to draw conclusions based on
causality. Future drafts will also consider what policies should be considered, in light of these findings.

It is worthy of mention that upward mobility is a specific, not holistic, metric of economic success. This outcome intends to capture the ability of lower-income individuals to reach higher economic classes. It is not necessarily a measure of current economic prosperity, and critics have noted that mobility metrics provide little insight into current standards of living or recent economic growth (Gilbert, 2017). Nevertheless, it remains a useful metric for commenting on the presence of opportunity nationwide.

Data

Outcome Data

The primary outcome of interest for this paper is household upward mobility, though four-year college completion rates and individual upward mobility are also analyzed at points. Data for all outcomes are collected from the Opportunity Atlas data set, which is publicly available on the Opportunity Insights website. This sample was constructed from 20.5 million income tax records from children born from 1978-1983, along with the income tax records of their parents. It includes individuals who were either born and raised in the U.S. or who immigrated at a very young age, and who were claimed as dependents by at least one parent. The sample is comprehensive, including 96.2 percent of children in the 1978-1983 birth cohorts (Chetty, Friedman, Hendren, Jones, & Porter, 2018).

To quantify upward mobility, Chetty et al construct a national income ranking for every birth cohort in the sample. To do so, every person in the sample is assigned a rank according to
their location in the income distribution of all individuals born in the same year, as measured in 2014-15. A parental income ranking is also constructed, comparing the incomes of all parents with a child in that birth cohort. To translate this ranking into a measure of mobility, the Opportunity Insights project then calculates the predicted percentile rank in the national income distribution among all individuals whose parents were ranked at a certain income percentile in a single county. In this analysis, “upward mobility” is based on this predicted rank from the 25th percentile, with one exception: the upward mobility metric for the non-economically disadvantaged subgroup is defined as the predicted percentile rank for individuals whose parents were ranked at the 75th income percentile. This ranking method, while perhaps less intuitive than directly comparing intergenerational income levels, helps mitigate outlier bias and increase stability in comparison to income measurements using dollar amounts (Chetty, Friedman, Hendren, Jones, & Porter, 2018). Further, note that the unit of analysis for this research is the county; while Opportunity Insights also provides tract-level outcome data, I conduct my analysis at the county level in order to maximize covariate availability. Chetty et al. further differentiate upward mobility by constructing household and individual mobility outcomes. To calculate the former, children are ranked based on their position in the household income ranking, for which their income and that of their spouse are combined, based on marital status in 2014-15. Individual mobility is constructed based only on the child’s annual income. I use household mobility as the primary outcome in this paper, given it is less prone to annual fluctuations and generally preferred in the Opportunity Insights research (Chetty, Hendren, Kline, & Saez, 2014). However, to develop a more robust understanding of how outcomes vary by rurality, I incorporate both individual mobility along with four-year college
degree completion in my analysis. The four-year completion outcome is also available through the Opportunity Atlas, and it represents the fraction of children in the sample who attain a four-year college degree.

Finally, outcome data is available by subgroup, including gender and race/ethnicity. I also construct a non-economically disadvantaged subgroup, using the mobility outcome for children born into households at the 75th income percentile. While categorizing by subgroup is useful, there remain challenges with data availability in more remote counties, particularly by race. Table 1 highlights the number of observations available for each subgroup outcome in rural and nonrural counties, along with means for both.

*Covariate Data*

To analyze which factors may correlate to rural upward mobility, I incorporate a series of covariates that are likely to have some relationship with upward mobility generally, or that may display some differences in value in rural areas, based on a review of the literature. I classify these covariates into one of four buckets: demographic-, education-, economic-, or institution-related factors. The demographic covariates include income segregation, fraction above the poverty line, fraction black, household income per capita, and fraction with single mothers. Education covariates include per-pupil school expenditures, number of colleges per capita, and three measures of educational opportunity. Gini coefficient, fraction with commute time less than 15 minutes, teenage labor force participation, unemployment rate, and in- and out-migration comprise the economic covariates, while institutional covariates are local tax rate per capita, fraction religious, violent crime rate, and a social capital index created using principal components analysis on county-level voter turnout rates, Census response rates, and number of
political, civic, recreational, business, and non-profit establishments per capita (Goetz & Rupasingha, 2014). While these buckets are not always mutually exclusive, this is an initial attempt at gathering a holistic view of what drives rural mobility, with an eye toward tangible policy takeaways.

With the exception of the educational opportunity variables, all of the bucketed covariates are collected at the county level from the publicly available dataset used in Chetty & Hendren (2017). These covariates are themselves gathered from a variety of sources, which are documented in Table A1. I standardize each covariate to have a mean value of 0 and standard deviation of 1, to enable comparability among regression estimates.

The educational opportunity variables used in this analysis are drawn from the Stanford Education Data Archive (SEDA, version 3.0), which provides nationally standardized student achievement and growth estimates for nearly every county, district, and school in the U.S. SEDA data is gathered from the U.S. Department of Education’s EDFacts database, along with the Common Core of Data and Civil Rights Data Collection. The data collected includes annual third through eighth-grade assessment scores in mathematics and English/Language Arts (ELA) from each state, spanning the 2008-09 through 2015-16 school years. SEDA then links these raw scores to a national scale using the National Assessment of Educational Progress, enabling comparison of achievement scores across the nation for the first time. Per SEDA’s interpretation, the key measures of educational opportunity include third grade achievement scores and growth rates. The former are considered a measure of early childhood opportunities, which are closely tied to socioeconomic resources within the child’s community, both in and out of school (Reardon, 2019). Growth rates, or learning rates, are measured as the grade slope on mean
achievement from grades three through eight and are intended to reflect the amount of material learned annually, which is largely a reflection of school quality. I also include eighth grade achievement scores, which encompass both early childhood opportunities and school quality. However, in full models, eighth grade achievement is omitted to avoid collinearity. I standardize all three covariates to have county mean values of 0 and standard deviations of 1.

In addition to achievement and growth scores, I include the SEDA county-level socioeconomic composite for the fixed effects analysis. This composite, which is only available for white, black, and Hispanic subgroups, is constructed using estimates of median family income, the proportion of adults with a bachelor’s degree or higher, proportion of unemployed adults, proportion of households receiving SNAP benefits, household poverty rates, and proportion of households with children that are headed by a single mother (Fahle et al, 2018).

One drawback with utilizing SEDA data is that the assessments in question are gathered from the 2008-09 through 2015-16 school years, whereas most other covariates are collected closer to the year 2000, and which is nearer the period when individuals in the relevant birth cohorts were adolescents. However, both SEDA and the Opportunity insights team have conducted consistency analyses, finding that outcomes for both achievement and mobility change very little over time; Chetty and coauthors find that the power of these estimates to predict future outcomes declines by just 1% per year (Chetty, Friedman, Hendren, Jones, & Porter, 2018). While this difference in years could potentially create a noisier analysis, SEDA is now one of the most comprehensive educational data sources, and the only one to provide data on learning rates, making it an invaluable addition to this research.
Finally, I incorporate three regional dimension covariates to gauge whether the “type” of rural county is relevant to upward mobility. These dimensions include U.S. Census geographic division, level of rurality, and economic subtype. I recognize the notion of rurality is not easily defined; the U.S. government itself employs dozens of definitions of “rural.” I ultimately utilize the U.S. Department of Agriculture’s 2003 Rural-Urban Continuum Code, which assigns each county a classification of 1 (most urban) through 9 (most rural) based on population density and adjacency to a metropolitan area (Cromartie, 2019). This continuum is used to construct a rural indicator variable, for which counties with RUCCs 1-3 have values of 0, while RUCCs 4-9 have values of 1. To classify counties by economic subtype, I use the 2004 USDA county typology metric, which identifies counties dominated economically by industries including farming, manufacturing, mining, federal and state government, and services. ERS assigns these classifications based on proportion of labor earnings in relation to a set threshold; in other words, if a county’s agricultural earnings sit above a certain proportion of total county earnings, then this county is eligible for the farming-dependent classification (Pender, 2019). The six categories are mutually exclusive, with counties classified into the category in which they are the largest degree above the set threshold. Distributions of each covariate, along with definitions for geographic division and rural-urban continuum code, are listed in Tables A2-A4.

Descriptive Statistics

Of the 3,143 U.S. counties in 2010, this project’s core sample includes 3,136, or 99.8 percent. 2,047 counties, or 65 percent of this remaining sample, are rural. While seven counties are excluded because of missing upward mobility data, and four are excluded because of missing
Chetty covariate data, the comprehensiveness of this sample is a marked improvement over previous studies of rural mobility based on the Chetty 2014 dataset, which lacked outcome data for 336 rural counties (Weber, Fanni, Cordes, & Johnson, 2017). Among counties in the primary mobility sample, 79 are missing college outcome values, of which 78 are rural.

The sample used for simple regression analysis is dependent on availability of each covariate; Table 2 shows the number of observations and other descriptive statistics for each variable. Regressions on regional dimensions and covariate buckets are fit on a consistent sample of 1,797 rural counties, which excludes counties with missing covariate values. This sample represents 87.5 percent of rural counties in the US. For the fixed effects analysis, I omit observations with missing values for SEDA achievement, growth, and socioeconomic covariates.

Table 1 highlights the mean values and standard deviations in rural and nonrural areas, along with the p-value of the difference between those means, for household mobility, individual mobility, and four-year degree completion outcomes. These estimates are weighted by the number of children in each subgroup with parents below the median income percentile. To the question of where mobility is higher, the answer is dependent upon the metric of mobility and subgroup examined. Rural regions report higher levels of overall household mobility but lower levels of overall individual mobility and college completion. Looking only at overall means, one might be tempted to make a blanket assertion that mobility is higher in rural areas, but upon closer examination, this is likely driven by higher male mobility and the sizable male population. For every subgroup except males, nonrural individual mobility is greater than rural; for household mobility, the same is true for black, Hispanic, and Asian subgroups. Unsurprisingly, in
light of existing literature, college outcome values are lower in rural than nonrural counties for every subgroup, with the exception of the Native American subgroup.

Further, notable gender differences emerge depending on the outcome in question. Female household mobility rates are statistically equal by geography and higher than mean values for males. However, when examining individual mobility, female nonrural is statistically higher than female rural, and male values are greater in both geographies. This suggests that marital status more strongly influences female life outcomes than male, especially in rural counties.

[Table 1 here]

The mean values, standard deviations, and number of observations for each covariate are shown in Table 2. There are significant differences in rural and nonrural means for nearly all covariates. Notably, rural areas generally perform worse economically; they are more impoverished and have higher unemployment rates. However, they are less segregated racially and by income, and have greater levels of social capital and religious participation, along with lower crime rates. In terms of education, there is no significant difference in growth rates between rural and nonrural counties. Rural counties report lower average test scores despite higher per-pupil spending.

[Table 2 here]

Methodology

To develop a rudimentary understanding of the factors most correlated with upward mobility and college completion, along with the extent to which the effects of these factors
differ between rural and nonrural areas, I fit the simple regression model outlined in Equation 1 for each continuous covariate, using all counties in the sample. This model will indicate the strength of the correlation between the covariate and mobility in rural counties ($\beta_1$) along with the difference in this correlation between rural and nonrural counties ($\beta_4$). The model takes the form:

\[
\hat{Y}_i = \beta_0 + \beta_1 \text{Covariate}_i + \beta_2 i_{i}^{\text{Nonrural}} + \beta_3 \text{Pop Below p50}_i + \beta_4 \text{Covariate}_i \times i_{i}^{\text{Nonrural}} + \epsilon_i,
\]

where $\hat{Y}_i$ is the estimated average for either household upward mobility or four-year college completion in county $i$, $i_{i}^{\text{Nonrural}}$ is a binary variable indicating whether the county is rural or nonrural, and $\text{Pop Below p50}$ is the natural log of the number of children in the county with parents whose incomes fall below the national median. This population control is included to more accurately gauge the association between covariates and outcomes among counties with similar numbers of lower-income children.

Recognizing the potential for endogeneity among these covariates, I next fit a series of multivariate regressions using only rural counties. This model is outlined below:

\[
\hat{Y}_i = \beta_0 + \mathbf{X}_i \mathbf{B} + \epsilon_i,
\]

where $\mathbf{X}_i$ is a vector of covariates in rural county $i$. For the regional dimension analysis and any other models involving categorical covariates, $\mathbf{X}_i$ will include dummy variables for all but one category of the variable in question. In this draft, model 2 is only fit on the household upward mobility outcome.

After a broad analysis of factors associated with rural mobility and college completion, I fit a series of within-county fixed effects models, using race- and gender-specific outcomes and
education covariates for each county to further evaluate the relationship between educational
opportunity and the outcomes in question. These models take the form:

\[
\hat{Y}_{si} = \alpha_s + \gamma_i + \beta_1 Grd3\text{ Achievement}_{si} + \beta_2 Growth_{si} + X_{si} B + \epsilon_{si},
\]

where \textit{Grd3 Achievement} and \textit{Growth} are subgroup-specific values of the educational
opportunity covariates, \(\hat{Y}_{i}\) is the estimated average for the outcome in question for subgroup \(s\) in
county \(i\), and where \(X_{si}\) is a vector of covariates, including dummy indicators for student
subgroups and the SEDA socioeconomic composite. This model will be fit for household mobility,
individual mobility, and college completion outcomes, and for race/ethnicity and gender
subgroups.

The foremost limitation of this methodology is that the estimates from these models
cannot be interpreted causally. Further, while county-level analysis is useful for identifying
trends across communities, it insufficiently granular to illuminate resource trends at the
neighborhood or individual level. However, this methodology still enables descriptive insight into
factors which could be central to the success of children in rural communities.

**Results**

Table 3 reports coefficients from simple regressions on both household mobility and
college outcomes for each of the continuous covariates. I find statistically significant differences
in effects between rural and nonrural areas for nearly all covariates, but slightly fewer significant
differences for effects on college outcomes. Notably, growth rates, third grade achievement, and
eighth grade achievement all report positive coefficients on achievement scores and growth
rates for mobility and college completion. Growth rates are more strongly associated with
nonrural mobility than rural, while achievement scores appear to matter more for rural mobility but less for rural college outcomes, compared to their effects on nonrural outcomes. Among other covariates, demographic factors including fraction black and fraction with single mothers have a strong negative correlation with rural mobility but a statistically greater effect in nonrural areas, while the social capital index positively predicts mobility but is less strongly predictive in nonrural areas. The significance of these factors will be explored in more depth using multivariate regressions.

[Table 3 here]

I next limit my analysis to household mobility outcomes in rural areas. Table 4 highlights the results of multivariate regressions examining the effects of regional dimensions. I find that individually, level of rurality accounts for 13 percent of variation in mobility, economic dependency for 20 percent, and division for a striking 50 percent, while controlling for all three dimensions explains 59 percent of the variation. Even when controlling for all dimensions, all regions except the most remote (RUCC 8) display a significant negative association with mobility. While effects of the manufacturing subtype are somewhat weakened by including all dimensions, farming and mining subtypes maintain a strong positive correlation with the outcome in question. Some divisional differences also remain significant, with Southern divisions predicting lower levels of mobility.

[Table 4 here]

Having developed a preliminary understanding of the relationship between upward mobility and regional characteristics, I incorporate covariate buckets into the models. As outlined previously, these buckets include demographics, education, economics, and institutions, in
addition to regional dimensions. Findings are highlighted in Table 5. Each bucket except education explains more than half of the variation in household mobility outcomes, though the complete model explains 82 percent of variation, indicating a fair amount of covariate endogeneity. Controlling for all covariates, there is still some degree of regional variation present, as less remote counties are negatively associated with mobility. This suggests that regional differences represent more than differences in demographics or other key covariates.

Notably, despite a larger coefficient than many other covariates in Table 3, achievement is insignificant after including other covariates in the full model. While growth remains significant, its coefficient is smaller than that of several other covariates, suggesting education may not play as crucial of a role in driving upward mobility as hypothesized. Demographics, meanwhile, strongly predict rural mobility; interestingly, once controlling for all variables, fraction of individuals above the poverty line is negatively correlated, suggesting that counties with a larger number of nonpoor individuals are worse at improving outcomes for children in the lowest quartile. Covariates such as social capital and teenage labor force participation also remain significantly predictive in the full model.

Finally, I also fit the full model to the college completion outcome, as shown in the final column of Table 5. When controlling for all variables, I find third grade achievement significantly predictive of college completion, while many regional dimensions are rendered insignificant, suggesting the type of rural county matters less for college completion than upward mobility. In this vein, only 39 percent of variation in college attendance is explained by this regression, suggesting greater individual-level variation in college outcomes.

[Table 5 here]
Table 6 reports results of the fixed effects model on socioeconomic status, race, and education for both household mobility and college completion outcomes. Interestingly, growth rates appear to matter significantly for nonrural mobility, though not rural mobility. While third grade achievement is significantly correlated with household mobility within both rural and nonrural counties, the R-squared of these models is essentially unchanged when education covariates are included alongside socioeconomic status and race, suggesting that while the correlations between education and mobility are largely meaningful, they pale in comparison to the effect of race. Further, while more variation is explained in nonrural models than rural ones, this can likely be attributed to greater racial variation in nonrural areas than in rural areas.

Turning to college outcomes, educational opportunity seems to display significant correlations, as one might expect. In both rural and nonrural areas, the effect of both black and socioeconomic covariates disappear once achievement scores and growth rates are added. This is heartening, suggesting that an individual’s chances at college completion are likely owed to their achievement and growth more than their demographic status. As in Table 5, the models explain less variation in college completion than mobility outcomes.

Table 7 highlights the results of a fixed effects regression on gender and educational opportunity. In this table I add individual mobility outcomes, given the significant gender differences between household and individual mobility highlighted in Table 1. Also, given the insignificance of some coefficients, I calculate p-values to determine whether including education variables significantly adds explanatory power to the model. Generally, I find little predictive effects of education when examining gender differences, particularly within rural
counties, though there is a significant association between third grade achievement and rural individual mobility. This potentially suggests that females can improve life outcomes by increasing third grade achievement. On the other hand, female household and individual mobility are actually lower in nonrural counties with higher growth rates; however, education appears to account for little, if any, variation in college outcomes or upward mobility once gender is included.

[Table 7 here]

Discussion

This project attempts to understand the factors most strongly correlated to rural upward mobility, particularly the significance of educational opportunity, by utilizing the latest data from the Opportunity Insights and Stanford Education Data Archive databases. Four primary conclusions can be drawn from this analysis.

First, there are significant differences between levels of upward mobility and four-year college completion between rural and nonrural areas, though the question of where mobility is higher depends on the specific subgroup and outcome in question. While rural overall household mobility is higher, this is largely driven by higher mobility for rural males than nonrural ones, as for nearly other subgroup, mobility is higher in nonrural areas. Further, average college completion is generally greater in nonrural counties.

Second, educational opportunity significantly predicts mobility and college outcomes. Even when controlling for all covariate buckets, growth rates retain a significant correlation with
rural household mobility. I also find that growth rates matter more for nonrural mobility than rural, particularly when examining within-county variation and controlling for gender.

Third, while educational opportunity is significantly predictive, it is less so than other covariates. In the full covariate model, the effect of growth rates is substantially smaller than that of demographic variables such as fraction black and fraction with single mothers. This is also confirmed by the results of fixed effects regressions, which finds that race and socioeconomic status explain the majority of variation in mobility, and that adding measures of achievement and growth rates leaves R-squared values virtually unchanged.

Fourth, the type of rural county matters when examining mobility. While some regional dimensions diminish in effect when controlling for other covariates, even among rural counties, those which are less remote have negative correlations to household mobility outcomes. Certain geographic divisions also significantly predict mobility; Southern divisions have a strong negative association, while Midwestern states in the West North Central division correlate positively. I also find farming and mining economies to correlate positively with mobility, while manufacturing-dependent counties display a significant negative relationship.

This research faces several limitations. First, it is unable to comment on causality between these covariates and the outcomes in question. It is also unable to gauge the effects of growing up in a rural area in the wake of events such as the 2008 economic crisis, from which rural areas were slower to recover than nonrural ones. Finally, by analyzing mean values across counties, I am unable to examine individual-level mobility, thereby obfuscating edge cases which may provide useful insight into the experience of overcoming hardship in rural America.
In future drafts, I will continue to analyze the relationship between educational opportunity and rural upward mobility and will comment on tangible policy solutions. While rural America often falls to the periphery of academic research, my hope is that this project will not only highlight the benefits and challenges of a rural upbringing, but will also draw attention to the need to develop a more concerted effort toward ensuring equality of opportunity across the rural-urban divide.
## Table 1: Outcome Descriptive Statistics

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Household mobility</th>
<th>Individual mobility</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Nonrural</td>
<td>Rural</td>
<td>Nonrural</td>
</tr>
<tr>
<td>All (Economically Disadvantaged)</td>
<td>0.406</td>
<td>1089</td>
<td>0.414</td>
</tr>
<tr>
<td>Male</td>
<td>0.390</td>
<td>1089</td>
<td>0.406</td>
</tr>
<tr>
<td>Female</td>
<td>0.422</td>
<td>1089</td>
<td>0.422</td>
</tr>
<tr>
<td>Black</td>
<td>0.326</td>
<td>959</td>
<td>0.318</td>
</tr>
<tr>
<td>Asian</td>
<td>0.567</td>
<td>782</td>
<td>0.515</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.430</td>
<td>1009</td>
<td>0.429</td>
</tr>
<tr>
<td>Native American</td>
<td>0.337</td>
<td>637</td>
<td>0.318</td>
</tr>
<tr>
<td>White</td>
<td>0.450</td>
<td>1089</td>
<td>0.448</td>
</tr>
<tr>
<td>Non-economically Disadvantaged</td>
<td>0.576</td>
<td>1089</td>
<td>0.590</td>
</tr>
</tbody>
</table>

Means weighted by number of children below 50th percentile in each subgroup, except non-economically disadvantaged, which is weighted by number of children in the county as of 2000.
Table 2: Covariate Descriptive Statistics

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Nonrural</th>
<th>Rural</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std Dev</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>Income segregation</td>
<td>0.68</td>
<td>1.09</td>
<td>1088</td>
</tr>
<tr>
<td>Fraction above poverty line</td>
<td>0.40</td>
<td>0.81</td>
<td>1088</td>
</tr>
<tr>
<td>Fraction black</td>
<td>0.12</td>
<td>0.94</td>
<td>1088</td>
</tr>
<tr>
<td>Household income per capita</td>
<td>0.53</td>
<td>1.15</td>
<td>1088</td>
</tr>
<tr>
<td>Fraction with single mothers</td>
<td>0.09</td>
<td>0.97</td>
<td>1088</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Std Dev</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>School expenditures per pupil</td>
<td>-0.07</td>
<td>0.61</td>
<td>1065</td>
</tr>
<tr>
<td>Number colleges per capita</td>
<td>-0.07</td>
<td>0.44</td>
<td>1088</td>
</tr>
<tr>
<td>Growth rates</td>
<td>0.01</td>
<td>0.91</td>
<td>1086</td>
</tr>
<tr>
<td>Grade 3 achievement</td>
<td>0.20</td>
<td>0.94</td>
<td>1086</td>
</tr>
<tr>
<td>Grade 8 achievement</td>
<td>0.20</td>
<td>0.98</td>
<td>1086</td>
</tr>
<tr>
<td>Economy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Std Dev</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>Fraction with commute &lt; 15 mins</td>
<td>-0.71</td>
<td>0.65</td>
<td>1088</td>
</tr>
<tr>
<td>Gini coefficient, bottom 99%</td>
<td>0.19</td>
<td>1.08</td>
<td>1083</td>
</tr>
<tr>
<td>Teenage labor force participation</td>
<td>-0.16</td>
<td>0.77</td>
<td>1088</td>
</tr>
<tr>
<td>In-migration</td>
<td>0.73</td>
<td>1.10</td>
<td>1088</td>
</tr>
<tr>
<td>Out-migration</td>
<td>0.71</td>
<td>1.11</td>
<td>1088</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.26</td>
<td>0.77</td>
<td>1088</td>
</tr>
<tr>
<td>Institutions factors</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>Std Dev</td>
<td>N</td>
<td>Mean</td>
</tr>
<tr>
<td>Local tax rate per capita</td>
<td>0.02</td>
<td>0.46</td>
<td>1088</td>
</tr>
<tr>
<td>Penn State social capital index</td>
<td>-0.27</td>
<td>0.82</td>
<td>1084</td>
</tr>
<tr>
<td>Fraction religious</td>
<td>-0.25</td>
<td>0.82</td>
<td>1088</td>
</tr>
<tr>
<td>Violent crime rate</td>
<td>0.14</td>
<td>0.94</td>
<td>1018</td>
</tr>
<tr>
<td>Demographics</td>
<td>KFR</td>
<td>College</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>--------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Income segregation</td>
<td>-0.00672 ***</td>
<td>0.0229 ***</td>
<td>0.0343 ***</td>
</tr>
<tr>
<td>Fraction above poverty line</td>
<td>0.0248 ***</td>
<td>-0.00159 ***</td>
<td>0.0117 ***</td>
</tr>
<tr>
<td>Fraction black</td>
<td>-0.0314 ***</td>
<td>0.00612 ***</td>
<td>-0.0140 ***</td>
</tr>
<tr>
<td>Household income per capita</td>
<td>0.0251 ***</td>
<td>-0.00464 *</td>
<td>0.0227 ***</td>
</tr>
<tr>
<td>Fraction with single mothers</td>
<td>-0.0439 ***</td>
<td>0.0187 ***</td>
<td>-0.0238 ***</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Per-pupil expenditures</td>
<td>0.00323 **</td>
<td>0.0219 ***</td>
<td>0.00863 ***</td>
</tr>
<tr>
<td>Number colleges per capita</td>
<td>-0.00295 **</td>
<td>-0.0006</td>
<td>0.0027</td>
</tr>
<tr>
<td>Growth rates</td>
<td>0.0055 ***</td>
<td>0.0085 ***</td>
<td>0.0023</td>
</tr>
<tr>
<td>Grade 3 achievement</td>
<td>0.0269 ***</td>
<td>-0.0077 ***</td>
<td>0.0236 ***</td>
</tr>
<tr>
<td>Grade 8 achievement</td>
<td>0.0319 ***</td>
<td>-0.0070 ***</td>
<td>0.0260 ***</td>
</tr>
<tr>
<td>Economics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction with commute &lt; 15 mins</td>
<td>0.0276 ***</td>
<td>-0.0193 ***</td>
<td>0.0293 ***</td>
</tr>
<tr>
<td>Gini coefficient, bottom 99%</td>
<td>-0.0300 ***</td>
<td>0.0206 ***</td>
<td>-0.0133 ***</td>
</tr>
<tr>
<td>Teenage labor force participation</td>
<td>0.0349 ***</td>
<td>-0.00621 **</td>
<td>0.0308 ***</td>
</tr>
<tr>
<td>In-migration</td>
<td>-0.0130 ***</td>
<td>0.0147 ***</td>
<td>-0.0207 ***</td>
</tr>
<tr>
<td>Out-migration</td>
<td>-0.0121 ***</td>
<td>0.0149 ***</td>
<td>-0.0174 ***</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.0284 ***</td>
<td>0.0167 ***</td>
<td>-0.0256 ***</td>
</tr>
<tr>
<td>Institutions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local tax rate per capita</td>
<td>0.00345 ***</td>
<td>0.0421 ***</td>
<td>0.0032 *</td>
</tr>
<tr>
<td>Social capital index</td>
<td>0.0350 ***</td>
<td>-0.0134 ***</td>
<td>0.0367 ***</td>
</tr>
<tr>
<td>Fraction religious</td>
<td>0.0246 ***</td>
<td>-0.0063 **</td>
<td>0.0178 ***</td>
</tr>
<tr>
<td>Violent crime rate</td>
<td>-0.0230 ***</td>
<td>0.0150 ***</td>
<td>-0.0149 ***</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001
Table 4: Multivariate Regressions by Regional Dimension

<table>
<thead>
<tr>
<th>RUCC 4</th>
<th>RUCC 5</th>
<th>RUCC 6</th>
<th>RUCC 7</th>
<th>RUCC 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.071</td>
<td>-0.055</td>
<td>-0.059</td>
<td>-0.036</td>
<td>-0.032</td>
</tr>
<tr>
<td>0.005</td>
<td>0.007</td>
<td>0.004</td>
<td>0.004</td>
<td>0.005</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Divisions</th>
<th>Econ subtypes</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle Atlantic</td>
<td>0.013</td>
<td>0.021</td>
</tr>
<tr>
<td>East North Central</td>
<td>-0.003</td>
<td>0.009</td>
</tr>
<tr>
<td>West North Central</td>
<td>0.068</td>
<td>0.055</td>
</tr>
<tr>
<td>South Atlantic</td>
<td>-0.055</td>
<td>-0.058</td>
</tr>
<tr>
<td>East South Central</td>
<td>-0.05</td>
<td>-0.051</td>
</tr>
<tr>
<td>West South Central</td>
<td>0.000</td>
<td>-0.007</td>
</tr>
<tr>
<td>Mountain</td>
<td>0.036</td>
<td>0.025</td>
</tr>
<tr>
<td>Pacific</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>Farming</td>
<td>0.054</td>
<td>0.029</td>
</tr>
<tr>
<td>Mining</td>
<td>0.014</td>
<td>0.026</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>-0.031</td>
<td>-0.006</td>
</tr>
<tr>
<td>Government</td>
<td>-0.018</td>
<td>-0.003</td>
</tr>
<tr>
<td>Services</td>
<td>-0.002</td>
<td>0.013</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.479</td>
<td>0.439</td>
</tr>
<tr>
<td>N</td>
<td>1797</td>
<td>1797</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.13</td>
<td>0.50</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001
Table 5: Multivariate Regressions by Covariate Bucket

<table>
<thead>
<tr>
<th>Demographics</th>
<th>Education</th>
<th>Economics</th>
<th>Institutions</th>
<th>All</th>
<th>All + dimensions</th>
<th>College</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Income</td>
<td>-0.006 **</td>
<td>-0.001</td>
<td>0.001</td>
<td>-0.001</td>
<td></td>
<td>0.004</td>
</tr>
<tr>
<td>segregation</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fraction above</td>
<td>-0.009 ***</td>
<td>-0.014 ***</td>
<td>-0.005 ***</td>
<td>-0.024 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>poverty line</td>
<td>0.002</td>
<td>0.001</td>
<td>0.002</td>
<td></td>
<td></td>
<td>0.003</td>
</tr>
<tr>
<td>Fraction black</td>
<td>0.004 *</td>
<td>0.012 ***</td>
<td>0.014 ***</td>
<td>0.015 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Household</td>
<td>0.017 ***</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
<td></td>
<td>0.011 **</td>
</tr>
<tr>
<td>income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>per capita</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
<td></td>
<td>0.004</td>
</tr>
<tr>
<td>Fraction with</td>
<td>-0.05 ***</td>
<td>-0.041 ***</td>
<td>-0.034 ***</td>
<td>-0.018 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>single mothers</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
<td></td>
<td>0.004</td>
</tr>
<tr>
<td>Per-pupil</td>
<td>0.012 ***</td>
<td>0.003 **</td>
<td>0.003 **</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>expenditures</td>
<td>0.002</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>-0.002</td>
<td>-0.001</td>
<td>-0.001</td>
<td></td>
<td></td>
<td>0.003 *</td>
</tr>
<tr>
<td>colleges</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>per capita</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td>0.001</td>
</tr>
<tr>
<td>Growth rates</td>
<td>0.018 ***</td>
<td>0.001</td>
<td>0.002 **</td>
<td>0.005 **</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 3</td>
<td>0.033 ***</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.013 ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>achievement</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td>0.003</td>
</tr>
<tr>
<td>Fraction with</td>
<td>0.012 ***</td>
<td>0.011 ***</td>
<td>0.006 ***</td>
<td>0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>commute &lt; 15</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td>0.003</td>
</tr>
<tr>
<td>mins</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini coefficient,</td>
<td>-0.018 ***</td>
<td>-0.007 ***</td>
<td>-0.004 ***</td>
<td>-0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>bottom 99%</td>
<td>0.001</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td>0.002</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Teenage labor force</td>
<td>0.022</td>
<td>***</td>
<td>0.012</td>
<td>***</td>
<td>0.008</td>
<td>***</td>
</tr>
<tr>
<td>participation</td>
<td>0.001</td>
<td></td>
<td>0.001</td>
<td></td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>In-migration</td>
<td>-0.012</td>
<td>***</td>
<td>-0.004</td>
<td>*</td>
<td>-0.005</td>
<td>**</td>
</tr>
<tr>
<td>Out-migration</td>
<td>-0.003</td>
<td></td>
<td>0.001</td>
<td></td>
<td>-0.001</td>
<td></td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>-0.011</td>
<td>***</td>
<td>-0.003</td>
<td>***</td>
<td>-0.003</td>
<td>**</td>
</tr>
<tr>
<td>Local tax rate per capita</td>
<td>0.028</td>
<td>***</td>
<td>0.006</td>
<td>**</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>Social capital index</td>
<td>0.027</td>
<td>***</td>
<td>0.013</td>
<td>***</td>
<td>0.009</td>
<td>***</td>
</tr>
<tr>
<td>Fraction religious</td>
<td>0.015</td>
<td>***</td>
<td>0.004</td>
<td>***</td>
<td>0.005</td>
<td>***</td>
</tr>
<tr>
<td>Violent crime rate</td>
<td>-0.015</td>
<td>***</td>
<td>-0.001</td>
<td>-0.003</td>
<td>**</td>
<td>0</td>
</tr>
<tr>
<td>RUCC 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.008</td>
<td>*</td>
</tr>
<tr>
<td>RUCC 5</td>
<td>-0.011</td>
<td>**</td>
<td>0.005</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUCC 6</td>
<td>-0.005</td>
<td>*</td>
<td>-0.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUCC 7</td>
<td>-0.008</td>
<td>***</td>
<td>-0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RUCC 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.002</td>
<td>0.005</td>
</tr>
<tr>
<td>Middle Atlantic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.018</td>
<td>**</td>
</tr>
<tr>
<td>Region</td>
<td>Intercept</td>
<td>0.005</td>
<td>-0.044</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------</td>
<td>-----------</td>
<td>--------</td>
<td>--------</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>East North Central</td>
<td>0.006</td>
<td>0.012</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West North Central</td>
<td>0.022</td>
<td>***</td>
<td>-0.013</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Atlantic</td>
<td>0.006</td>
<td>0.013</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>East South Central</td>
<td>-0.018</td>
<td>**</td>
<td>-0.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>West South Central</td>
<td>0.005</td>
<td>-0.023</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mountain</td>
<td>0.017</td>
<td>**</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pacific</td>
<td>0.028</td>
<td>***</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farming</td>
<td>0.005</td>
<td>*</td>
<td>0.013</td>
<td>**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mining</td>
<td>0.018</td>
<td>***</td>
<td>-0.026</td>
<td>***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manufacturing</td>
<td>-0.009</td>
<td>***</td>
<td>-0.008</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Government</td>
<td>0.002</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Services</td>
<td>0.002</td>
<td>0.007</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**N** 1797  
**R-squared** 0.54 0.31 0.62 0.57 0.77 0.82 0.386

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors shown in italics.
Table 6: Fixed Effects Analysis by Race & Socioeconomic Status

<table>
<thead>
<tr>
<th></th>
<th>Rural</th>
<th>Nonrural</th>
<th>Rural</th>
<th>Nonrural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>-0.108***</td>
<td>-0.061***</td>
<td>-0.105***</td>
<td>-0.071***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.005)</td>
<td>(0.001)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-0.041***</td>
<td>-0.019***</td>
<td>-0.024***</td>
<td>-0.035***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.002)</td>
</tr>
<tr>
<td>SES</td>
<td>0.02***</td>
<td>0.015***</td>
<td>0.015***</td>
<td>0.01***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Grade 3 achievement</td>
<td>0.008***</td>
<td>0.007***</td>
<td>0.029***</td>
<td>0.039***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.006)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Growth</td>
<td>0.001</td>
<td>0.004*</td>
<td>0.018**</td>
<td>0.026***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.006)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.461***</td>
<td>0.456***</td>
<td>0.459***</td>
<td>0.445***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>N</td>
<td>4200</td>
<td>4200</td>
<td>4200</td>
<td>3005</td>
</tr>
<tr>
<td>Adj. R-squared</td>
<td>0.62</td>
<td>0.64</td>
<td>0.64</td>
<td>0.73</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors shown in parentheses.
Table 7: Fixed Effects Analysis by Gender

<table>
<thead>
<tr>
<th></th>
<th>KFR</th>
<th>College</th>
<th>KIR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rural</td>
<td>Nonrural</td>
<td>Rural</td>
</tr>
<tr>
<td>Female</td>
<td>0.015***</td>
<td>0.015***</td>
<td>0.023***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Grade 3</td>
<td>0.002</td>
<td>0.002</td>
<td>-0.003</td>
</tr>
<tr>
<td>achievement</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>Growth</td>
<td>0.003</td>
<td>-0.007***</td>
<td>-0.008</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.432***</td>
<td>0.432***</td>
<td>0.401***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>N</td>
<td>4063</td>
<td>4063</td>
<td>2172</td>
</tr>
<tr>
<td>Adj. R-squared</td>
<td>0.19</td>
<td>0.19</td>
<td>0.54</td>
</tr>
<tr>
<td>P-value</td>
<td>0.277</td>
<td>0.000</td>
<td>0.362</td>
</tr>
</tbody>
</table>

** p < 0.05, ** p < 0.01, *** p < 0.001. Standard errors shown in parentheses.
 References


## Appendix

### Table A1: Continuous Covariate Data Sources

<table>
<thead>
<tr>
<th>Category</th>
<th>Data Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographics</strong></td>
<td></td>
</tr>
<tr>
<td>Income segregation</td>
<td>Opportunity Atlas – 2000 Census SF3 Sample Data Table P052</td>
</tr>
<tr>
<td>Fraction above poverty line</td>
<td>Opportunity Atlas – 2000 Census SF3 Sample Data Table P087</td>
</tr>
<tr>
<td>Fraction black</td>
<td>Opportunity Atlas – 2000 Census SF1 100% Data Table P008</td>
</tr>
<tr>
<td>Household income per capita</td>
<td>Opportunity Atlas – 2000 Census SF3 Sample Data Table P054</td>
</tr>
<tr>
<td>Fraction with single mothers</td>
<td>Opportunity Atlas – 2000 Census SF3 Sample Data Table P015</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>Number colleges per capita</td>
<td>Opportunity Atlas – IPEDS 2000</td>
</tr>
<tr>
<td>Growth rates</td>
<td>SEDA v3.0</td>
</tr>
<tr>
<td>Grade 3 achievement</td>
<td>SEDA v3.0</td>
</tr>
<tr>
<td>Grade 8 achievement</td>
<td>SEDA v3.0</td>
</tr>
<tr>
<td><strong>Economics</strong></td>
<td></td>
</tr>
<tr>
<td>Fraction with commute &lt; 15 mins</td>
<td>Opportunity Atlas – 2000 Census SF3 Sample Data Table P031</td>
</tr>
<tr>
<td>Teenage labor force participation</td>
<td>Opportunity Atlas – Tax Records, Extended Sample</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>Opportunity Atlas – Local Area Unemployment Statistics</td>
</tr>
<tr>
<td><strong>Institutions</strong></td>
<td></td>
</tr>
<tr>
<td>Fraction religious</td>
<td>Opportunity Atlas – Association of Religion Data Archives</td>
</tr>
<tr>
<td>Violent crime rate</td>
<td>Opportunity Atlas – Uniform Crime Reports</td>
</tr>
</tbody>
</table>

Source: Chetty & Hendren 2017
Table A2: Definitions and Distribution of U.S. Census Division

<table>
<thead>
<tr>
<th>Region</th>
<th>Division</th>
<th>States</th>
<th>Nonrural</th>
<th>Rural</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northeast</td>
<td>New England</td>
<td>CT, ME, MA, NH, RI, VT</td>
<td>34</td>
<td>33</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>Middle Atlantic</td>
<td>NJ, NY, PA</td>
<td>89</td>
<td>61</td>
<td>150</td>
</tr>
<tr>
<td>Midwest</td>
<td>East North Central</td>
<td>IL, IN, MI, OH, WI</td>
<td>173</td>
<td>264</td>
<td>437</td>
</tr>
<tr>
<td></td>
<td>West North Central</td>
<td>IA, KS, MN, MO, NE, ND, SD</td>
<td>112</td>
<td>505</td>
<td>617</td>
</tr>
<tr>
<td>South</td>
<td>South Atlantic</td>
<td>DE, DC, FL, GA, MD, NC, SC, VA, WV</td>
<td>289</td>
<td>299</td>
<td>588</td>
</tr>
<tr>
<td></td>
<td>East South Central</td>
<td>AL, KY, MS, TN</td>
<td>118</td>
<td>246</td>
<td>364</td>
</tr>
<tr>
<td></td>
<td>West South Central</td>
<td>AR, LA, OK, TX</td>
<td>143</td>
<td>325</td>
<td>468</td>
</tr>
<tr>
<td>West</td>
<td>Mountain</td>
<td>AZ, CO, ID, MT, NV, NM, UT, WY</td>
<td>62</td>
<td>218</td>
<td>280</td>
</tr>
<tr>
<td></td>
<td>Pacific</td>
<td>AK, CA, HI, OR, WA</td>
<td>69</td>
<td>96</td>
<td>165</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Total</strong></td>
<td><strong>1,089</strong></td>
<td><strong>2,047</strong></td>
<td><strong>3,136</strong></td>
</tr>
</tbody>
</table>

Table A3: Definitions and Distribution of Rural-Urban Continuum Codes

<table>
<thead>
<tr>
<th>RUCC</th>
<th>Description</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonrural</td>
<td>County in metro area with 1 million population or more</td>
<td>414</td>
</tr>
<tr>
<td></td>
<td>County in metro area of 250,000 to 1 million population</td>
<td>325</td>
</tr>
<tr>
<td></td>
<td>County in metro area of fewer than 250,000 population</td>
<td>350</td>
</tr>
<tr>
<td>Rural</td>
<td>Nonmetro county with urban population of 20,000 or more, adjacent to a metro area</td>
<td>218</td>
</tr>
<tr>
<td></td>
<td>Nonmetro county with urban population of 20,000 or more, not adjacent to a metro area</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>Nonmetro county with urban population of 2,500-19,999, adjacent to a metro area</td>
<td>608</td>
</tr>
<tr>
<td></td>
<td>Nonmetro county with urban population of 2,500-19,999, not adjacent to a metro area</td>
<td>451</td>
</tr>
<tr>
<td></td>
<td>Nonmetro county completely rural or less than 2,500 urban population, adjacent to a metro area</td>
<td>235</td>
</tr>
<tr>
<td></td>
<td>Nonmetro county completely rural or less than 2,500 urban population, not adjacent to a metro area</td>
<td>430</td>
</tr>
<tr>
<td></td>
<td><strong>Total</strong></td>
<td><strong>3,136</strong></td>
</tr>
</tbody>
</table>

Table A4: Distribution of Economic Subtypes

<table>
<thead>
<tr>
<th>Description</th>
<th>Nonrural</th>
<th>Rural</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonspecialized</td>
<td>333</td>
<td>616</td>
<td>949</td>
</tr>
<tr>
<td>Farming-dependent</td>
<td>37</td>
<td>400</td>
<td>437</td>
</tr>
<tr>
<td>Mining-dependent</td>
<td>15</td>
<td>112</td>
<td>127</td>
</tr>
<tr>
<td>Manufacturing-dependent</td>
<td>319</td>
<td>584</td>
<td>903</td>
</tr>
<tr>
<td>Federal/state government-dependent</td>
<td>159</td>
<td>222</td>
<td>381</td>
</tr>
<tr>
<td>Services-dependent</td>
<td>226</td>
<td>113</td>
<td>339</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>1,089</strong></td>
<td><strong>2,047</strong></td>
<td><strong>3,136</strong></td>
</tr>
</tbody>
</table>