

Ascending or Not: School Profiles and Intergenerational Mobility

MACS 30122 Winter 2025

Ascending Team

Carrie Huang, Jeanette Wu, Kunjian Li, Shirley Zhang

University of Chicago

March 5th, 2025

1 Introduction

2 Data Collection

3 Data Cleaning and Wrangling

4 Data Analysis and Visualization

5 Main Takeaways

Research Questions and Significance

Research Questions

- Which factors/characteristics of school profile contribute to the prediction of intergenerational mobility?

Research Significance

- Address a critical social issue by exploring how educational quality affects economic mobility.
- Inform policies aimed at reducing educational disparities
- Foster equitable opportunities for upward mobility
- Contribute to broader discussions about education and social justice.

Hypothesis

- The racial composition of a school is expected to be a significant predictor of mobility (Jang and Reardon 2019, Matheny et.al 2023).
- Schools with a higher proportion of students from lower-income families are likely to exhibit lower rates of social mobility (Jang and Reardon 2019, Scherger and Savage 2010, Beckett 2024).
- Schools with greater academic resources such as lower student-to-teacher ratios and access to SAT prep courses are expected to foster higher upward mobility (Blizard 2020).
- Schools with more positive reviews are anticipated to be associated with greater upward mobility (Beckett 2024).

- 1 Introduction
- 2 Data Collection**
- 3 Data Cleaning and Wrangling
- 4 Data Analysis and Visualization
- 5 Main Takeaways

School Profiles

School Statistics

Academics and Student Demographics, with Academics covering areas such as College Preparation, Advanced Courses, and Test Scores and Student Demographics covering Students Ethnicity, Family Income and so on.

School Reviews

We collected reviews on schools from

- Parents/Guardians
- School Staffs
- Alumni
- Family Members
- and so on...

School Profiles (Cont'd)

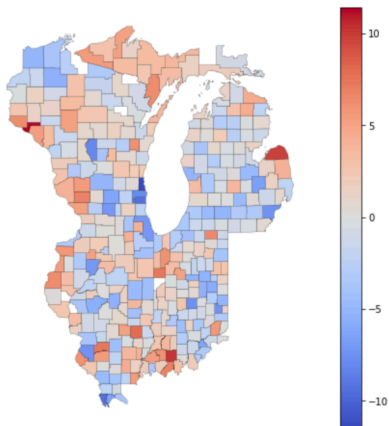
Challenges

- We met a problem with the dynamic websites that when scrapping the `Greatschools.org`, *Selenium* slows down the process.
- The government data doesn't provide sufficient granularity to answer the questions.

Solutions

- We collected schools data and reviews utilizing the `Greatschools.org` API.
- We use the census and government data to supplement the school data collected from `Greatschools.org`.

Intergenerational Mobility



Counties with the highest upward mobility, shown in dark red, are located in eastern Wisconsin, northern Michigan, southern Illinois, and western Indiana.

Source

Chetty et al. (2014)

Figure 1: Upward Mobility with State FE

- 1 Introduction
- 2 Data Collection
- 3 Data Cleaning and Wrangling**
- 4 Data Analysis and Visualization
- 5 Main Takeaways

Four-layer Structure

Data Transformation

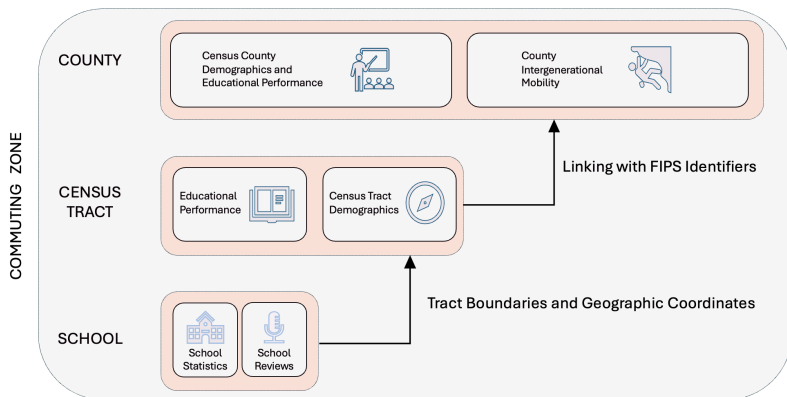


Figure 2: Illustration of the Four-Layer Structure

- 1 Introduction
- 2 Data Collection
- 3 Data Cleaning and Wrangling
- 4 Data Analysis and Visualization**
 - Sentiment Analysis
 - Causal Inference
- 5 Main Takeaways

- 1 Introduction
- 2 Data Collection
- 3 Data Cleaning and Wrangling
- 4 Data Analysis and Visualization**
 - Sentiment Analysis
 - Causal Inference
- 5 Main Takeaways

Sentiment Analysis

Example 1: Highly Positive



Parent / Guardian

★★★★★ July 27, 2023

My oldest son graduated from Kenwood and he often spoke about how supportive the teachers were towards him and how they encouraged him to excel and do his best. He loved that the administration was always supportive and easy to talk to as well. He knew not only the teachers but also the admin pushed him and expected great things from him. We enjoyed the experience so much that now my other son is looking forward to beginning his high school year there as a freshmen this fall and becoming a bronco. My family highly recommends Kenwood Academy.

Example 2: Highly Negative



Other (school staff, family member, recent alum, etc)

★☆☆☆☆ October 04, 2014

This school is prison, a place which you'll find yourself dreading arrival upon every morning. I feel only absolute sympathy for any future generations that will be forced to endure it.

- 1 Introduction
- 2 Data Collection
- 3 Data Cleaning and Wrangling
- 4 Data Analysis and Visualization**
 - Sentiment Analysis
 - Causal Inference
- 5 Main Takeaways

Feature Engineering

First Stage: Feature Selection via Correlation Matrix

- Remove highly correlated features (e.g., Math and English grades)
- Reduced from 173 features to 28
- Selected features come from: Rating, Demographics (Subgroup, Gender, Ethnicity), Teachers & Staff, College Preparation, Courses & Programs

Second Stage: Lasso-based Feature Selection

- Further reduce features to 9 for Fixed Effect Regression
- Selected Features:

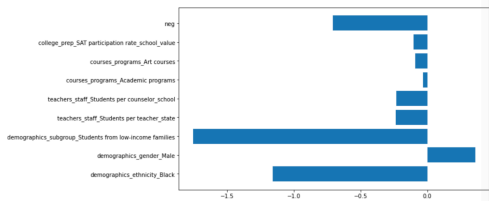


Figure 3: Feature Importance via LASSO

Causal Inference

Table 1: OLS Regression Results at County Level

	Coef.	Std. Err.	t	P> t
Constant	49.4712	3.164	15.637	0.000
Black	-0.1855	0.021	-8.639	0.000
Male	0.1060	0.060	1.777	0.077
Low-income	-0.1359	0.015	-8.878	0.000
Students per teacher	-0.0632	0.019	-3.392	0.001
Negative Comments	-43.9415	11.169	-3.934	0.000
...
State-FE	Yes			

Causal Inference (Cont'd)

Regression Results

Notably, a higher proportion of Black students and students from low-income families are associated with lower mobility, while negative ratings and availability of study resources also play a crucial role. This finding aligns with broader research on systemic barriers to economic mobility faced by marginalized groups.

Robust Analysis

Parallel Dataset at the Commuting Zone Level

Table 2: OLS Regression Results at Commuting Zone Level

	coef	std err	t	P> t
Constant	38.1562	7.841	4.866	0.000
Black	-0.2235	0.059	-3.812	0.000
Male	0.3006	0.157	1.918	0.056
Low-income	-0.1283	0.032	-3.959	0.000
Students per teacher	-0.0427	0.039	-1.084	0.279
Negative Comments	-61.5299	29.083	-2.116	0.035
...
State-FE	Yes			

Robust Analysis (Cont'd)

Random Forest Method at the County Level

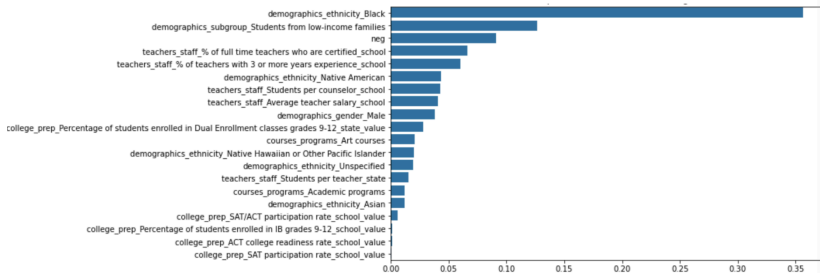


Figure 4: Feature Importance with Random Forest at County Level

- 1 Introduction
- 2 Data Collection
- 3 Data Cleaning and Wrangling
- 4 Data Analysis and Visualization
- 5 Main Takeaways**

Insights and Improvements

Insights

Our project indicates that race is a predominant determinant of intergenerational mobility, followed by the influence of income and educational resources. This effect remains robust across both county-level and commuting zone-level analyses.

Improvements

- We could collect data for multiple periods.
- We could collect national data.
- We may conduct experiments to identify causality.

- [1] R. Chetty, N. Hendren, P. Kline, and E. Saez, “Where is the land of opportunity? the geography of intergenerational mobility in the united states,” *The Quarterly Journal of Economics*, 2014.
- [2] M. Beckett, “How class impacts on education and social mobility,” in *Student Transitions from Further Education to University*, (Switzerland), pp. 77–99, Springer International Publishing, 2024.
- [3] H. Jang and S. F. Reardon, “Uneven progress: Recent trends in academic performance among u.s. school districts,” *AERA Open*, 2019.
- [4] S. Scherger and M. Savage, “Cultural transmission, educational attainment and social mobility,” *The Sociological Review*, 2010.

- [5] K. T. Matheny, M. E. Thompson, C. Townley-Flores, and S. F. Reardon, “Uneven progress: Recent trends in academic performance among u.s. school districts,” *American Educational Research Journal*, vol. 60, no. 3, pp. 447–485, 2023.
- [6] Z. D. Blizard, “Has the allocation of certain teachers impacted student achievement and upward economic mobility? the case of forsyth county, nc elementary schools,” *Education and Urban Society*, vol. 53, no. 7, pp. 778–806, 2021.

Thank You