Cluster Analysis of Economic Mobility in the United States

May 3, 2024 | Shaun Campbell | COSC 6520

1 Introduction

Despite being extolled as the land of opportunity, the United States ranks quite low in terms of economic mobility, coming in at number 27 out of 82 countries ranked by the World Economic Forum Global Social Mobility Index[1]. Research by Raj Chetty et al has confirmed that the United States ranks low compared to other developed nations, and furthermore there is a large disparity in economic mobility across geographical areas of the US[2].

A well-functioning economy should allow its participants a fair chance of achieving success and not systemically hold down the lower classes. In other words, one's ability to climb the economic ladder should be based on merit, decisions, and to some extent luck, rather than being predetermined by the circumstances of their birth. The US's low ranking in upward mobility and situations like steadily increasing wealth inequality[3] make this a pertinent issue in need of being addressed.

This project focuses on analyzing the relationships between demographics, financial well-being, and economic mobility in the United States. Clustering areas of the US based on such characteristics allows for the assessment of geographic trends across the county and helps identify groups of citizens where opportunities for economic advancement are most lacking. Once areas of the country are clustered into groups, relationships between economic mobility, financial well-being, and demographic and economic characteristics can be explored to identify factors that have the most influence on upward mobility in deficient clusters. These relationships can then be leveraged with targeted programs and funding, such as increased funding for public schools, programs that incentivize post-secondary education, and programs to incentivize home ownership. Tailoring programs and funding to address specific deficiencies within certain areas of the US will result in efficient public resource allocation and allow more citizens a chance to achieve the American Dream.

2 Statistical Methodology

First, a model to predict an individual's financial well-being score based on information available through the census was created. The dataset used was the Financial Well-Being Survey ("FWBS"), in which the Consumer Financial Protection Bureau algorithmically determined financial well-being scores based on survey responses (Appendix 1). Linear regression was performed and the resulting model was satisfactory with an adjusted R^2 value of 0.46 and a mean absolute error of 8.01 (Appendix 2).

Next, census data was compiled from the 2019 American Community Survey ("ACS") 5-year estimates. Percentage estimates were retrieved for all census tracts in the US to mirror the indicator variables in the FWBS regression model. Additional ACS data was also retrieved for clustering purposes. The census data was supplemented with financial health metrics data from the Urban Institute. This dataset contains estimates not readily available in the census data – specifically the percent of the area that has $2,000 in savings and the area's estimated median net worth, the latter of which was used as a proxy for amount in savings. A summary of the entire dataset is in Appendix 3.

Agglomerative clustering was then performed to divide census tracts into groups for further analysis. The data was scaled to prevent variables with a higher magnitude from dominating the distance measures, such as median income in dollars versus high school graduation rate as a percentage. A random sample of 5,000 tracts was taken to perform the hierarchal clustering for computational efficiency.

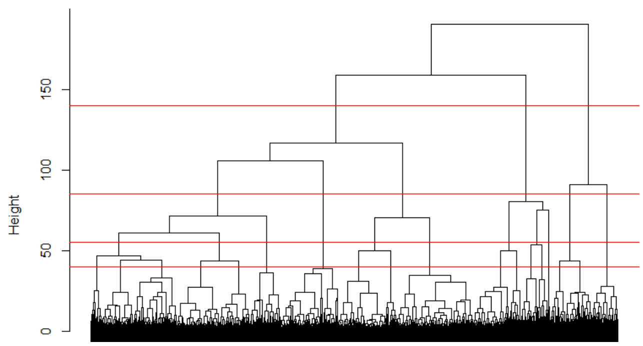


Figure 1: Dendrogram of clustering with evaluated cut points

A dendrogram was generated to visualize the clusters (Figure 1). Various cut points were explored and the mean and median values for each variable in each cluster were evaluated. 18 clusters provided a good level of granularity with meaningful dissimilarity between clusters, but also retained a high number of tracts in each cluster allowing for greater generalization utility.

K-means clustering was then used to cluster all census tracts, and the cluster number was added to the original (non-scaled) data. The coefficients from the FWBS regression model were applied to the ACS tract percentages to create estimated financial well-being scores. Finally, economic mobility data was obtained from The Opportunity Atlas. For each tract, the fraction of people in the top 20% of income who had parents in the bottom 25% of income is used as a measure for economic mobility (very similar to the metric used by Chetty et al[4]). The rationale behind this metric is that being born to lower income parents and becoming a high earner demonstrates upward mobility, thus areas where this occurs more often can be said to be more economically mobile.

Linear regression models were fit to investigate the relationship between financial well-being scores and economic mobility in each cluster (Appendix 4). A linear model was then created to predict economic mobility based on the ACS variables for all tracts (Appendix 5). A subset of predictor variables was used to address multicollinearity issues and the target economic mobility variable was logged to address violation of homoscedasticity of residuals. Models were then repeatedly fit for each cluster so the coefficients could be analyzed (Appendix 6). Finally, historic census data from 2010-2019 was used with the general predictive economic mobility model to investigate time-series trends in Wisconsin census tracts.

2 Results

Clustering all the census tracts in the US allows for an assessment of the geographical areas that are similar (and dissimilar) in terms of economic and demographic characteristics. Figure 2 displays the US with the clusters color-coded. Appendix 7 contains a table of median values for each cluster which can be referenced to assess geographic trends in the US and Appendix 8 contains national-level medians for comparison. For example, cluster 14 (purple) is prominent in the Bible Belt. This cluster is entirely rural, has high poverty (median 22.6%), low median income ($37,976), and very low economic mobility (median 7.8%). The east coast contains many tracts in cluster 1 (red), which has high economic mobility (median 21.29%), low poverty (median 3.5%) and high median income ($114,063).

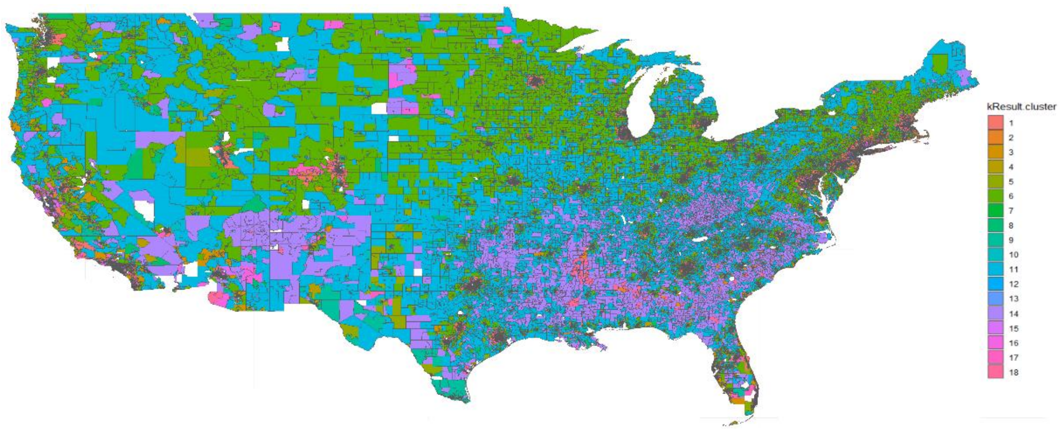


Figure 2: Map of US displaying color-coded clusters

On a state level, the poorest state of Mississippi[5] has a majority of census tracts in clusters 14 and 11, which are both rural and have low economic mobility. In comparison, the richest state of New Jersey[6] has many census tracts in cluster 1 and cluster 15, the latter of which is the wealthiest cluster with a median income of $153,700 and median net worth of $655,833. Wisconsin has an interesting geographic tract makeup with the majority of the state's area being in the rural clusters of 11 and 6, but with high concentrations of the wealthy urban cluster 1 in Dane and Waukesha Counties. Cluster maps of these states are included in Appendix 9. Visualizing the geographic locations of the clustered tracts reveals that there are striking geographic trends in terms of economic and demographic characteristics of the US population.

Linear regression models regressing economic mobility on financial well-being score reveal weak linear relationships within clusters (R^2 values range from 0.0008 to 0.0539 – see Appendix 4). The coefficients for financial well-being score are highest for clusters 15, 1, 3 and the R^2 values are highest as well. These clusters have above-average economic mobility and above-average median incomes. Increasing a tract's financial well-being in these clusters will tend to increase economic mobility at a faster rate than for other clusters, and financial well-being explains more variance in economic mobility. This could be explained by surmising that financial well-being is strongly relational to wealth, which is supported by the FWBS regression coefficients (see Appendix 2). Affluent areas likely have better education, job opportunities, and community programs that lead to better upward mobility. If wealthy areas already have social and economic infrastructure that supports economic advancement, increases in affluence could therefore be expected to have a more direct and immediate effect on upward mobility.

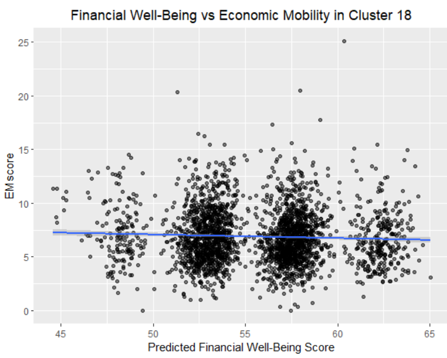


Figure 3: Scatter Plot of Financial Well-Being and Economic Mobility in Cluster 18

A relationship of note is that of cluster 18, shown in Figure 3. The coefficient is negative and explains about zero variance in economic mobility (R^2 = 0.0024). The median economic mobility is the lowest among the 18 clusters at 6.56% and is low on average across all levels of financial well-being. This means that even tracts in this cluster with high financial well-being (and conjecturally high affluence) tend to have low economic mobility. This tract has the highest median poverty rate (38.7%) and unemployment rate (15.8%), and the lowest median income ($25,111) and net worth ($25,694). The fact that this cluster also has the highest median Black population percentage (82.3%) highlights the systemic racial disparities that exist in the US. Since increasing financial well-being will not generally improve economic mobility in these tracts, the tracts should be evaluated for other solutions that can create more equity in terms of economic mobility.

From regressing economic mobility on census characteristics (see Appendix 6), the alterable factors that influence cluster 18's economic mobility positively are high school graduation rate, post-secondary education, and home ownership, visualized in Figure 4. The coefficients indicate that increasing home ownership rate, high school graduation rate, and bachelor degree or higher rate to the national medians would be expected to increase the median predicted economic mobility by around 1.28. While doing so would bring the median economic mobility score of the cluster to 7.84%, this is still well below the national median of 12.34%. This indicates the current social and economic state of the United States fails to provide an equal economic opportunity to all citizens.

There are some overall characteristics of the coefficients that are worth noting. Male population percentage is negatively associated with economic mobility in 13 of the 18 clusters. This may be explained if it's true that men are more likely to pursue manual occupations that have less earning potential. Black population is negatively associated in 15 of the 18 clusters, again demonstrating systemic racial disparities that are present in the US. Home ownership is positively associated in 15 of 18 clusters. This could mean either that home ownership is a stepping stone to climbing the economic ladder or that those who have done so tend to own a home. High school graduation rate is positive in 16 of 18, and similarly post-secondary education in 14 of 18. This makes intuitive sense because increased education is typically associated with higher earning potential, but highlights what could be gained by improving educational opportunities across the county.

The time-series analysis of economic mobility trends in Wisconsin resulted in some interesting findings but is not directly related to the rest of the analysis so is included as Appendix 10.

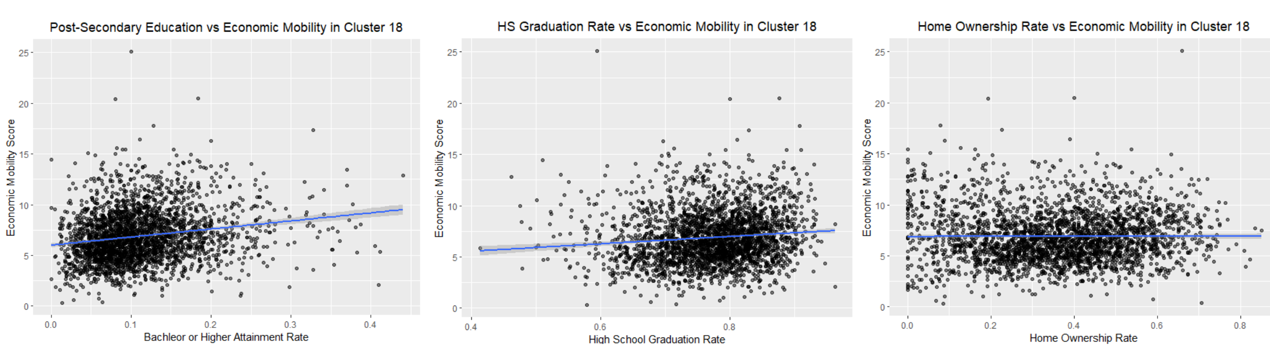


Figure 4: Scatter Plots of Economic Mobility and Influencing Factors in Cluster 18

4 Alternative Approaches

Economic and demographic information could be evaluated at a broad level and generalized to the entire US population. However, this approach would ignore the fact that the US has a huge variance in geographic and demographic structure. Clustering areas in the US into smaller groups allows for generalizations to be made on a more granular level and improves understanding of how economic and demographic factors interact with economic mobility within different groups. Doing so also identifies smaller groups that are deficient in mobility measures, and what factors within those groups should be targeted to maximize efficient use of finite resources. Given innate collinearity in the studied characteristics, prediction of economic mobility is a prime candidate for PCA; however, the coefficients are of interest to evaluate the relationships and thus PCA was not used in this study.

5 Conclusion

The economic mobility measure used in this study should be 20% for all areas of the US because that would indicate perfect equality in economic outcomes independent of birth circumstances. However, the national median of 12.34% indicates this is not the case and in some areas of the US it is far from the truth. The fact that more affluent areas have higher economic mobility offers evidence that living in wealthy areas provides opportunities that may not otherwise be present.

Increasing wealth among mobility-deficient populations would be a difficult thing to achieve, so non-economic factors that contribute to upward mobility should be targeted instead. These factors consistently include high school graduation rates and bachelor degree (or higher) attainment rates. Given finite resources, programs should target these factors within groups that trail most in mobility. Doing so would increase overall economic mobility in the US while also decreasing the variance resulting in a more level field of economic outcomes.

Specifically, it is recommended that areas aligning with characteristics of cluster 18 be prioritized: urban areas with high poverty, low income, low graduation rates, and very low economic mobility. While this may seem intuitive, this study demonstrates just how much these groups lag behind national averages and offers evidence that high school graduation rates and post-secondary education have positive associations with economic mobility within this group. Directing funding and incentives to improve these factors could improve economic mobility for not only a large subset of the US population, but the subset that is most harmed by the current social and economic systems in the county.

6 References

1. The Global Social Mobility Report 2020. (2020). *World Economic Forum*.
2. *Economic mobility: Measuring the American Dream*. HUD USER. (n.d.). https://www.huduser.gov/portal/pdredge/pdr_edge_featd_article_071414.html
3. Monaghan, A. (2014, November 13). *US wealth inequality - top 0.1% worth as much as the bottom 90%*. The Guardian. https://www.theguardian.com/business/2014/nov/13/us-wealth-inequality-top-01-worth-as-much-as-the-bottom-90
4. Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). *Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States*.
5. *Top 10 poorest states in the U.S.* Friends Committee On National Legislation. (2023, November 6). https://www.fcnl.org/updates/2023-11/top-10-poorest-states-us
6. Johnson, S. R. (2023, November 22). *The 10 Richest States in America*. U.S. News. https://www.usnews.com/news/best-states/slideshows/10-wealthiest-states-in-america

Appendix 1

Financial Well-Being Data

Available here: https://ffiec.cfpb.gov/data-publication/three-year-national-loan-level-dataset/2018

Data dictionary: https://ffiec.cfpb.gov/documentation/publications/loan-level-datasets/lar-data-fields

Variable	Type	Description
male	Binary	1 if respondent is male
military	Binary	1 if respondent is in the military
absorb_shock	Binary	1 if respondent believes they could come up with $2,000 in 30 days
unemployed	Binary	1 if respondent is unemployed
snap	Binary	1 if respondent is receiving SNAP benefits
poverty	Binary	1 if respondent is below the poverty level
social_security	Binary	1 if respondent receives social security
health_ins	Binary	1 if respondent has health insurance
Age18_24	Binary	1 if respondent is age 18-24
Age25_34	Binary	1 if respondent is age 25-34
Age35_44	Binary	1 if respondent is age 35-44
Age45_54	Binary	1 if respondent is age 45-54
Age55_74	Binary	1 if respondent is age 55-74
Hispanic	Binary	1 if respondent is Hispanic
Black	Binary	1 if respondent is Black, non-Hispanic
Other	Binary	1 if respondent is Other, non-Hispanic
incomelt20000	Binary	1 if respondent's income is less than $20,000
income20000_29999	Binary	1 if respondent's income is $20,000-$29,999
income30000_39999	Binary	1 if respondent's income is $30,000-$39,999
income40000_49999	Binary	1 if respondent's income is $40,000-$49,999
income50000_74999	Binary	1 if respondent's income is $50,000-$74,999
income75000_99999	Binary	1 if respondent's income is $75,000-$99,999
income100000_149999	Binary	1 if respondent's income is $100,000-$149,999
savings0	Binary	1 if respondent has $0 in savings
savings1_99	Binary	1 if respondent has $1-$99 in savings
savings100_999	Binary	1 if respondent has $100-$999 in savings
savings1000_4999	Binary	1 if respondent has $1,000-$4,999 in savings
savings5000_19999	Binary	1 if respondent has $5,000-$19,999 in savings
savings20000_74999	Binary	1 if respondent has $20,000-$74,999 in savings
savingsUnknown	Binary	1 if respondent did not know savings or declined to answer
hvNA	Binary	1 if respondent does not own a home
hvlt150000	Binary	1 if respondent's home is worth less than $150,000
hv150000_249999	Binary	1 if respondent's home is worth $150,000-$249,999
hv250000_399999	Binary	1 if respondent's home is worth $250,000-$399,999

Appendix 2

Financial Well-Being Predictive Model

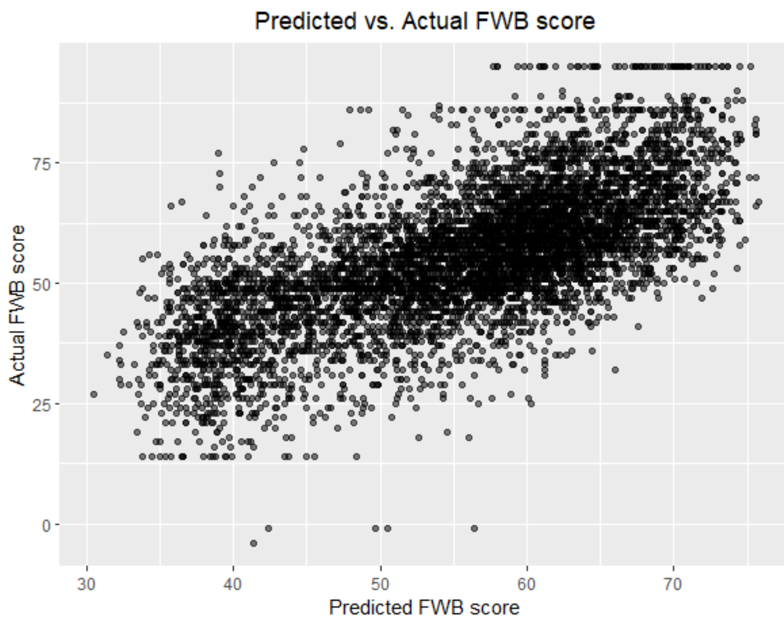
a) Regression Output

Variable	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	62.79	0.90	69.47	0.00
male	0.08	0.27	0.29	0.77
military	1.00	0.35	2.83	0.00
absorb_shock	7.85	0.40	19.39	0.00
unemployed	-2.55	0.70	-3.64	0.00
snap	-1.87	0.52	-3.60	0.00
poverty	1.11	0.74	1.49	0.14
social_security	3.22	0.44	7.28	0.00
health_ins	0.61	0.32	1.90	0.06
Age18_24	-0.02	0.81	-0.03	0.98
Age25_34	-1.87	0.67	-2.78	0.01
Age35_44	-2.46	0.69	-3.59	0.00
Age45_54	-2.96	0.66	-4.49	0.00
Age55_74	-1.26	0.49	-2.58	0.01
Hispanic	1.46	0.41	3.56	0.00
Black	2.25	0.45	5.04	0.00
Other	-1.77	0.60	-2.96	0.00
incomelt20000	-5.04	0.85	-5.92	0.00
income20000_29999	-5.91	0.68	-8.75	0.00
income30000_39999	-5.23	0.62	-8.40	0.00
income40000_49999	-3.48	0.65	-5.31	0.00
income50000_74999	-3.27	0.52	-6.29	0.00
income75000_99999	-2.57	0.52	-4.95	0.00
income100000_149999	-1.87	0.49	-3.83	0.00
savings0	-16.20	0.75	-21.63	0.00
savings1_99	-16.96	0.74	-22.79	0.00
savings100_999	-14.01	0.61	-22.85	0.00
savings1000_4999	-11.95	0.53	-22.35	0.00
savings5000_19999	-7.67	0.50	-15.38	0.00
savings20000_74999	-3.97	0.52	-7.69	0.00
savingsUnknown	-6.66	0.52	-12.93	0.00
hvNA	-2.16	0.52	-4.14	0.00
hvlt150000	-1.34	0.56	-2.41	0.02
hv150000_249999	-1.21	0.53	-2.31	0.02
hv250000_399999	-0.35	0.51	-0.69	0.49

b) Performance metrics

Adjusted R-squared: 0.4579
F-statistic: 159.8
MAPE: 0.1953278
MAE: 8.006852

c) Predicted versus actual values



Appendix 3

Census data set. Bold values indicate that the variables were used in clustering.

Urban Institute data available here: https://datacatalog.urban.org/dataset/financial-health-and-wealth-dashboard-2022

Opportunity Atlas data available here: https://www.opportunityatlas.org/

Variable	Type	Description	Source
Tract	Categorical	Census Tract	Census
Name	Categorical	Name of census tract	Census
male	**Numeric (percentage)**	**percentage of tract that is male**	**Census**
Age18_24	Numeric (percentage)	percentage of tract that is age 18-24	Census
Age25_34	Numeric (percentage)	percentage of tract that is age 25-34	Census
Age35_44	Numeric (percentage)	percentage of tract that is age 34-44	Census
Age45_54	Numeric (percentage)	percentage of tract that is age 45-54	Census
Age55_74	Numeric (percentage)	percentage of tract that is age 55-74	Census
Age_75	Numeric (percentage)	percentage of tract that is age 75+	Census
Hispanic	**Numeric (percentage)**	**percentage of tract that is Hispanic**	**Census**
White	Numeric (percentage)	percentage of tract that is White, Non-Hispanic	Census
Black	**Numeric (percentage)**	**percentage of tract that is Black, Non-Hispanic**	**Census**
Other	**Numeric (percentage)**	**percentage of tract that is Other, Non-Hispanic**	**Census**
MedAge	**Numeric**	**Median age of census tract**	**Census**
military	**Numeric (percentage)**	**percentage of tract that is military**	**Census**
unemployed	**Numeric (percentage)**	**percentage of tract that is unemployed**	**Census**
incomelt20000	Numeric (percentage)	percentage of tract that has income less than $20,000	Census
income20000_29999	Numeric (percentage)	percentage of tract that has income $20,000-29,999	Census
income30000_39999	Numeric (percentage)	percentage of tract that has income $30,000-39,999	Census
income40000_49999	Numeric (percentage)	percentage of tract that has income $40,000-49,999	Census
income50000_74999	Numeric (percentage)	percentage of tract that has income $50,000-$74,999	Census
income75000_99999	Numeric (percentage)	percentage of tract that has income $75,000-$99,999	Census
income100000_149999	Numeric (percentage)	percentage of tract that has income $100,000-$149,999	Census
income150000+	Numeric (percentage)	percentage of tract that has income over $150,000	Census
social_security	**Numeric (percentage)**	**percentage of tract that receives social security**	**Census**
snap	**Numeric (percentage)**	**percentage of tract that receives SNAP benefits**	**Census**
health_ins	**Numeric (percentage)**	**percentage of tract that has health insurance**	**Census**
poverty	**Numeric (percentage)**	**percentage of tract that is below the poverty level**	**Census**
MedIncome	**Numeric**	**Median income of census tract**	**Census**
hvNA	**Numeric (percentage)**	**Percent of tract that does not own a home**	**Census**
hvlt150000	Numeric (percentage)	percentage of tract that has a home value less than $150,000	Census
hv150000_249999	Numeric (percentage)	percentage of tract that has a home value $150,000-249,999	Census
hv250000_399999	Numeric (percentage)	percentage of tract that has a home value $250,000-$399,999	Census
MedHomeVal	**Numeric**	**Median home value of census tract**	**Census**
own_home	**Numeric (percentage)**	**percentage of tract that owns a home**	**Census**
MedRent	**Numeric**	**Median rent of census tract**	**Census**
married	**Numeric (percentage)**	**percentage of tract that is married**	**Census**
foreign_born	**Numeric (percentage)**	**percentage of tract that is foreign born**	**Census**
no_hs	**Numeric (percentage)**	**percentage of tract that did not graduate high school**	**Census**
bachelor_or_higher	**Numeric (percentage)**	**percentage of tract that has a bachelor degree or higher**	**Census**
computer	**Numeric (percentage)**	**percentage of tract that has a computer**	**Census**
internet	**Numeric (percentage)**	**percentage of tract that has internet**	**Census**
rural	**Numeric (percentage)**	**percentage of tract that is rural**	**Census**
absorb_shock	**Numeric (percentage)**	**percentage of tract that could raise $2,000 in 30 days**	**Urban Institute**
credit_score	**Numeric**	**Mean credit score of census tract**	**Urban Institute**
net_worth	**Numeric**	**Mean net worth of census tract**	**Urban Institute**
savings0	Binary	1 if mean net worth of tract is $0	Urban Institute
savings1_99	Binary	1 if mean net worth of tract is $1-$99	Urban Institute
savings100_999	Binary	1 if mean net worth of tract is $100-$999	Urban Institute
savings1000_4999	Binary	1 if mean net worth of tract is $1,000-$4,999	Urban Institute
savings5000_19999	Binary	1 if mean net worth of tract is $5,000-$19,999	Urban Institute
savings20000_74999	Binary	1 if mean net worth of tract is $20,000-$74,999	Urban Institute
savingsUnknown	Binary	1 if mean net worth of tract is unknown	Urban Institute
FWBscore	Numeric	Financial well-being score (predicted) of tract	
EMscore	Numeric	Fraction of people in the top 20% of income who had parents in the bottom 25%	The Opportunity Atlas

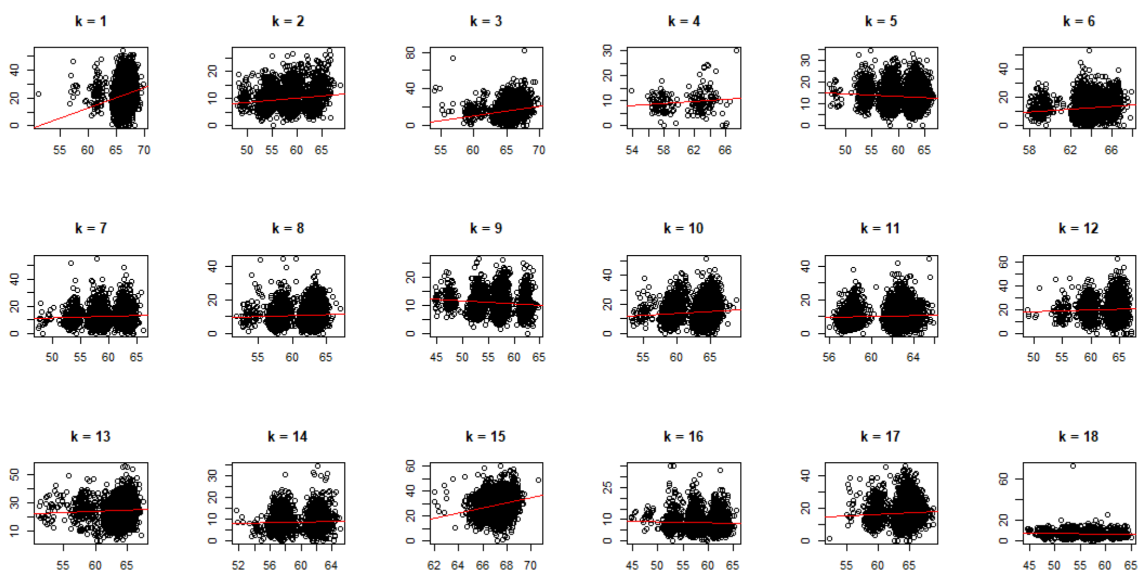
Appendix 4

Linear regression of economic mobility on financial well-being score for each cluster

a) Regression summary statistics for each cluster

cluster	r.squared	intercept	coef	p.coef	f.stat
1	0.0405	-76.0558	1.4777	0.0000	285.7690
2	0.0211	1.2941	0.1480	0.0000	66.6516
3	0.0520	-51.9366	1.0489	0.0000	144.7034
4	0.0246	-4.2734	0.2285	0.0261	5.0232
5	0.0074	20.3468	-0.1151	0.0000	32.8962
6	0.0137	-17.4213	0.4706	0.0000	75.2959
7	0.0046	5.1296	0.1204	0.0000	16.5230
8	0.0032	3.7946	0.1103	0.0000	26.7184
9	0.0110	16.4846	-0.0996	0.0000	31.7726
10	0.0075	-0.6551	0.2405	0.0000	61.3608
11	0.0027	2.2031	0.1289	0.0001	16.1394
12	0.0024	10.6378	0.1458	0.0162	5.7925
13	0.0025	15.6205	0.1422	0.0392	4.2585
14	0.0009	5.6394	0.0482	0.1293	2.3028
15	0.0539	-109.3302	2.0612	0.0000	123.1879
16	0.0008	10.4372	-0.0330	0.0848	2.9724
17	0.0025	5.3391	0.1784	0.0005	12.0985
18	0.0024	9.0063	-0.0377	0.0113	6.4277

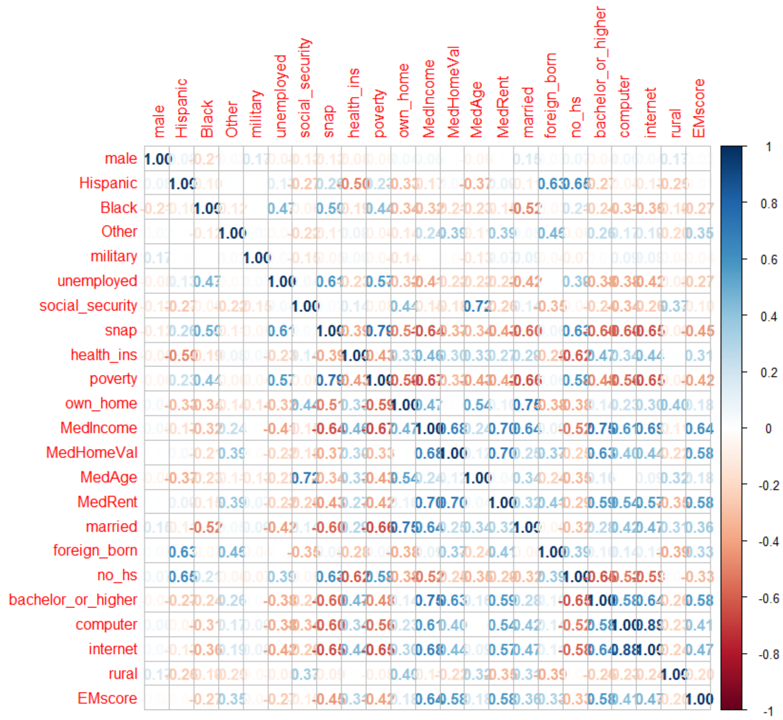
b) Scatter plots of economic mobility (x-axis) versus predicted financial well-being score (y-axis) for each cluster



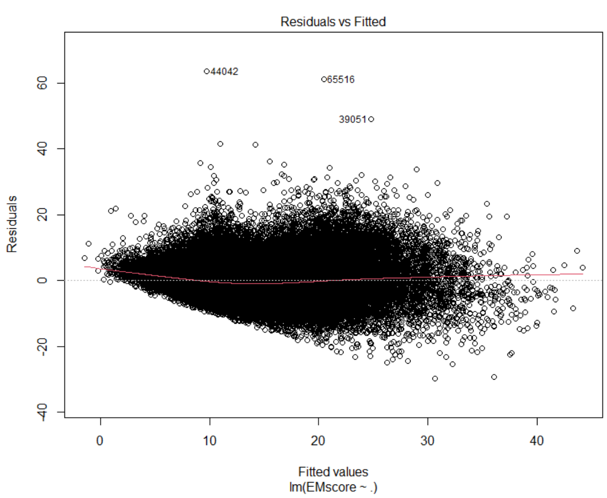
Appendix 5

Linear model to predict economic mobility within all tracts based on ACS variables

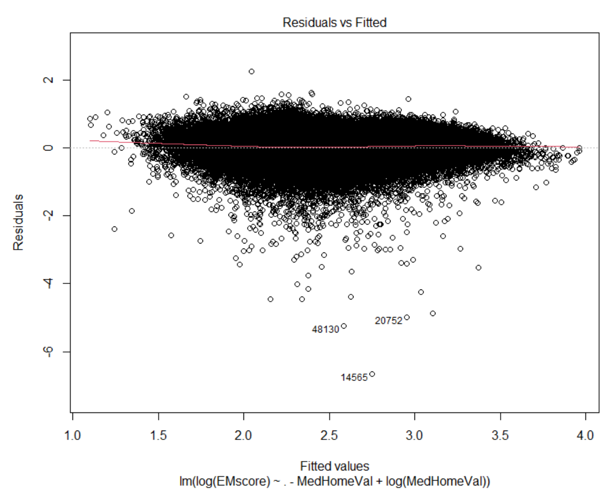
a) Correlation matrix of ACS variables. Several variables were removed due to high prevalence of correlation.



b) Residuals vs fitted values with non-logged and logged target variable



Residuals vs fitted values with non-logged target variable



Residuals vs fitted values with logged target variable

c) Regression output for models with non-logged and logged target variable

	Emscore ~ .	log(Emscore) ~ .
(Intercept)	4.180 (0.000)	0.104 (0.026)
male	-2.720 (0.000)	-0.213 (0.000)
Hispanic	-0.260 (0.134)	0.220 (0.000)
Black	-2.228 (0.000)	-0.142 (0.000)
Other	3.745 (0.000)	0.197 (0.000)
military	-0.900 (0.195)	-0.352 (0.000)
unemployed	7.209 (0.000)	0.324 (0.000)
social_security	2.850 (0.000)	0.160 (0.000)
poverty	-2.553 (0.000)	-0.361 (0.000)
own_home	6.563 (0.000)	0.355 (0.000)
foreign_born	17.729 (0.000)	
no_hs	-8.704 (0.000)	-0.967 (0.000)
bachelor_or_higher	10.917 (0.000)	0.537 (0.000)
rural	-1.126 (0.000)	-0.079 (0.000)
MedHomeVal	0.000 (0.000)	0.168 (0.000)
log(MedHomeVal)		1.080 (0.000)
Adjusted R-squared	0.512	0.463
F-stat	5358	4406
MAE	3.961	1.352
RMSE	5.322	1.503

Appendix 6

Linear regression of logged economic mobility on ACS variables for each cluster

cluster	Intercept	male	Hispanic	Black	Other	military
1	-0.1182 (0.5767)	-0.4514 (0.014)	-0.371 (0)	-0.3113 (0)	-0.4836 (0)	-1.8183 (0.0051)
2	0.9146 (0)	-0.1461 (0.339)	0.3499 (1e-04)	0.0472 (0.2223)	0.1956 (0.2134)	-1.0653 (0.045)
3	0.8179 (0.0039)	-1.097 (0)	-0.2026 (0.1379)	-0.2112 (0.091)	1.0909 (0)	-3.4404 (0.0297)
4	-1.1057 (0.562)	1.2195 (0.072)	0.3201 (0.5809)	-0.7005 (0.1539)	0.4089 (0.6374)	-0.8402 (0.0199)
5	1.1386 (0)	-0.3962 (0.0034)	0.2029 (0)	-0.2176 (0)	0.8849 (0)	-1.0324 (0.1519)
6	1.0602 (1e-04)	0.9404 (0)	0.7452 (0)	-1.3887 (0)	0.3788 (0.0721)	-0.2889 (0.7488)
7	0.8527 (1e-04)	-0.6249 (0)	0.2787 (4e-04)	-0.4421 (0)	0.6709 (0)	-1.1889 (3e-04)
8	-0.1345 (0.4624)	-0.2663 (0.0512)	0.3868 (0)	-0.1872 (0)	0.2097 (0.0458)	-1.7418 (0.0029)
9	0.71 (0)	0.0211 (0.8807)	0.8784 (0)	0.0307 (0.6875)	0.8367 (0)	-0.609 (0.7322)
10	-0.5182 (0.0143)	-0.3494 (0.0341)	0.3096 (0)	-0.2214 (0)	0.123 (0.3289)	-2.2983 (0)
11	2.6123 (0)	0.6558 (0)	1.6025 (0)	-0.2509 (0.0012)	0.8577 (0)	-3.583 (0.0037)
12	1.6186 (0)	-1.0259 (0)	-0.5864 (0)	-0.8756 (0)	0.206 (0.0637)	-0.218 (0.704)
13	1.5229 (0)	-1.2115 (0)	-0.6005 (0)	-0.8442 (0)	-0.5802 (0)	-1.6723 (0.0096)
14	1.2102 (0.0033)	0.2217 (0.2753)	0.8859 (0)	-0.002 (0.9667)	-0.1397 (0.0224)	-0.5662 (0.8128)
15	3.5024 (0)	-0.444 (0.0557)	-0.8879 (0)	-0.4101 (0.0148)	-0.3275 (0)	-1.998 (0.0147)
16	-0.1141 (0.5197)	-0.1945 (0.143)	0.2842 (0)	-0.3237 (0)	0.0961 (0.1733)	0.5412 (0.4123)
17	0.4786 (0.0035)	-0.3826 (0.0062)	-0.1796 (0)	-0.3371 (0)	0.2042 (0.0016)	-1.72 (0)
18	0.6234 (0.0023)	-0.1442 (0.3011)	0.384 (1e-04)	0.146 (0.0053)	-0.5177 (3e-04)	-6.4009 (0.0078)

…

cluster	unemployed	social_security	poverty	own_home	foreign_born	no_hs
1	1.4056 (0)	0.7413 (0)	-1.0966 (0)	0.328 (0)	1.8656 (0)	-0.6508 (0.0054)
2	0.1114 (0.4723)	-0.0121 (0.8758)	-0.7051 (0)	0.1963 (0)	0.8255 (0)	-1.3668 (0)
3	0.2874 (0.3447)	-0.2831 (0.001)	-1.6701 (0)	0.0963 (0.2192)	1.3626 (0)	-0.648 (0.0568)
4	-0.4335 (0.5007)	0.6529 (0.3634)	0.2376 (0.692)	-0.9051 (0.0677)	1.6288 (0.2782)	1.9252 (0.1872)
5	0.2695 (0.079)	0.1526 (0.0297)	-0.6887 (0)	0.0084 (0.7945)	0.5449 (0)	-0.7051 (0)
6	-1.2451 (1e-04)	-0.2506 (0.0124)	-1.2315 (0)	-0.561 (0)	1.0357 (0.0026)	-2.982 (0)
7	0.466 (0.0553)	0.4219 (7e-04)	-0.0931 (0.1854)	-0.073 (0.2091)	1.3128 (0)	-2.2313 (0)
8	0.2686 (0.1008)	0.2872 (0)	-0.6225 (0)	0.1374 (7e-04)	1.1824 (0)	-1.6963 (0)
9	0.2163 (0.1072)	0.1945 (0.0057)	-0.1938 (0.0066)	0.0747 (0.0601)	0.2282 (1e-04)	-0.4242 (0)
10	0.7136 (0.0018)	0.5137 (0)	-0.935 (0)	0.0909 (0.0318)	1.6143 (0)	-1.9174 (0)
11	-0.4381 (0.0629)	-0.1927 (0.0354)	-0.7787 (0)	0.0402 (0.6772)	-0.6038 (0.0427)	-2.1116 (0)
12	0.5076 (0.1841)	0.8657 (0)	-0.142 (0.295)	0.0338 (0.5512)	1.5636 (0)	-1.6155 (0)
13	0.4295 (0.1031)	0.136 (0.0973)	0.171 (0.1965)	0.3932 (0)	1.1433 (0)	0.3844 (0.0029)
14	0.3041 (0.1776)	-0.2043 (0.1129)	-0.3459 (0.0224)	0.4969 (0)	-0.7611 (0.0106)	-0.5155 (0.0018)
15	0.8339 (0.0163)	0.1446 (0.0753)	-0.7524 (0.0046)	0.464 (0)	1.2824 (0)	-0.0631 (0.8383)
16	-0.0451 (0.7554)	0.3072 (2e-04)	-0.1911 (0.0166)	0.1274 (0.0095)	1.2148 (0)	-1.0843 (0)
17	0.6487 (3e-04)	0.2327 (6e-04)	-0.7406 (0)	0.1039 (0.0015)	1.1728 (0)	-1.0178 (0)
18	0.0158 (0.8803)	0.1726 (0.0355)	-0.0462 (0.561)	0.2514 (0)	0.8773 (0)	-0.713 (0)

…

cluster	bachelor_or_higher	rural	log(MedHomeVal)	adj.r.squared	fstatistic
1	0.2023 (0)	-0.2808 (0)	0.2078 (0)	0.2022	123.4937
2	-0.0122 (0.8767)	-0.1124 (0.0934)	0.1215 (0)	0.2907	91.3619
3	0.4842 (0)	-0.1565 (0)	0.1835 (0)	0.2411	60.6658
4	0.7175 (0.0145)	0.1197 (0.4836)	0.2146 (0.1669)	0.0855	2.3163
5	0.3827 (0)	0.1018 (0.0058)	0.1138 (0)	0.2630	112.8734
6	-0.3733 (1e-04)	0.1368 (0)	0.145 (0)	0.1017	44.9186
7	-0.1014 (0.1361)	-0.0617 (0.6705)	0.1495 (0)	0.2061	67.5074
8	0.2448 (2e-04)	-0.2477 (0)	0.2101 (0)	0.1724	125.8292
9	0.9732 (0)	0.256 (0)	0.0733 (0)	0.2555	71.3181
10	0.1378 (0.0043)	-0.361 (0)	0.249 (0)	0.1654	116.6623
11	0.3688 (0.01)	0.1688 (0)	-0.0499 (0.0256)	0.1170	56.5426
12	-0.1519 (0.0773)	-0.2414 (0.1264)	0.1195 (0)	0.2391	55.7768
13	0.7431 (0)	-0.4479 (0)	0.1173 (0)	0.5041	124.8025
14	0.554 (0.0213)	0.0073 (0.8387)	0.0433 (0.185)	0.0717	15.6683
15	0.2965 (2e-04)	-0.3352 (0)	-0.0476 (0.0142)	0.2395	49.6432
16	0.1143 (0.256)	-0.2707 (1e-04)	0.195 (0)	0.2699	98.4392
17	0.4318 (0)	-0.2868 (0)	0.1723 (0)	0.3041	152.7576
18	0.5354 (2e-04)	0.0782 (0.1515)	0.0984 (0)	0.1147	26.0039

Appendix 7

Median values of variables in each cluster

cluster	male	Hispanic	Black	Other	military	unemployed	social_security
1	0.493	0.058	0.021	0.076	0.000	0.033	0.294
2	0.457	0.055	0.663	0.039	0.000	0.078	0.294
3	0.470	0.051	0.014	0.036	0.000	0.042	0.508
4	0.589	0.182	0.132	0.101	0.496	0.073	0.020
5	0.500	0.602	0.050	0.059	0.000	0.059	0.241
6	0.506	0.024	0.005	0.023	0.000	0.033	0.363
7	0.502	0.110	0.120	0.093	0.000	0.049	0.185
8	0.482	0.061	0.041	0.041	0.000	0.050	0.356
9	0.501	0.781	0.052	0.018	0.000	0.076	0.241
10	0.488	0.054	0.035	0.050	0.000	0.034	0.310
11	0.502	0.026	0.009	0.023	0.000	0.048	0.406
12	0.493	0.111	0.045	0.142	0.000	0.035	0.171
13	0.493	0.150	0.024	0.509	0.000	0.044	0.287
14	0.492	0.027	0.107	0.022	0.000	0.076	0.428
15	0.489	0.078	0.014	0.168	0.000	0.034	0.296
16	0.487	0.161	0.173	0.061	0.000	0.093	0.307
17	0.492	0.283	0.066	0.104	0.000	0.049	0.259
18	0.458	0.029	0.823	0.020	0.000	0.158	0.342

cluster	snap	health_ins	poverty	absorb_shock	net_worth	own_home	MedIncome
1	0.022	0.972	0.035	0.800	310453	0.871	114063
2	0.202	0.900	0.192	0.530	44939	0.521	44853
3	0.045	0.957	0.070	0.728	223880	0.799	68243
4	0.026	0.984	0.074	0.617	90658	0.002	56495
5	0.159	0.847	0.168	0.605	83621	0.506	53533
6	0.060	0.952	0.074	0.671	143106	0.847	68866
7	0.107	0.909	0.199	0.646	78511	0.315	48801
8	0.136	0.928	0.141	0.622	102857	0.641	50440
9	0.281	0.759	0.295	0.478	34324	0.398	35988
10	0.052	0.953	0.064	0.690	136792	0.769	76618
11	0.114	0.910	0.134	0.584	95959	0.803	51269
12	0.032	0.956	0.086	0.764	134012	0.387	90637
13	0.068	0.945	0.092	0.742	241532	0.560	82686
14	0.197	0.874	0.226	0.515	70893	0.733	37976
15	0.012	0.980	0.037	0.855	655833	0.837	153700
16	0.305	0.890	0.303	0.583	75641	0.404	33292
17	0.079	0.917	0.088	0.726	210575	0.653	75297
18	0.406	0.890	0.387	0.452	25694	0.370	25111

cluster	MedHomeVal	MedAge	MedRent	married	foreign_born	no_hs	bachelor_or_higher
1	394600	42.700	1590	0.667	0.093	0.033	0.537
2	133400	35.900	984	0.296	0.063	0.128	0.212
3	280800	55.600	1196	0.509	0.074	0.054	0.388
4	193300	22.600	1504	0.793	0.052	0.020	0.305
5	241700	33.200	1209	0.450	0.322	0.253	0.167
6	191000	44.800	854	0.609	0.019	0.071	0.245
7	197800	31.100	1019	0.272	0.115	0.079	0.410
8	133700	39.900	825	0.417	0.041	0.106	0.208
9	122400	30.600	896	0.391	0.335	0.399	0.079
10	217600	40.200	1089	0.540	0.056	0.053	0.379
11	122800	44.100	712	0.550	0.015	0.131	0.161
12	608450	34.900	1721	0.349	0.197	0.039	0.676
13	597650	39.600	1656	0.543	0.449	0.122	0.376
14	86200	42.000	633	0.459	0.013	0.205	0.120
15	902200	44.500	2362	0.669	0.186	0.027	0.696
16	96500	33.900	774	0.280	0.086	0.205	0.132
17	306350	36.600	1447	0.508	0.212	0.112	0.298
18	70500	34.400	741	0.165	0.022	0.220	0.100

cluster	computer	internet	credit_score	rural	FWBscore	EMscore
1	0.969	0.939	744	0.000	66.314	21.290
2	0.884	0.778	663	0.000	58.408	9.220
3	0.915	0.854	730	0.000	65.323	15.330
4	0.995	0.966	697	0.000	61.809	9.200
5	0.909	0.814	692	0.000	62.142	12.900
6	0.908	0.829	727	0.993	63.733	12.040
7	0.937	0.845	706	0.000	61.527	11.470
8	0.878	0.797	703	0.000	62.485	10.000
9	0.824	0.675	663	0.000	57.083	10.450
10	0.944	0.893	716	0.000	64.007	14.020
11	0.849	0.742	692	1.000	62.483	9.310
12	0.964	0.918	737	0.000	64.286	19.055
13	0.938	0.887	735	0.000	64.108	24.245
14	0.765	0.631	668	1.000	58.288	7.800
15	0.975	0.950	756	0.000	67.208	28.580
16	0.827	0.706	690	0.000	59.838	7.810
17	0.949	0.893	715	0.000	64.386	16.025
18	0.758	0.599	653	0.000	56.511	6.560

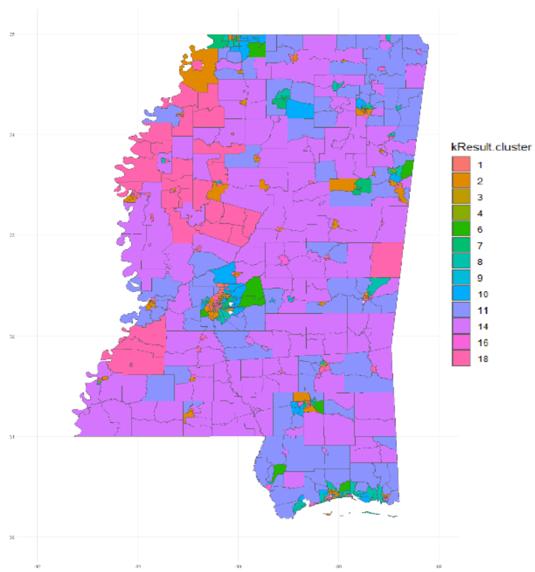
Appendix 8

National-level medians of variables

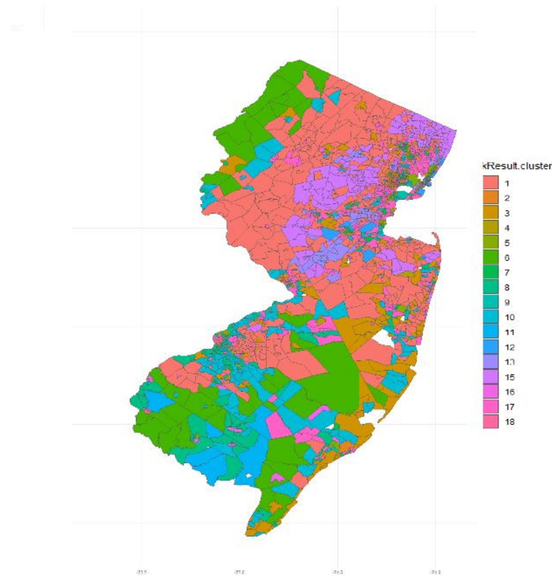
Variable	Median
male	0.491
Hispanic	0.076
Black	0.040
Other	0.050
military	0.000
unemployed	0.048
social_security	0.318
snap	0.096
health_ins	0.930
poverty	0.115
absorb_shock	0.651
net_worth	114909
own_home	0.685
MedIncome	59589
MedHomeVal	193300
MedAge	39
MedRent	1009
married	0.485
foreign_born	0.075
no_hs	0.098
bachelor_or_higher	0.257
computer	0.909
internet	0.832
credit_score	707
rural	0.000
FWBscore	63.082
EMscore	12.340

Appendix 9

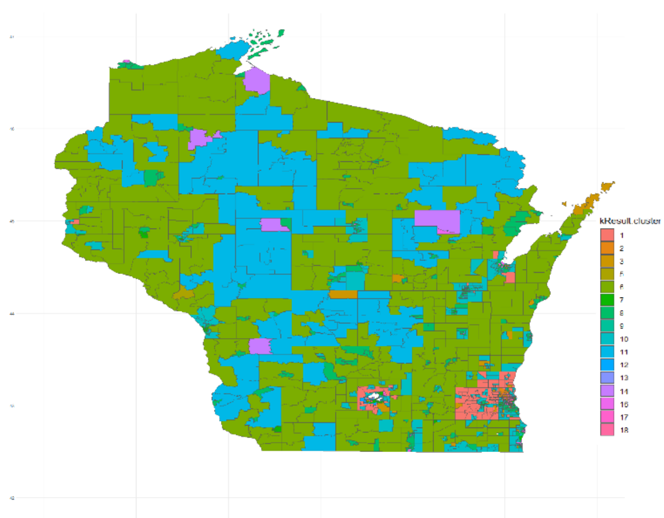
Cluster maps of Mississippi, New Jersey, and Wisconsin



Cluster map of Mississippi



Cluster map of New Jersey



Cluster map of Wisconsin

Appendix 10

Time-series analysis of Wisconsin economic mobility

a) Figure 10.1 displays time series plots and trends for the Wisconsin tracts with the 3 highest trends and the 3 lowest trends.

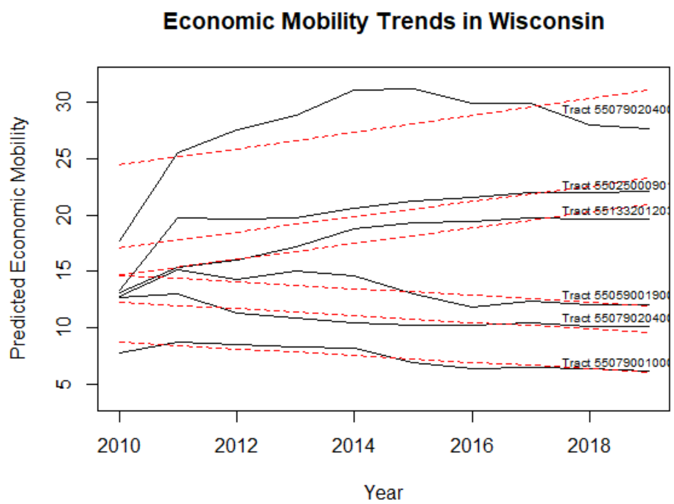


Figure 10.1: Economic Mobility Trends for 6 WI Tracts

b) The table below shows the intercept, trend, R^2 and p-values for the trend models in the 6 tracts shown in Figure 10.1. The trend models are statistically significant in 5 of the 6 tracts. The R^2 value is highest for tract 55133201203 meaning that the trend explains 80.8% of the variance in economic mobility score, and also means the trend is a good fit to the increasing values of economic mobility in the tract.

Tract	Trend	Intercept	R2	p.value
55025003200	0.740	23.699	0.237	0.087
55133201203	0.697	13.985	0.808	0.000
55025000901	0.685	16.436	0.588	0.006
55079001000	-0.294	8.998	0.730	0.001
55059001900	-0.297	14.966	0.396	0.030
55079020400	-0.299	12.592	0.693	0.002

c) The table below shows how the census variables changed from 2010 to 2019 for tract 55133201203, which has a high trend value and high R^2. Bachelor attainment rate increased, foreign-born population percentage increased, unemployment decreased, median home value (inflation-adjusted) increased, “other” population percentage increased, and high school dropout percentage decreased.

year	pred EM score	bache lor or higher	foreig n born	mil itar y	unem ploye d	social securi ty	povert y	own home	Median Home Val	male	Hispanic	White	Black	Other	no hs	rural
2010	13.10	0.29	0.025	0	0.057	0.378	0.011	0.856	243800	0.494	0.01	0.976	0.006	0.008	0.051	0
2011	15.35	0.507	0.037	0	0.058	0.372	0.02	0.854	242500	0.471	0.013	0.969	0.004	0.014	0.04	0
2012	16.00	0.501	0.053	0	0.054	0.342	0.012	0.868	242100	0.447	0.022	0.946	0.007	0.025	0.028	0
2013	17.16	0.523	0.094	0	0.063	0.342	0.013	0.877	229400	0.436	0.015	0.909	0.009	0.067	0.028	0
2014	18.78	0.569	0.147	0	0.06	0.353	0.017	0.839	230200	0.433	0.018	0.844	0.005	0.133	0.028	0
2015	19.32	0.545	0.189	0	0.057	0.346	0.013	0.817	232200	0.437	0.038	0.794	0.002	0.166	0.039	0
2016	19.43	0.539	0.2	0	0.056	0.346	0.014	0.806	239700	0.448	0.031	0.77	0.001	0.198	0.042	0
2017	19.78	0.588	0.194	0	0.044	0.343	0.012	0.798	261900	0.462	0.017	0.756	0.004	0.223	0.046	0
2018	19.66	0.58	0.181	0	0.03	0.372	0.011	0.827	273800	0.48	0.018	0.749	0.006	0.227	0.043	0
2019	19.61	0.613	0.132	0	0.026	0.336	0.007	0.879	282500	0.465	0.019	0.776	0.01	0.195	0.023	0

d) Figure 10.2 shows economic mobility scores of census tracts in Milwaukee County and Figure 10.2 shows the trend values of the tracts. Milwaukee county has many tracts with low economic mobility, but it is promising that there are many tracts with improving trends, especially in the middle of the county where current mobility values are typically very low. It is concerning though that several tracts have decreasing trends. Some tracts are missing due to changes in census designations between 2010 and 2019.

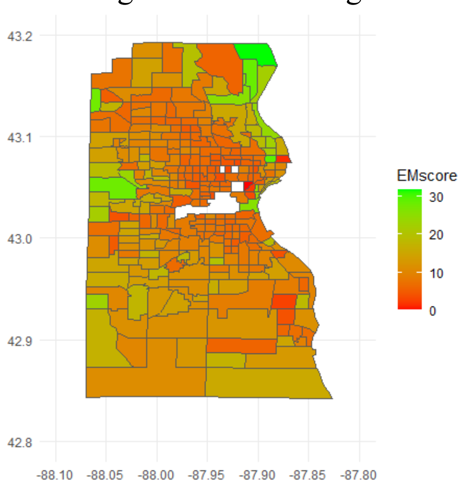


Figure 10.2: Economic mobility scores in Milwaukee County

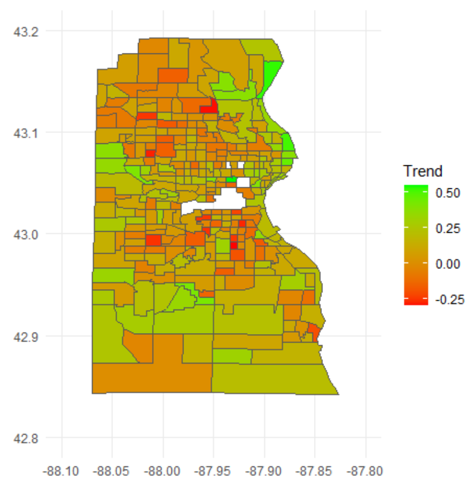


Figure 10.3: Economic mobility trends in Milwaukee County

e) Figure 10.4 displays a histogram of the trends in Wisconsin. ~82% of tracts in Wisconsin have a positive trend from 2010-2019, which is promising evidence that Wisconsin is generally improving its economic mobility over the course of time.

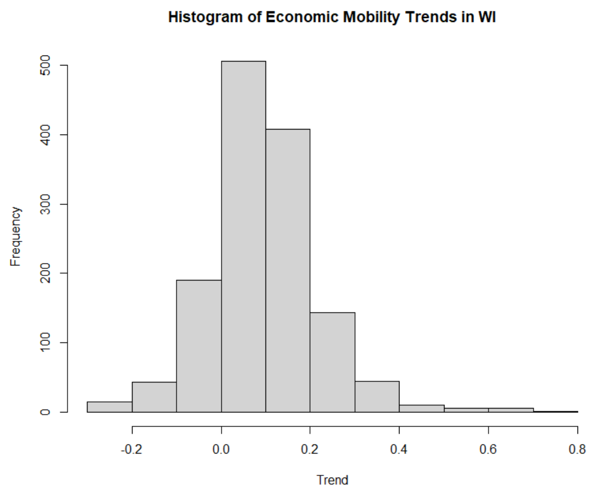


Figure 10.4: Histogram of trend values in Wisconsin

Appendix 12

R Code

1_fwb_predictive.R

Creates predictive regression model using FWBS data.

2_census_data.R

Retrieves and transforms census data and data from Urban Institute and Opportunity Atlas.

3_clustering.R

Performs hierarchical and k-means clustering and creates map visualizations.

4_forecasting.R

Creates time series trend models for Wisconsin census tracts.