Senior Thesis 301
Department of Economics, Vassar College

The Shape of Segregation and Intergenerational Mobility in Cities: An Interdisciplinary Approach

Kai Matheson

February 19, 2019

Abstract

The growing wealth gap in the United States points to the increasing importance of determining the causes of economic segregation and its impact on economic mobility. Yet, theories of class segregation, such as the concentric zone model, have not been empirically evaluated. Aspatial methods to measure segregation are often overly simplistic, ignoring critical aspects of the economic landscape. Evaluation of the relationship between economic mobility and economic segregation is sensitive to how one measures segregation. Thus, this work aims to explore methods of measuring economic segregation and to investigate the relationship between the shape or geographic structure of economic segregation and economic mobility in US cities. New spatial measures of economic segregation are developed, drawing from methods in applied mathematics and machine learning. These measures are then compared to existing ones. Given these measures, I explore the relationship between economic segregation and economic mobility.

Acknowledgements

I am filled with gratitude for each and every individual who has helped me through the process of dreaming up and creating this work.

Professor Pearlman, you made this process not just bearable but enjoyable. It was a pleasure to work with you, and to learn together. Thanks for believing in me.

Professor Frye and Professor Pradhan, thanks for the guidance and hours of conversations. Professor An, Professor Koechlin, Professor de Leeuw, and Professor Morin, thanks for giving me the skills to achieve this.

Mom, thank you for raising me, loving and accepting me, and being there for me. I hope I can not only make an impact on this unfair system but also make you proud. Dad and Nathan, thanks for always listening to me ramble, and being proud of me.

Ruth, thanks for being such a supportive mentor, friend, and godparent. Suze, I am proud to follow in your footsteps as I cross the stage in May. Derek, thank you for inspiring me to love numbers. Granny Cecily, your unwavering support and pride fill me with warmth.

April, Paige, Katie, Eli, Lauren, Mercy, Colby, Katherine, and Kyle, thank you for the late nights, the long talks, and the encouragement. To all my friends, thanks for being there to support me.

Monique and Saira, thank you for helping me cope.

Operation Understanding DC, thank you for shaping me.

Opportunity Insights, I am so excited to join you all and contribute to the work that you do. Thanks for being passionate for my passions.

I could not have done it without you all.

Love,

Kai

1. Introduction

The growing wealth gap in the United States points to the increasing importance of determining the causes of economic segregation and its impact on economic mobility. Yet, theories of class segregation, such as the concentric zone model, remain largely untested. Aspatial measures of segregation are often overly simplistic, ignoring critical aspects of the economic landscape. Evaluation of the relationship between economic mobility and economic segregation is sensitive to how one measures segregation. Thus, in this work, I aim to explore methods of measuring economic segregation and to investigate the relationship between these measures and economic mobility in US cities. New spatial measures of economic segregation are proposed. These are evaluated by comparing them to existing ones. I then use these measures to explore the relationship between economic segregation and economic mobility.

2. Literature Review

Imagine a city laid out like a typical checkerboard, with each square representing a neighborhood. The black squares represent a neighborhood made up of entirely wealthy people and the white squares represent a neighborhood made up of entirely poor people. In this checkerboard city, there is economic segregation *within* neighborhoods but not *between* neighborhoods– any one poor neighborhood is bordered by wealthy neighborhoods and vice versa. On a larger scale the city is, to a degree, economically integrated, though there is segregation within each neighborhood. Now imagine a checkerboard recolored so that the left half of the board's squares are black while the right half of the board's squares are white. There are low-income and high-income halves of this checkerboard city. This income map is much more segregated than the original checkerboard, though the level of *within*-neighborhood segregation is the same. As I will later discuss, this is one of the major flaws in typical methods of measuring "aspatial segregation." The "checker-board problem" highlights the fact that aspatial segregation measures ignore the spatial proximity of neighborhoods and instead only capture the composition within neighborhoods (Reardon, 2006).

The growing wealth gap in the United States, along with differences along class lines for measures of well-being such as life expectancy, point to the increasing importance of theoretically determining the nature and effects of class segregation and verifying them empirically. A big part of the problem is in choosing the right information to investigate. The manner in which we measure this information is equally important. Class segregation, as well as other factors such as cost of living, greatly play into the mechanisms of mobility, as prior theories as well as empirical work have investigated. However, as cost of living is not directly observable, and as there are many different ways to measure segregation, the different proxies or measures for these outcomes have potential to greatly influence what relationship we seem to find (Johnston et al., 2014). In measuring class segregation, it is crucial that these studies define exactly what they mean by segregation, and that they measure it spatially, as segregation is a spatial phenomenon.

Past studies have mostly relied on aspatial segregation largely influenced by arbitrary political divisions and boundaries (Dwyer, 2010). Even in the few studies that consider spatial segregation, they indirectly measure it through measures of segregation termed concentration, centralization, and clustering, which each in its own constitutes an aspatial measurement. These will be explained later on. A spatial measure of segregation is crucial to understanding the intensity of social exclusion among the disadvantaged and the degree of social closure achieved by the advantaged, rather than ignoring the location of neighborhoods within a metropolitan system (Dwyer, 2010). In this work, I take a more direct approach in measuring segregation spatially, by proposing, evaluating, and assessing measurements of spatial segregation and sprawl within metropolitan areas that rely on analysis of maps as images.

2.1. Why Segregation Matters for Mobility

Why would the relationship between upward mobility and segregation be an important one, or how might it work? How does segregation impact people, particularly children? Why might growing up in more segregated cities have long term negative consequences, even for individuals who move away?

Orfield and Lee (2005) study the modern process of resegregation in U.S. schools, and find that the level of concentrated poverty in a school is one of the biggest predictors of educational outcomes. While that study is focused on class segregation within schools, work by Mayer (2002) highlights the effects of Census tract-level class segregation on educational outcomes. She finds that increases in economic segregation within Census tracts in the same state hardly change average educational attainment but exacerbate the inequality between high-income and low-income children, with increases in segregation resulting in high-income children's increased educational attainment corresponding directly to low-income children's decreased educational attainment. Further, Mayer (2002) finds that economic inequality *within* tracts has little effect on low-income children's educational outcomes while changes in inequality *between* tracts lessen educational outcomes for low-income children. The theory behind this stems from school financing and neighborhood effects (Mayer, 2002). Educational outcomes are, of course, extremely linked with income, and thus better educational outcomes come hand in hand with higher percentiles in the income distribution. A poor kid starting out in a school with lots of concentrated poverty sees smaller educational gains and thus one would think is less upwardly mobile than a similar poor kid in a school with less concentrated poverty.

2.2. Existing Measures of Economic Segregation

It is important to consider how the existing literature defines and measures segregation. Might the estimated effects differ based on how segregation is measured? The majority of the literature utilizes aspatial measures of segregation, although the geographic component of segregation is crucial to how it manifests. Aspatial segregation measures ignore crucial questions– what is the shape that segregation takes? Are cities with multiple clusters of high-income neighborhoods or multiple clusters of low-income neighborhoods faring better than cities with only one, larger cluster of each? Are cities whose economic segregation takes shape through concentric circles inherently different than those whose segregation does not?

Mayer (2002) finds that between-neighborhood segregation matters considerably more than within-neighborhood segregation for educational outcomes of low-income children. Within-neighborhood segregation is primarily what aspatial measures pick up on. The scale of neighborhood that one looks at is also very sensitive to which cities are marked as having high or low segregation. Using maps that are drawn on a very granular level (Census tract), but without borders, we are not as sensitive to scale and get a more accurate picture of the segregation landscape. We must also address the fact that aspatial measures of segregation additionally fall victim to both the "checker-board problem" and the "modifiable areal unit problem" (Reardon, 2006).

The "modifiable areal unit problem" (MAUP) arises from population data typically being collected, aggregated, and reported for spatial units based on political divisions that may have no relationship with meaningful social or spatial divisions. This method of aggregating data implicitly assumes that people within these regions are more similar than the people on the boundaries of the regions who may be geographically located nearer to one another but in separate municipalities. Unless spatial boundaries correspond to meaningful social boundaries, all measures of segregation that rely on aggregates are sensitive to the drawing of boundaries (Reardon, 2006). While I am not able to address this problem completely, my approach of using maps without drawn borders that are aggregated on the Census tract level blur the boundaries more implicitly relative to other approaches that either ignore between-neighborhood segregation or handle the boundaries explicitly.

In the past, aspatial segregation has been measured using a dissimilarity index which ranks cities like New York as highly segregated, though it is thought to be a very diverse city. Both the dissimilarity index and the index of segregation indicate average situations without any indication of variation around that situation or the range of contexts that individual members of groups experience (Johnston et al., 2014). We also need to consider locked-in versus locked-out segregation versus congregation.

2.2.1. Existing Measures of Economic Segregation and their Equations

Name: Neighborhood Sorting Index (NSI)

Equation:

$$NSI = \frac{\sigma_N}{\sigma_H} = \frac{\sqrt{(\sum_{n=1}^{N} h_n(\bar{y_n} - \bar{y}))/H}}{\sqrt{(\sum_{i=1}^{H}(y_i - \bar{y})^2)/H}} \tag{1}$$

where y is income, i indexes households, n indexes neighborhoods, h_n is the number of households in neighborhood n, and H and N are the total number of households and neighborhoods respectively.

Interpretation/background: The proportion of income variation that lies between subareas; the ratio of the standard deviation of subarea mean incomes (weighted by subarea population) to the standard deviation of income in the regional population

(Reardon, 2006); square root of correlation ratio of between-tract income variance over the total income variance (Reardon and Bischoff, 2011; Jargowsky, 1996).

Data needed: Counts of households by income brackets or PUMS sample

Citation: Jargowsky (1996)

Additional notes: Between-tract variance is the household-weighted variance of the neighborhood means; total variance of income can be calculated using the PUMS individual-level sample data or to use the published counts of households by income brackets and make some assumptions about the distribution of households within income brackets–based on testing and comparison with PUMS estimates, Jargowsky (1996) assumes linear distributions in lower brackets and Pareto distributions in the brackets above the MSA mean, described in his Appendix A; this measure is the most well-known in sociology Reardon and Bischoff (2011); because NSI does not satisfy the property of being independent from income inequality, it may confound changes in residential sorting by income with differences in income distributions across time, place, and groups (Reardon and Bischoff, 2011).

Name: Generalized Neighborhood Sorting Index (Generalized NSI)

Equation:

$$GNSI_k = \sqrt{\frac{\sum_{i=1}^{H}(m_{ki} - M)^2}{\sum_{i=1}^{H}(y_i - M)^2}} \tag{2}$$

where m_{ki} is the mean household income in the kth order-expanded community from a household i.

Since individual household data is not available, and income data is normally provided as summaries for geographical neighborhood boundaries e.g. census tracts, we change the equation for GNSI to the following:

$$GNSI_k = \sqrt{\frac{\sum_{n=1}^{N} h_n (m_{kn} - M)^2}{\sum_{i=1}^{H} (y_i - M)^2}} \tag{3}$$

where m_{kn} is the mean household income in the kth order expansion from a neighborhood n (Jargowsky and Kim, 2005).

Interpretation/background: Spatial index of income segregation based on the NSI incorporating a "distance-decay" effect (Reardon, 2006); key difference from NSI is that the numerator of GNSI incorporates a flexible moving window for the calculation of a neighborhood's economic level which is larger thana the neighborhood itself, e.g. all neighborhoods whose centroids are within a certain distance r of the given neighborhood's centroid (Jargowsky and Kim, 2005).

Data needed: Counts of households by income brackets or PUMS sample

Citation: Jargowsky (1997); Jargowsky and Kim (2005)

A note on entropy: Entropy is commonly used in physics and information theory to measure the randomness of a system. Theil brought entropy to the social sciences with the Theil Index and more. It can also be thought of as the amount of information needed to describe a probability distribution. It is used in a lot of segregation measures, but it can only be calculated for discrete distributions (Roberto, 2015).

Theil Index:

$$I = \frac{1}{N} \sum_{i=1}^{N} \frac{x_i}{\bar{x}} \log \frac{x_i}{\bar{x}} \tag{4}$$

where x_i is the income of groups of earners, $\bar{x}$ is the average income. When all incomes are equal (all individuals earn the mean income), there is no inequality so $I = 0$ (Roberto, 2015).

Name: Rank-order information theory index (H^R)

Equation: Let p be income percentile ranks in a given income distribution i.e. $p = F(Y)$ where Y is income and F is the cumulative income density function. In other words, p is the proportion of the population with incomes below a certain threshold. For any p, we can compute the residential segregation between those with income ranks less than p and those with income ranks greater than or equal to p. Let $H(p)$ denote the value of the traditional information theory index of segregation computed between the two groups. Let $E(p)$ denote the entropy of the population when divided into these two groups.

$$E(p) = p \log_2 \frac{1}{p} + (1-p) \log_2 \frac{1}{1-p} \tag{5}$$

$$H(p) = 1 - \sum_j \frac{t_j E_j(p)}{T E(p)} \tag{6}$$

where T is the population of the metropolitan area and t_j is the population of neighborhood j. Then the rank-order information theory index H^R can be written as

$$H^R = 2\ln(2) \int_0^1 E(p) H(p) dp \tag{7}$$

The 2000 census reported 16 income categories, which allow us to compute $H(p)$ at 15 values of p. They then approximate the function $H(p)$ over $(0, 1)$ by fitting an mth-order (i.e. 4th-order) polynomial to the values, weighting each point by the square of $E(p)$:

$$H(p) \approx \beta_0 + \beta_1 p + \beta_2 p^2 + ... + \beta_m p^m + \varepsilon_p,\ \varepsilon_p \sim N(0, \frac{\sigma^2}{E(p)^2}) \tag{8}$$

If $\hat{\beta_k}$ is the kth coefficient from this model, then

$$\hat{H^R} = \hat{\beta_0} + \frac{1}{2}\hat{\beta_1} + ... + (\frac{2}{(m+2)^2} + 2\sum_{n=0}^{m} \frac{(-1)^m - n({}_mC_n)}{(m-n+2)^2})\hat{\beta_m} \tag{9}$$

where ${}_mC_n \approx m!/(n!(m-n)!)$ is the binomial coefficient (the number of distinct combinations of n elements from a set of size m). Then, we can estimate other $H(p)$ at different values of p by using the fitted polynomial $H(p)$ equation.

Interpretation: Ratio of within-unit (tract) income rank variation to overall (metropolitan area) income rank variation; 0=no income segregation (income distribution in each census tract mirrors that of the region as a whole) and 1=complete income segregation (no variation in income in any census tract and only across census tracts); H^R is a weighted average of the binary income segregation at each point in the income distribution, where the weights are proportional to entropy $E(p)$ which is maximized when $p = 0.5$ and minimized at $p = 0$ or $p = 1$– these weights assume that segregation between those above and below the median is more important than segregation between those above and below the 90th percentile.

Data needed: Categorical income data with a lot of bins

Citation: Reardon and Bischoff (2011); Reardon et al. (2006, 2018)

Additional notes: Aspatial; spatial version uses definitions of "local" ranging from radii of 500-4,000 meters– arbitrary radii aren't good; this is the segregation measure used in Chetty and Hendren (2016); Roberto (2015) argues that this index measures relative diversity and not segregation, so it is problematic to interpret it as a measure of segregation– rather, it measures relative homogeneity, comparing the diversity of local areas to overall diversity of a region rather than measuring the difference between the local and overall proportions of each group (Roberto, 2015)

Tools: Stata RANKSEG module to compute rank-order segregation measures with finite sample-bias correction from Reardon et al. (2018); rankorderseg function from OasisR package for R to compute rank-order segregation measures from Reardon (2011).

Name: Segregation of poverty

Equation: $H(p25)$, rank-order information theory index (Chetty and Hendren, 2016).

Interpretation: Uneven distribution of low- and non-low-income households among neighborhoods (Reardon and Bischoff, 2011).

Citation: Reardon and Bischoff (2011); Chetty and Hendren (2016)

Name: Segregation of affluence

Equation: Same as segregation of poverty, but $p = 75$ instead of $p = 25$ (Chetty and Hendren, 2016).

Interpretation: Uneven distribution of high-income and non-high-income households among neighborhoods (Reardon and Bischoff, 2011).

Citation: Reardon and Bischoff (2011); Chetty and Hendren (2016)

Name: Nearest neighbor spatial ordering; Monocentric spatial ordering

Equation: Given an initial "parade" of neighborhood per capita incomes (y_j) ranked such that $y_1 \leq y_2 \leq ... \leq y_j \leq y_J$, a spatial ordering is a reranking of the original parade in a way that reflects the jth neighborhood's spatial position.

There are many ways to do spatial orderings, but the two that Dawkins (2007) suggests are Nearest Neighbor Spatial ordering and Monocentric Spatial Ordering:

Nearest Neighbor Spatial Ordering: Holding neighborhood incomes constant, assign each neighborhood the income parade ranking of the most spatially proximate nearby neighborhood.

Monocentric Spatial ordering: Define X which is equal to the distance from a given point within the region to the centroid of each neighborhood. Holding neighborhood incomes constant, rank the neighborhood income parade by ascending or descending values of X.

Using these spatial orderings, we can create a Spatial Ordering Index (S_r), the ratio of the covariances, modifying the standardized spatial Gini index from Dawkins (2004) as follows:

$$S_r = G_r/G_B \tag{10}$$

where S_r is a spatial ordering index calculated from the rth spatial ordering, G_B is the Gini index of between-neighborhood income segregation, and G_r is a spatial Gini index calculated from either a nearest neighbor or monocentric spatial ordering. This makes S_r a measure of the degree to which a given spatial configuration of neighborhoods results in a reordering of neighborhood income distribution used to construct G_B– so it is a direct measure of the importance of the checkerboard phenomena.

G_r and G_B can be calculated using covariance-based formulas:

$$G_r = \frac{2\sum_{j=1}^{J}(\bar{R}_{j(n)} - \frac{H+1}{2})y_j}{HY} \tag{11}$$

$$G_B = \frac{2\sum_{j=1}^{J}(\bar{R}_j - \frac{H+1}{2})y_j}{HY} \tag{12}$$

where y_j is aggregate household income earned by residents of the jth neighborhood, Y is aggregate household income earned by residents of the region, $\bar{R}_j$ is the average rank of per capita income earned by neighborhood j within the overall neighborhood per capita income distribution, and $\bar{R}_{j(n)}$ is the average spatial rank of per capita income earned by neighborhood j within the overall neighborhood per capita income distribution.

To calculate $\bar{R}_j$, first rank neighborhoods in ascending order by per capita income. Let N_j be "the cumulative number of households within each successive position within the neighborhood income parade" (I don't really understand this). Let N_{j-1} be the cumulative total of households one rank lower within the income parade. The jth neighborhood's average rank is then equal to $(N_j + N_{j-1} + 1)/2$.

To calculate $\bar{R}_{j(n)}$ for a nearest neighbor spatial ordering, pair each neighborhood with its most spatially proximate neighbor ($j(n)$), and assign the jth neighborhood a rank equal to the rank of $j(n)$. $N_{j(n)}$ is the cumulative number of households within each successive position within the spatially ordered income parade, and $N_{j-1(n)}$ is the cumulative number of households one rank lower within the spatially ordered income parade. The jth neighborhood's average rank is then equal to $(N_{j(n)} + N_{j-1(n)} + 1)/2$.

Then, we can plug in G_B and G_r to S_r to yield:

$$S_r = \frac{\sum_{j=1}^{J}(\bar{R}_{j(n)} - \frac{H+1}{2})y_j}{\sum_{j=1}^{J}(\bar{R}_j - \frac{H+1}{2})y_j} \tag{13}$$

The paper proposes a couple more ways of how to calculate the segregation index which I will not go into detail of here.

Interpretation: Can be thought of as the ratio of two covariances

Data needed: "Average neighborhood income per household"/"neighborhood per capita income", data on individual households?!!?!

Citation: Dawkins (2007)

Additional notes: Spatial ordering index using spatial autocorrelation– the only measure of economic segregation I have found thus far that uses spatial autocorrelation; Claims to be less sensitive to presence of outliers, satisfies the "principle of transfers," and is "flexible enough to quantify a variety of spatial patterns of segregation"; this measure seems only to pick up on between-neighborhood segregation and ignores the aspect that mean income ("neighborhood income") is not reflective of the income of everyone in the census tract and that income could sometimes vary greatly (or not) within neighborhoods, making means or medians not representative of the actual phenomenon, rendering the initial ranking to be somewhat uninformative

Some additional measures:

- Divergence Index aims to address the need for a decomposable measure of segregation, based on relative entropy, a measure of difference between two probability distributions; aspatial (Roberto, 2015)

- Rank-order variance ratio index; rank-order square root index (Reardon, 2011)

- Theil index of income segregation is based on relative entropy and is interpreted as one minus the ratio of within-neighborhood variation in income percentile ranks to the overall variation in percentile ranks (Watson 2006, 2009) (Reardon, 2011; Roberto, 2015)

- Neighborhood Disparity Index (NDI) uses information on contiguity to arrive at individual disparity values for individual parcels, and an indexed value for an area (Chakravorty, 1996)

- Segregation indices discussed in Florida and Mellander (2015): poverty segregation index, wealth segregation index, overall segregation index, all based on Index of Dissimilarity developed by Massey and Denton, compares the spatial distribution of a selected group of people with all others in that location (Florida and Mellander, 2015)

2.3. Theories of Urban Segregation

We now draw away from the empirics behind segregation and more towards theories of class segregation, which have yet to be empirically evaluated. There are many contradictory theories of urban land use debating the ways in which class segregation manifests in cities. Among these are the concentric zone model, the multiple nuclei model, and a more patchwork model.

Many urbanists believe American cities are undergoing demographic inversion, which refers to the process by which suburbs become the principal region where low-income individuals settle due to the increasing cost of living in central cities (Ehrenhalt, 2012). From 2000 to 2008-2012, the percentage of suburban poor in the United States increased by 139%, which is nearly three times more than within cities. By 2008-2012, 46% of all non-rural poor residents living in concentrated poverty lived in the suburbs (Kneebone, 2014). Large metropolitan suburbs house about one-third of low-income Americans, a greater share than big cities, small metropolitan areas, or rural areas. Through the 2000s, suburban poverty increased at a rate five times more than what we have seen within cities (Tomer et al., 2011).

Studies have shown that the primary reason for central city poverty is access to the public transportation system. Upward mobility, that is the capacity of increasing one's social or economic position, is currently higher in cities with less sprawl, as measured by commute times to work (Chetty et al., 2014). This suggests that job access, economic segregation, and transportation access are strongly correlated and interdependent.

Urban economics literature has developed different types of models to analyze the socioeconomic distribution of a city. One of the most well-known monocentric city models, created by Ernest W. Burgess in 1925, posits that the large American city can be generalized to have a central business district surrounded by a zone of transition including other industries, followed by inner-city poor residences and then high-income residences located in the suburbs. This is contextualized within an industrial city, where the wealthy prefer the suburbs because of pollution and violence downtown, while low-income individuals prioritize lower transit costs. This model is still in use today but is seen as outdated as the post-industrial processes of gentrification and displacement become increasingly more common. An older monocentric city model was created in 1841 by J.G. Kohl, based on the pre-industrial cities of continental Europe, in which the high-income population was housed in the city center and the low-income communities resided farther from the city. With demographic inversion occurring in many large American cities, the Kohl model becomes relevant again in such cases.

There are other models, however, that do not assume a ring-like structure to the city (Hoyt's sector, Harris and Ullman's multiple nuclei).

However, there is no consensus over how to evaluate which model is best for certain cities. Further, there has been almost no work done on directly analyzing the shape of economic segregation. The existence of this debate highlights the need for research done on evaluating which models are the most accurate, or for which types of cities. There has been little work done to empirically evaluate the theories aside from individual case studies based on subjective observation alone. In assuming and evaluating whether the structure of class segregation in cities fits these various models using data-driven methodology, I add to the literature on urban land use by validating or invalidating theories of class segregation on a larger scale.

2.4. Towards Spatial Measures of Segregation

I am interested in spatial measures. There are five very common categories of spatial segregation measures: evenness, exposure, concentration, centralization, and clustering (Dwyer, 2010). While these measures implicitly consider geography, the spatial measures presented in this paper more explicitly capture the geographic arrangements of wealth.

The only paper that has come close to what I would like to study is work by Dwyer (2010) on predicting the spatial form of class segregation using metropolitan factors of suburbanization and levels of class and racial inequality. She measures the spatial form of class segregation by looking at combinations of measures of concentration, centralization, and clustering.

What Dwyer (2010) calls the five "spatial dimensions" of segregation are evenness, exposure, concentration, centralization, and clustering. They are defined in Table 1. She argues that by combining those three spatial dimensions, she is able to characterize the economic geography of cities. She defines hypersegregation as multiple overlapping dimensions of segregation (Dwyer, 2010). Dwyer (2010) aims to assess how widely the concentric zone model of class segregation holds in metropolital areas and investigate whether there are alternative spatial forms. She looks into whether levels of class and racial inequality can predict the spatial form of class segregation. She finds...

Dimension	Definition	Examples
Evenness	degree to which a group is spread in equal proportions among another group across neighborhoods	dissimilarity index, Theil index
Exposure	likelihood a member of one group will come into contact with a member of the other group	isolation index, interaction index
Concentration	land area taken up by one group compared to another, whether a group resides in a relatively small portion of the metro area or is spread out over more space	
Centralization	degree to which groups are located near the center of the metropolitan area versus the periphery	
Clustering	whether one group is located in neighborhoods near other neighborhoods dominated by the same group versus the other group	

Table 1: The five spatial dimensions of segregation as defined by Dwyer (2010) and Reardon (2006).

That makes this paper the second-ever study to analyze the spatial dimensions of class segregation. But, rather than relying on a combination of aspatial measures of "spatial dimensions," as is done by Dwyer (2010), my methods evaluate spatial segregation in cities drawing on methods from applied mathematics and computational science in order to directly assess and create measures of spatial segregation using actual income maps of cities as the bases for assessment. I then add to the literature on determinants of upward mobility by more thoughtfully assessing the relationship between class segregation and mobility using these measures.

3. Data

3.1. Data Description

Absolute upward mobility is defined as the mean rank of children whose parents are in the 25th percentile of the income distribution, with values potentially ranging from 1 to 100. I source data on a Census tract level of this measure of absolute upward mobility from Chetty et al. (2018). They predict this outcome using data for 20.5 million children who were born in the U.S. or are authorized immigrants who came to the U.S. in childhood, born between 1978 and 1983. The authors link the children to their parents' pre-tax household incomes from 1994-2000, and measure children's pre-tax incomes in 2014-2015, when they are ages 31-37 (Chetty et al., 2018).

I source data on determinants of mobility on the levels of Census tract, Metropolitan Statistical Area, or county, depending on data availability at the most granular level, in order of preference as stated. See Appendix A for information on which variables are found on which level (to be added later). It is a weakness of the data that I am not able to find measures all on the Census tract level, as this poses... (what problems does this pose? I just feel like it's not good).

I source information on top industry by MSA from (blank). Controls for baseline city characteristics would include median age, median household income, some measure of cost of living (perhaps Consumer Price Index, but I cannot find an aggregate dataset), and some measure of poverty by city. I also want to measure the white and black populations of the city, and inequality. *Within-neighborhood* economic inequality could be measured through the common Gini coefficient for each census tract, the proportion of the city that is super wealthy (say, in the top 1%), or through the maps by measuring distribution of color.

As education is crucial to mobility, I want to measure the quality of public schools in an area. This could be measured by inputs, such as mean public school expenditures per student or mean class size, or it could be measured by outputs, such as test scores, dropout rates, and percent who continue to college.

Different health outcomes such as life expectancy and obesity also hold weight. Housing measures, such as distribution of house prices, the percentage of renters in a city, and the proportion of single-family homes are also important covariates. Measures of family structure, such as teenage birth rates, proportion single parents, proportion adults divorced, and proportion adults married are also telling but have also been found in the literature to explain away certain race effects. I must also account for local tax and government expenditure policies.

Crime rates are important as well. Violent crime rate specifically has been used as a social capital index in past work. There are other social capital indices to consider, such as one related to the proportion who vote, respond to the census, and participate in community organizations. More examples are the proportion of religious individuals or the number of bowling alleys in an area. Rates of migration may also be important.

Some of these measures are already well-defined while others are not. Transit infrastructure affects low income people's access to various parts of the city, so some measure of public transit quality or accessibility as well as walkability of the city are crucial to include. These data are not readily available, measurable characteristics, however.

I construct measures of segregation and sprawl on the MSA level, and limit the scope of my data to only those census tracts which fall within an MSA, assigning MSA-level features to each Census tract within a given MSA. These measures are described in detail in Section 3.2.1.

3.2. Data Construction

I create an indicator of whether an MSA is a "megacity" by clustering the cities into two groups by population and GDP.

3.2.1. Measures of Shape

For possible measures of shape, I draw from machine learning, applied mathematics and spatial statistics in order to characterize and measure segregation and sprawl. This involves creating maps in GIS of median household income by Census tract for entire Metropolitan Statistical Areas (MSAs) and bits of their surrounding regions (I use the year 2000, the tail-end of the years in which Chetty et al. (2018) measure parents' income used in calculating our estimates of upward mobility), converting those maps to images, and running analyses on them which may include converting the images to sets of points in software such as MATLAB.

In trying to investigate whether economic segregation is concentric, there are a few different approaches to consider. The first is to generate thousands of fake images of concentric cities by defining all the free parameters in a concentric city model and incorporating noise, perhaps utilizing Procrustes transformations or projections. Then, I can take those images as well as fake images of non-concentric cities, and train a convolutional neural network classifier to learn how to identify concentric cities. I can input the real images into the classifier, and get the probability that each image is concentric.

Another approach to measuring concentricity is to use feature extraction tools for concentric circles, given the center of the city. We could also do this for stripes and patches. This would probably require converting the images to collections of points, and computing the center of mass of the points. In applied mathematics, measures of eccentricity are often used to determine whether something is circular. By converting the images to sets of points, I can apply these methods to the maps.

There are ways of creating variables that leave a lot of decisions up to the machine. The problem with these approaches is that the variables will not be easily interpretable, although they correspond to properties of the images. The benefit is that it would largely be a plug-and-chug process. One means of variable creation is to train a variational autoencoder on the images, and use the latent variables from inside the model. Another approach is to run an unsupervised clustering algorithm on the maps and then hand-label the clusters to have cluster dummy variables– this would pick up on certain elements of economic segregation that these cities share, and account for those.

There are some more basic, understandable characteristics of segregation and sprawl that we can pull from the maps. In the past, sprawl has been measured through commute times to work, but with the use of maps, I can consider measuring sprawl by the distance from the city center to the furthest point in the MSA, or the average radius, controlled by population size. We can get the radius of the city, which could be defined as the distance from the center of the city to the farthest point on the border, or to the average border distance. Since all the maps will be the same size in terms of pixels, it is also important to include some sort of scaling variable for how large the city actually is in square miles.

As well, measures of contrast, for example, the difference in color between the 75th percentile and 25th percentile pixel values in the color distribution of the map, would serve as descriptors for how disparate income inequality is in a city– median household income of cities has been found in past works to have no relationship to intergenerational mobility, so we want to be able to differentiate between those cities which are homogenously middle-income and those that are made up of the extreme rich and the extreme poor. These measures would be absolute, as the colors on the maps are in reference to all cities, not city-specific.

Other methods from applied math include fractal dimension of boundaries and box counting. I want to look more into methods used in spatial statistics, geography, and topological data analysis. Topological data analysis tools could help to determine whether a city has multiple nuclei versus just one, with applications from persistence homology. I must continue to research how segregation and sprawl have been quantified in the past.

4. Model Description

I am exploring the relationship between upward mobility and, specifically, segregation and sprawl, while holding constant an extensive list of other city outcomes as discussed in Section 3.1 and adding in fixed effects.

I am considering a couple different levels of fixed effects. I have ring fixed effects for different proportions of bins of distance from the city center, as explained in Section 3.2.1. This accounts for differences in urban or suburban living associated with mobility that is constant across cities. I use region fixed effects, as designated by the U.S. Census regions (Appendix B) to account for unobservable characteristics that are constant throughout a region.

My main regression is seen in Equation 14.

$$\begin{aligned} UpwardMobility_s = \beta_0 + \beta_1 Segregation_s + \beta_2 Sprawl_m \\ + \sum_{i=1}^{k-1} \beta_i SegregationCluster_{im} + \delta_r \\ + \beta_s \mathbf{x}_s + \beta_m \mathbf{x}_m + \beta_c \mathbf{x}_c + \varepsilon_t \end{aligned} \quad (14)$$

where s is a Census tract, m is an MSA, r is a ring, and c is a county. The predictor $Segregation_s$ is based off a standard measure of segregation called the (blank). The predictor $Sprawl_m$ is the MSA-level measure of sprawl constructed from the maps. The $SegregationCluster_{im}$ variables represent the cluster fixed effects based off of the maps. The vector δ_r is of ring-level fixed effects. The construction of these variables can be read about in Section 3.2.1.

The vectors $\mathbf{x}_t, \mathbf{x}_m$, and $\mathbf{x}_c$ represent the controls discussed in Section 3.1 on different levels of geography.

The coefficients on measures of segregation and sprawl have a causal interpretation if the zero conditional mean assumption holds. That is, the errors are not correlated with the independent variables because we have accounted for all confounding factors. As well, they have a causal interpretation if there is no issue of reverse causality, or if segregation and sprawl are not caused by differences in mobility. Both of these assumptions are hard to test and verify, and are unlikely to be met based on the theory behind the relationship between segregation, sprawl and mobility. Thus, we must likely interpret relationships between the independent and dependent variables as correlations rather than causal relationships.

5. Estimation Techniques

K-means clustering algorithm basic explanation should go here.

6. Results

6.1. K-means clustering preliminary results

I have constructed the maps and clustered them into 6 groups using the K-means clustering algorithm. Please see some examples of maps within each cluster in Appendix C. Some current problems with my approach include:

- K=6 chosen arbitrarily
- Clustering seems currently to be mostly based on where there is whitespace
- Data issues - only 276 observations instead of the full 380

Even with these problems, I had some significant results, which is promising! See a coefficient plot of my regression of clusters on absolute upward mobility in Figure 1, where Cluster 1 is my reference group.

As you can see in Figure 1, both Cluster 3 and Cluster 5 have a statistically significant association with lower upward mobility in their metropolitan statistical areas when compared with that of Cluster 1. In Appendix C, you can see examples of images from each cluster. Cluster 1 seems to be made up of cities that are poor in the center and richer on the outsides, where yellow represents low income and red represents high income. Cluster 3, on the other hand, seems to be made up of cities that are rich in the center

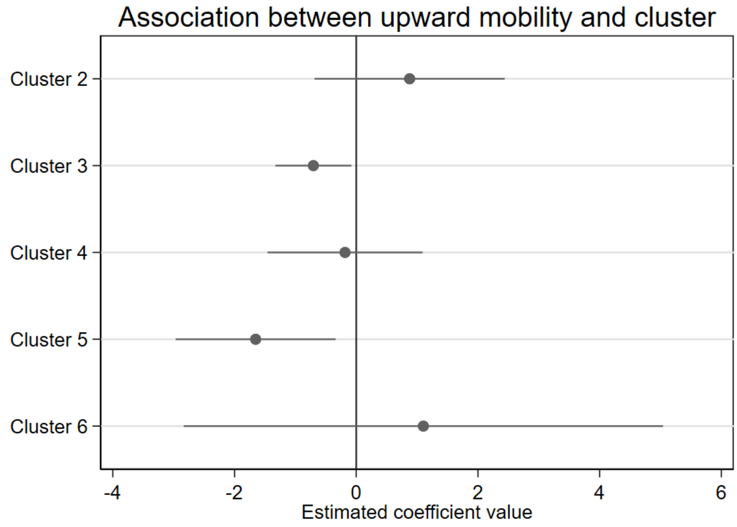


Figure 1: Coefficient plot of clusters on absolute upward mobility.

and poor on the outside. Cluster 5 seems to be made up of cities where there is water on the northeast side of the city (represented as white space in the maps).

These results support my hypothesis that mobility may be affected by the layout of economic segregation. Namely, that Cluster 3 (rich in center, poor on outside) is associated with lower upward mobility compared with Cluster 1 (poor in center, rich on outside). I have included a table in Appendix D which shows each MSA's cluster and estimate for absolute upward mobility, sorted by cluster and then by lowest to highest mobility.

Right now, it seems as though Clusters 1 and 3 are the only two clusters that are actually clustered based on economic geography alone– the rest seem clustered mostly based on where water is. Generally, Cluster 2 has water in the southwest, Cluster 4 has water in the southeast, Cluster 5 has water in the northeast, and Cluster 6 has water in the northwest. Once I figure out how to have the algorithm disregard white space in the images, then the clusters will be even more focused on the economic geography component and we may see more specific patterns within the clusters.

7. Conclusion

References

Chakravorty, S. (1996). A Measurement of Spatial Disparity: The Case of Income Inequality. *Urban Studies 33*(9), 1671–1686.

Chetty, R., J. N. Friedman, N. Hendren, M. R. Jones, and S. R. Porter (2018). The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility. Technical report, National Bureau of Economic Research.

Chetty, R. and N. Hendren (2016). The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates *. *The Quarterly Journal of Economics*.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the Land of Opportunity: The Geography of Intergenerational Mobility in the United States. *Quarterly Journal of Economics 129*(4).

Dawkins, C. J. (2007). Space and the measurement of income segregation. *Journal of Regional Science 47*(2), 255–272.

Dwyer, R. E. (2010). Poverty, prosperity, and place: the shape of class segregation in the age of extremes. *Social Problems 57*(1), 114–137.

Ehrenhalt, A. (2012). *The great inversion and the future of the American city.* Knopf.

Florida, R. and C. Mellander (2015). Segregated City: The Geography of Economic Segregation in America's Metros. Technical report, Martin Prosperity Institute.

Jargowsky, P. A. (1996). Take the money and run: Economic segregation in U.S. metropolitan areas. *American Sociological Review 61*, 984–998.

Jargowsky, P. A. (1997). *Poverty and place : ghettos, barrios, and the American city.*

Jargowsky, P. A. and J. Kim (2005). A Measure of Spatial Segregation: The Generalized Neighborhood Sorting Index.

Johnston, R., M. Poulsen, and J. Forrest (2014). Segregation matters, measurement matters. *Social-spatial segregation: Concepts, processes and outcomes*, 13–44.

Kneebone, E. (2014). The growth and spread of concentrated poverty, 2000 to 2008-2012. *The Brookings*.

Mayer, S. E. (2002). How economic segregation affects children's educational attainment. *Social forces 81*(1), 153–176.

Orfield, G. and C. Lee (2005). Why segregation matters: Poverty and educational inequality. *Civil Rights Project at Harvard University (The).*

Reardon, S. F. (2006). A conceptual framework for measuring segregation and its association with population outcomes. In *Methods in social epidemiology*, pp. 169–192.

Reardon, S. F. (2011). Measures of Income Segregation.

Reardon, S. F. and K. Bischoff (2011). Income Inequality and Income Segregation. *American Journal of Sociology 116*(4), 1092–1153.

Reardon, S. F., K. Bischoff, A. Owens, and J. B. Townsend (2018). Has Income Segregation Really Increased? Bias and Bias Correction in Sample-Based Segregation Estimates. *Demography 55*, 2129–2160.

Reardon, S. F., G. Firebaugh, D. O'Sullivan, and S. Matthews (2006). A new approach to measuring socio-spatial economic segregation. In *29th general conference of the International Association for Research in Income and Wealth, Joensuu, Finland.*

Roberto, E. (2015). The Divergence Index: A Decomposable Measure of Segregation and Inequality. Technical report.

Tomer, A., E. Kneebone, R. Puentes, and A. Berube (2011). Missed opportunity: Transit and jobs in metropolitan America.

Appendix

A. Geographic levels of determinants of mobility

table of level of geography of controls

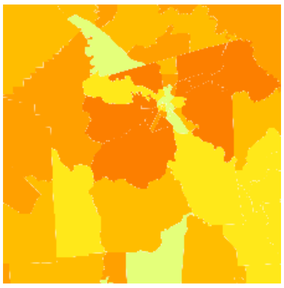
B. U.S. Census Regions

map showing census regions

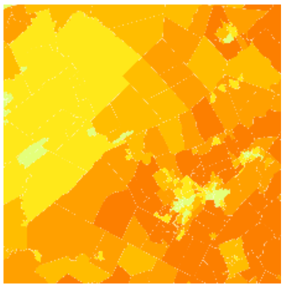
C. Some example images from each cluster

Figure 2: Cluster 1

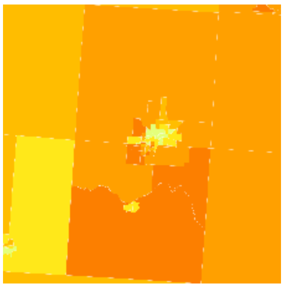
(a) MSA 120



(b) MSA 220

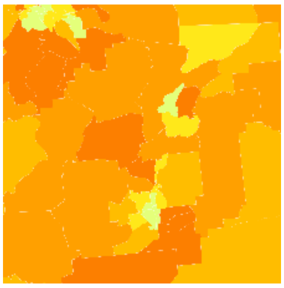


(c) MSA 240



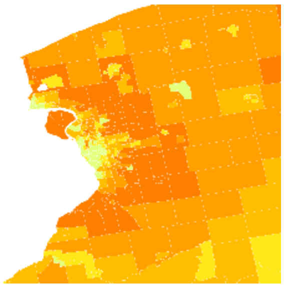
(d) MSA 320

(e) MSA 40

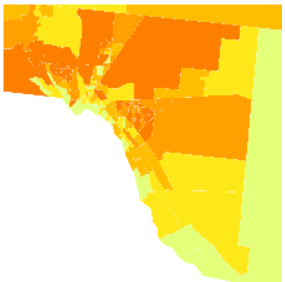


(f) MSA 450

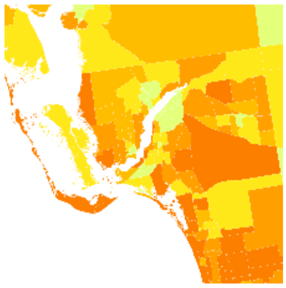
Figure 4: Cluster 2



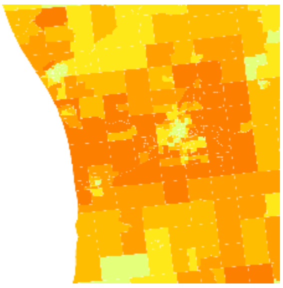
(a) MSA 1280



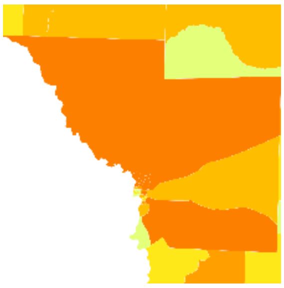
(b) MSA 2320



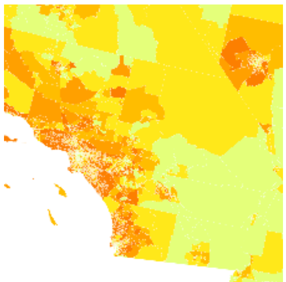
(c) MSA 2700



(d) MSA 3000

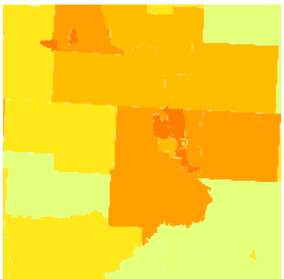


(e) MSA 4080

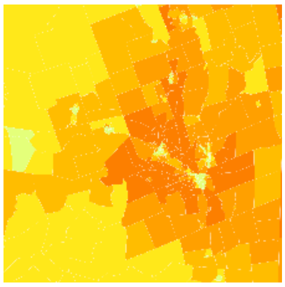


(f) MSA 4472

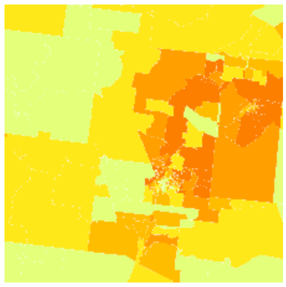
Figure 6: Cluster 3



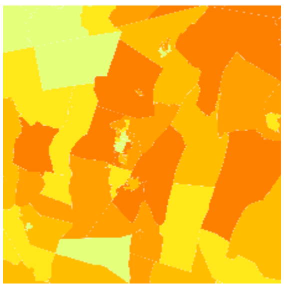
(a) MSA 1010



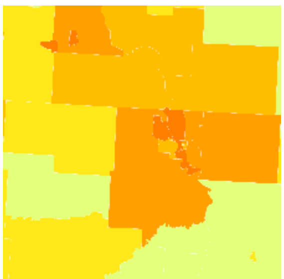
(b) MSA 160



(c) MSA 200



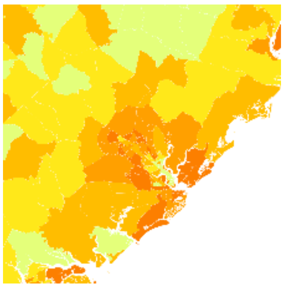
(d) MSA 280

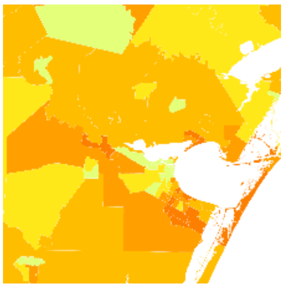


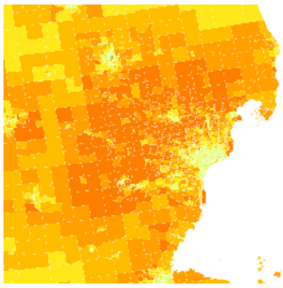
(e) MSA 1010

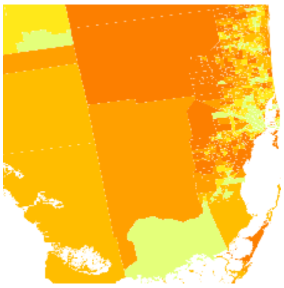
(f) MSA 1080

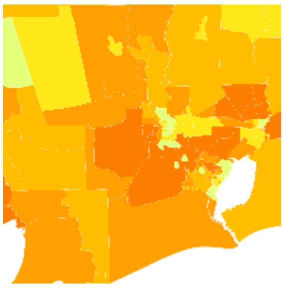
Figure 8: Cluster 4


(a) MSA 1440


(b) MSA 1880


(c) MSA 2162


(d) MSA 4992


(e) MSA 840

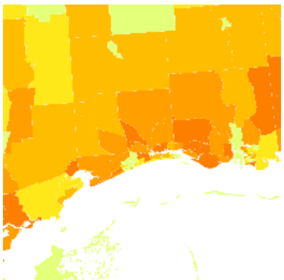
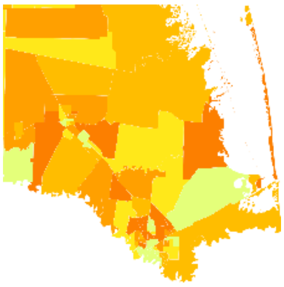
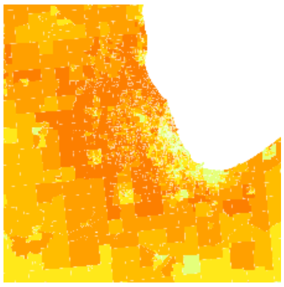
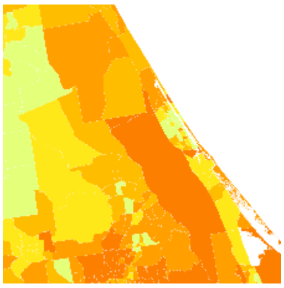

(f) MSA 920

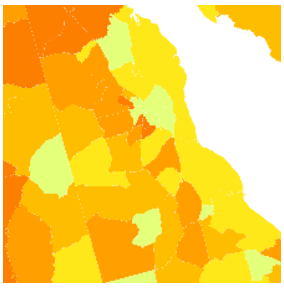
Figure 10: Cluster 5


(a) MSA 1122


(b) MSA 1240


(c) MSA 1602


(d) MSA 2020


(e) MSA 2190

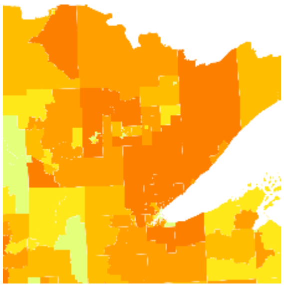
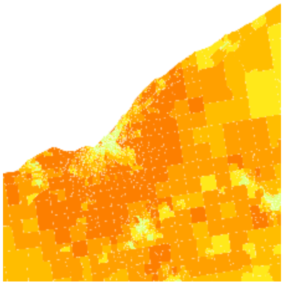
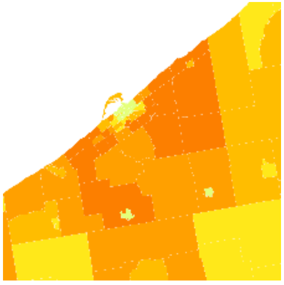

(f) MSA 2240

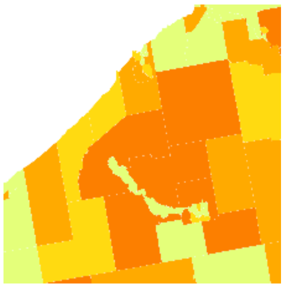
Figure 12: Cluster 6



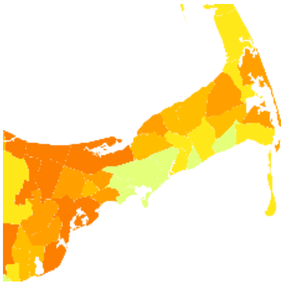
(a) MSA 1692



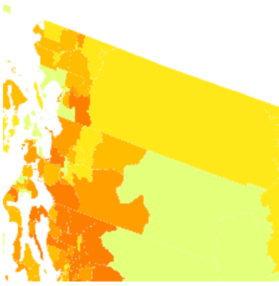
(b) MSA 2360



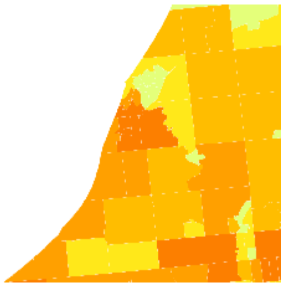
(c) MSA 3610



(d) MSA 740



(e) MSA 860



(f) MSA 870

D. Clustering algorithm results look-up table

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	34.375	Macon, GA	4680
1	34.375	Macon, GA	4680
1	34.375	Macon, GA	4680
1	34.391	Albany, GA	120
1	34.391	Albany, GA	120
1	34.96575	Athens-Clarke County, GA	500
1	34.96575	Athens-Clarke County, GA	500
1	34.96575	Athens-Clarke County, GA	500
1	35.17575	Greenville, NC	3150
1	35.4455	Columbus, GA-AL	1800
1	35.4455	Columbus, GA-AL	1800
1	35.4455	Columbus, GA-AL	1800
1	35.4455	Columbus, GA-AL	1800
1	35.63825	Charlotte-Gastonia-Concord, NC-SC	1520
1	35.63825	Charlotte-Gastonia-Concord, NC-SC	1520
1	35.63825	Charlotte-Gastonia-Concord, NC-SC	1520
1	35.63825	Charlotte-Gastonia-Concord, NC-SC	1520
1	35.63825	Charlotte-Gastonia-Concord, NC-SC	1520
1	35.63825	Charlotte-Gastonia-Concord, NC-SC	1520
1	36.0965	Jackson, MS	3560
1	36.0965	Jackson, MS	3560

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	36.0965	Jackson, MS	3560
1	36.1245	Lexington-Fayette, KY	4280
1	36.1245	Lexington-Fayette, KY	4280
1	36.1245	Lexington-Fayette, KY	4280
1	36.1245	Lexington-Fayette, KY	4280
1	36.1245	Lexington-Fayette, KY	4280
1	36.1245	Lexington-Fayette, KY	4280
1	36.2365	Winston-Salem, NC	3120
1	36.2365	Winston-Salem, NC	3120
1	36.2365	Winston-Salem, NC	3120
1	36.2365	Winston-Salem, NC	3120
1	36.32725	Greensboro-High Point, NC	3120
1	36.32725	Greensboro-High Point, NC	3120
1	36.33225	Pine Bluff, AR	6240
1	36.36425	Sumter, SC	8140
1	36.52225	Burlington, NC	3120
1	36.5415	Springfield, OH	2000
1	36.66875	Auburn-Opelika, AL	580
1	36.6885	Spartanburg, SC	3160
1	36.70825	Muncie, IN	5280
1	36.80275	Tallahassee, FL	8240
1	36.80275	Tallahassee, FL	8240
1	36.8625	Battle Creek, MI	3720

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	36.889	Muskegon-Norton Shores, MI a	3000
1	37.00675	Jackson, TN	3580
1	37.00675	Jackson, TN	3580
1	37.14625	Greenville, SC	3160
1	37.14625	Greenville, SC	3160
1	37.159	Dayton, OH	2000
1	37.159	Dayton, OH	2000
1	37.159	Dayton, OH	2000
1	37.243	Indianapolis, IN	3480
1	37.243	Indianapolis, IN	3480
1	37.243	Indianapolis, IN	3480
1	37.243	Indianapolis, IN	3480
1	37.243	Indianapolis, IN	3480
1	37.243	Indianapolis, IN	3480
1	37.243	Indianapolis, IN	3480
1	37.243	Indianapolis, IN	3480
1	37.267	Kokomo, IN	3850
1	37.267	Kokomo, IN	3850
1	37.28175	Tuscaloosa, AL	8600
1	37.30325	Warner Robins, GA	4680
1	37.37725	Monroe, LA	5200
1	37.50075	Huntsville, AL	3440
1	37.50075	Huntsville, AL	3440

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	37.5595	Birmingham-Hoover, AL	1000
1	37.5595	Birmingham-Hoover, AL	1000
1	37.5595	Birmingham-Hoover, AL	1000
1	37.5595	Birmingham-Hoover, AL	1000
1	37.56675	Goldsboro, NC	2980
1	37.574	Columbus, OH	1840
1	37.574	Columbus, OH	1840
1	37.574	Columbus, OH	1840
1	37.574	Columbus, OH	1840
1	37.574	Columbus, OH	1840
1	37.574	Columbus, OH	1840
1	37.78225	Shreveport-Bossier City, LA	7680
1	37.78225	Shreveport-Bossier City, LA	7680
1	37.806	Louisville, KY-IN	4520
1	37.806	Louisville, KY-IN	4520
1	37.806	Louisville, KY-IN	4520
1	37.806	Louisville, KY-IN	4520
1	37.806	Louisville, KY-IN	4520
1	37.806	Louisville, KY-IN	4520
1	37.91375	Knoxville, TN	3840
1	37.91375	Knoxville, TN	3840
1	37.91375	Knoxville, TN	3840
1	37.91375	Knoxville, TN	3840

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	37.91375	Knoxville, TN	3840
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.9255	Cincinnati-Middletown, OH-KY-IN	1640
1	37.96325	Toledo, OH	8400
1	37.96325	Toledo, OH	8400
1	37.96325	Toledo, OH	8400
1	37.96775	Kalamazoo-Portage, MI	3720
1	37.96775	Kalamazoo-Portage, MI	3720
1	38.121	Dothan, AL	2180
1	38.23525	Nashville-Davidson–Murfreesboro, TN	5360
1	38.23525	Nashville-Davidson–Murfreesboro, TN	5360
1	38.23525	Nashville-Davidson–Murfreesboro, TN	5360
1	38.23525	Nashville-Davidson–Murfreesboro, TN	5360

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	38.23525	Nashville-Davidson–Murfreesboro, TN	5360
1	38.23525	Nashville-Davidson–Murfreesboro, TN	5360
1	38.23525	Nashville-Davidson–Murfreesboro, TN	5360
1	38.23525	Nashville-Davidson–Murfreesboro, TN	5360
1	38.27075	Ocala, FL	5790
1	38.28575	Anniston-Oxford, AL	450
1	38.39675	Gainesville, FL	2900
1	38.4375	Rockford, IL	6880
1	38.4375	Rockford, IL	6880
1	38.448	Bloomington, IN	1020
1	38.50075	Milwaukee-Waukesha-West Allis, WI	5080
1	38.50075	Milwaukee-Waukesha-West Allis, WI	5080
1	38.50075	Milwaukee-Waukesha-West Allis, WI	5080
1	38.50075	Milwaukee-Waukesha-West Allis, WI	5080
1	38.51175	Clarksville, TN-KY	1660
1	38.51175	Clarksville, TN-KY	1660
1	38.751	South Bend-Mishawaka, IN-MI	7800
1	39.0105	Charlottesville, VA	1540
1	39.0105	Charlottesville, VA	1540
1	39.0105	Charlottesville, VA	1540
1	39.0105	Charlottesville, VA	1540
1	39.024	Jackson, MI	3520
1	39.085	Waco, TX	8800

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	39.233	Pueblo, CO	6560
1	39.2425	Lynchburg, VA	4640
1	39.2425	Lynchburg, VA	4640
1	39.2425	Lynchburg, VA	4640
1	39.2425	Lynchburg, VA	4640
1	39.2425	Lynchburg, VA	4640
1	39.278	Lima, OH	4320
1	39.319	Baton Rouge, LA	760
1	39.319	Baton Rouge, LA	760
1	39.319	Baton Rouge, LA	760
1	39.319	Baton Rouge, LA	760
1	39.3355	Gadsden, AL	2880
1	39.42625	Mansfield, OH	4800
1	39.50525	Lansing-East Lansing, MI	4040
1	39.50525	Lansing-East Lansing, MI	4040
1	39.50525	Lansing-East Lansing, MI	4040
1	39.52325	Parkersburg-Marietta, WV-OH	6020
1	39.52325	Parkersburg-Marietta, WV-OH	6020
1	39.52975	Decatur, IL	2040
1	39.53675	Alexandria, LA	220
1	39.59475	Fort Wayne, IN	2760
1	39.59475	Fort Wayne, IN	2760
1	39.59475	Fort Wayne, IN	2760

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	39.6135	Roanoke, VA	6800
1	39.6135	Roanoke, VA	6800
1	39.6135	Roanoke, VA	6800
1	39.6135	Roanoke, VA	6800
1	39.874	Springfield, IL	7880
1	39.874	Springfield, IL	7880
1	39.898	Columbia, MO	1740
1	40.02325	Austin-Round Rock, TX	640
1	40.02325	Austin-Round Rock, TX	640
1	40.02325	Austin-Round Rock, TX	640
1	40.02325	Austin-Round Rock, TX	640
1	40.02325	Austin-Round Rock, TX	640
1	40.0345	Canton-Massillon, OH	1320
1	40.0345	Canton-Massillon, OH	1320
1	40.07625	Jonesboro, AR	3700
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760
1	40.16775	Kansas City, MO-KS	3760
1	40.3955	Grand Rapids-Wyoming, MI	3000
1	40.469	Youngstown-Warren-Boardman, OH-PA	9320
1	40.469	Youngstown-Warren-Boardman, OH-PA	7610
1	40.469	Youngstown-Warren-Boardman, OH-PA	9320
1	40.47525	Santa Fe, NM	7490
1	40.48625	Killeen-Temple-Fort Hood, TX	3810
1	40.48625	Killeen-Temple-Fort Hood, TX	3810
1	40.511	Evansville, IN-KY	2440
1	40.511	Evansville, IN-KY	2440
1	40.511	Evansville, IN-KY	2440
1	40.511	Evansville, IN-KY	2440
1	40.51425	Lubbock, TX	4600
1	40.643	Lawrence, KS	4150
1	40.75475	Fresno, CA	2840
1	40.788	Florence-Muscle Shoals, AL	2650
1	40.788	Florence-Muscle Shoals, AL	2650
1	40.95525	Medford, OR	4890
1	41.03725	Weirton-Steubenville, WV-OH	8080
1	41.03725	Weirton-Steubenville, WV-OH	8080
1	41.03725	Weirton-Steubenville, WV-OH	8080

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	41.038	Elkhart-Goshen, IN	2330
1	41.057	San Antonio, TX	7240
1	41.057	San Antonio, TX	7240
1	41.057	San Antonio, TX	7240
1	41.057	San Antonio, TX	7240
1	41.18525	Peoria, IL	6120
1	41.18525	Peoria, IL	6120
1	41.18525	Peoria, IL	6120
1	41.224	Janesville, WI	3620
1	41.298	Tyler, TX	8640
1	41.30275	Chico, CA	1620
1	41.3385	Sherman-Denison, TX	7640
1	41.39325	Colorado Springs, CO	1720
1	41.47075	Lake Charles, LA	3960
1	41.588	Wichita, KS	9040
1	41.588	Wichita, KS	9040
1	41.588	Wichita, KS	9040
1	41.7265	Denver-Aurora, CO	2080
1	41.7265	Denver-Aurora, CO	2080
1	41.7265	Denver-Aurora, CO	2080
1	41.7265	Denver-Aurora, CO	2080
1	41.7265	Denver-Aurora, CO	2080
1	41.7265	Denver-Aurora, CO	2080

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	41.72725	Madera, CA	2840
1	41.74425	College Station-Bryan, TX	1260
1	41.76675	Champaign-Urbana, IL	1400
1	41.8365	Springfield, MO	7920
1	41.8365	Springfield, MO	7920
1	41.8365	Springfield, MO	7920
1	41.84975	Terre Haute, IN	8320
1	41.84975	Terre Haute, IN	8320
1	41.84975	Terre Haute, IN	8320
1	41.943	Redding, CA	6690
1	42.01275	Merced, CA	4940
1	42.08725	Corvallis, OR	1890
1	42.15875	Spokane, WA	7840
1	42.1765	Lawton, OK	4200
1	42.183	Modesto, CA	5170
1	42.32275	Oklahoma City, OK	5880
1	42.32275	Oklahoma City, OK	5880
1	42.32275	Oklahoma City, OK	5880
1	42.32275	Oklahoma City, OK	5880
1	42.32275	Oklahoma City, OK	5880
1	42.37175	St. Joseph, MO-KS	7000
1	42.37175	St. Joseph, MO-KS	7000
1	42.60775	Binghamton, NY	960

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	42.60775	Binghamton, NY	960
1	42.6105	Stockton, CA	8120
1	42.6105	Topeka, KS	8440
1	42.6385	Joplin, MO	3710
1	42.6385	Joplin, MO	3710
1	42.66225	Omaha-Council Bluffs, NE-IA	5920
1	42.66225	Omaha-Council Bluffs, NE-IA	5920
1	42.66225	Omaha-Council Bluffs, NE-IA	5920
1	42.66225	Omaha-Council Bluffs, NE-IA	5920
1	42.66225	Omaha-Council Bluffs, NE-IA	5920
1	42.719	Harrisburg-Carlisle, PA	3240
1	42.719	Harrisburg-Carlisle, PA	3240
1	42.719	Harrisburg-Carlisle, PA	3240
1	42.87325	Yakima, WA	9260
1	43.00875	Owensboro, KY	5990
1	43.09025	San Angelo, TX	7200
1	43.1885	Amarillo, TX	320
1	43.1885	Amarillo, TX	320
1	43.42325	Cumberland, MD-WV	1900
1	43.42325	Cumberland, MD-WV	1900
1	43.45275	Lincoln, NE	4360
1	43.47225	Davenport-Moline-Rock Island, IA-IL	1960
1	43.47225	Davenport-Moline-Rock Island, IA-IL	1960

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	43.47225	Davenport-Moline-Rock Island, IA-IL	1960
1	43.6315	Victoria, TX	8750
1	43.79425	U a-Rome, NY	8680
1	43.79425	U a-Rome, NY	8680
1	43.85625	Wichita Falls, TX	9080
1	43.85625	Wichita Falls, TX	9080
1	43.93	Longview, TX	4420
1	43.93	Longview, TX	4420
1	44.0005	Missoula, MT	5140
1	44.17225	York-Hanover, PA	9280
1	44.295	Reading, PA	6680
1	44.303	Rapid City, SD	6660
1	44.35	Des Moines, IA	2120
1	44.35	Des Moines, IA	2120
1	44.35	Des Moines, IA	2120
1	44.54025	Madison, WI	4720
1	44.54275	Oshkosh-Neenah, WI	460
1	44.6695	Abilene, TX	40
1	44.68775	Lancaster, PA	4000
1	44.70975	Kennewick-Richland-Pasco, WA	6740
1	44.70975	Kennewick-Richland-Pasco, WA	6740
1	44.79475	Pittsburgh, PA	6280
1	44.79475	Pittsburgh, PA	6280

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	44.79475	Pittsburgh, PA	6280
1	44.79475	Pittsburgh, PA	6280
1	44.79475	Pittsburgh, PA	6280
1	44.79475	Pittsburgh, PA	6280
1	44.87	Wheeling, WV-OH	9000
1	44.87	Wheeling, WV-OH	9000
1	44.87	Wheeling, WV-OH	9000
1	45.106	Allentown-Bethlehem-Easton, PA-NJ	240
1	45.106	Allentown-Bethlehem-Easton, PA-NJ	240
1	45.106	Allentown-Bethlehem-Easton, PA-NJ	240
1	45.17525	Cheyenne, WY	1580
1	45.38525	Fort Collins-Loveland, CO	2670
1	45.423	Williamsport, PA	9140
1	45.756	Salt Lake City, UT	7160
1	45.79075	Great Falls, MT	3040
1	45.946	Scranton–Wilkes-Barre, PA	7560
1	45.946	Scranton–Wilkes-Barre, PA	7560
1	45.946	Scranton–Wilkes-Barre, PA	7560
1	46.32125	Sheboygan, WI	7620
1	46.41475	Waterloo-Cedar Falls, IA	8920
1	46.46225	Ogden-Clearfield, UT	7160
1	46.46225	Ogden-Clearfield, UT	7160
1	46.473	Iowa City, IA	3500

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
1	46.66625	Lebanon, PA	3240
1	46.69825	Green Bay, WI	3080
1	46.88675	Cedar Rapids, IA	1360
1	47.0925	La Crosse, WI-MN	3870
1	47.0925	La Crosse, WI-MN	3870
1	47.26875	Eau Claire, WI	2290
1	47.26875	Eau Claire, WI	2290
1	47.322	State College, PA	8050
1	48.73425	Johnstown, PA	3680
1	49.023	Appleton, WI	460
1	49.023	Appleton, WI	460
1	49.279	Fargo, ND-MN	2520
1	49.279	Fargo, ND-MN	2520
1	49.4555	Casper, WY	1350
1	49.4665	Midland, TX	5800
1	49.87475	Wausau, WI	8940
1	50.1655	Rochester, MN	6820
1	50.79175	Odessa, TX	5800
1	52.68125	Dubuque, IA	2200
1	52.77525	St. Cloud, MN	6980
1	52.77525	St. Cloud, MN	6980
2	38.98425	Tucson, AZ	8520
2	39.1325	Tampa-St. Petersburg-Clearwater, FL	8280

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
2	39.1325	Tampa-St. Petersburg-Clearwater, FL	8280
2	39.1325	Tampa-St. Petersburg-Clearwater, FL	8280
2	39.1325	Tampa-St. Petersburg-Clearwater, FL	8280
2	39.1355	Cape Coral-Fort Myers, FL	2700
2	41.853	Punta Gorda, FL	6580
2	42.35375	Buffalo-Niagara Falls, NY	1280
2	42.35375	Buffalo-Niagara Falls, NY	1280
2	43.621	Salinas, CA	7120
2	43.98675	Yuma, AZ	9360
2	44.081	McAllen-Edinburg-Pharr, TX	4880
2	44.27025	San Diego-Carlsbad-San Marcos, CA	7320
2	44.27175	El Paso, TX	2320
2	44.6555	Laredo, TX	4080
2	45.1465	San Luis Obispo-Paso Robles, CA	7460
3	33.72775	Memphis, TN-MS-AR	4920
3	33.72775	Memphis, TN-MS-AR	4920
3	33.72775	Memphis, TN-MS-AR	4920
3	33.72775	Memphis, TN-MS-AR	4920
3	33.72775	Memphis, TN-MS-AR	4920
3	35.18625	Durham, NC	6640
3	35.18625	Durham, NC	6640
3	35.18625	Durham, NC	6640
3	35.6075	Fayetteville, NC	2560

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
3	35.68375	Montgomery, AL	5240
3	35.68375	Montgomery, AL	5240
3	35.68375	Montgomery, AL	5240
3	35.829	Hickory-Lenoir-Morganton, NC	3290
3	35.829	Hickory-Lenoir-Morganton, NC	3290
3	35.829	Hickory-Lenoir-Morganton, NC	3290
3	35.829	Hickory-Lenoir-Morganton, NC	3290
3	36.06575	Savannah, GA	7520
3	36.06575	Savannah, GA	7520
3	36.06575	Savannah, GA	7520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.14175	Atlanta-Sandy Springs-Marietta, GA	520
3	36.1915	Columbia, SC	1760
3	36.1915	Columbia, SC	1760
3	36.415	Augusta-Richmond County, GA-SC	600
3	36.415	Augusta-Richmond County, GA-SC	600
3	36.415	Augusta-Richmond County, GA-SC	600
3	36.415	Augusta-Richmond County, GA-SC	600
3	36.415	Augusta-Richmond County, GA-SC	600
3	37.3975	Mobile, AL	5160
3	37.4455	Johnson City, TN	3660
3	37.4455	Johnson City, TN	3660
3	37.4455	Johnson City, TN	3660
3	37.806	Little Rock-North Little Rock, AR	4400
3	37.806	Little Rock-North Little Rock, AR	4400
3	37.806	Little Rock-North Little Rock, AR	4400
3	37.806	Little Rock-North Little Rock, AR	4400
3	37.84425	Richmond, VA	6760

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.84425	Richmond, VA	6760
3	37.9435	Chattanooga, TN-GA	1560
3	37.9435	Chattanooga, TN-GA	1560
3	37.9435	Chattanooga, TN-GA	1560
3	37.9435	Chattanooga, TN-GA	1560
3	37.9435	Chattanooga, TN-GA	1560
3	38.028	Raleigh-Cary, NC	6640
3	38.028	Raleigh-Cary, NC	6640
3	38.028	Raleigh-Cary, NC	6640
3	38.73275	Kingsport-Bristol-Bristol, TN-VA	3660
3	38.73275	Kingsport-Bristol-Bristol, TN-VA	3660
3	38.73275	Kingsport-Bristol-Bristol, TN-VA	3660

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
3	38.73275	Kingsport-Bristol-Bristol, TN-VA	3660
3	38.73275	Kingsport-Bristol-Bristol, TN-VA	3660
3	38.97325	Asheville, NC	480
3	38.97325	Asheville, NC	480
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.034	St. Louis, MO-IL	7040
3	39.101	Texarkana, TX-Texarkana, AR	8360
3	39.101	Texarkana, TX-Texarkana, AR	8360
3	39.16275	Orlando, FL	5960
3	39.16275	Orlando, FL	5960
3	39.16275	Orlando, FL	5960
3	39.16275	Orlando, FL	5960
3	39.23175	Decatur, AL	2030

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
3	39.23175	Decatur, AL	2030
3	39.33825	Lakeland, FL	3980
3	39.41275	Albuquerque, NM	200
3	39.41275	Albuquerque, NM	200
3	39.41275	Albuquerque, NM	200
3	39.52625	Flagstaff, AZ	2620
3	39.7105	Eugene-Springfield, OR	2400
3	39.873	Las Vegas-Paradise, NV	4120
3	40.353	Phoenix-Mesa-Scottsdale, AZ	6200
3	40.353	Phoenix-Mesa-Scottsdale, AZ	6200
3	40.3675	Huntington-Ashland, WV-KY-OH	3400
3	40.3675	Huntington-Ashland, WV-KY-OH	3400
3	40.3675	Huntington-Ashland, WV-KY-OH	3400
3	40.3675	Huntington-Ashland, WV-KY-OH	3400
3	40.3675	Huntington-Ashland, WV-KY-OH	3400
3	40.95225	Rochester, NY	6840
3	40.95225	Rochester, NY	6840
3	40.95225	Rochester, NY	6840
3	40.95225	Rochester, NY	6840
3	40.95225	Rochester, NY	6840
3	41.0555	Reno-Sparks, NV	6720
3	41.06025	Charleston, WV	1480
3	41.06025	Charleston, WV	1480

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
3	41.29275	Tulsa, OK	8560
3	41.29275	Tulsa, OK	8560
3	41.29275	Tulsa, OK	8560
3	41.29275	Tulsa, OK	8560
3	41.29275	Tulsa, OK	8560
3	41.297	Portland-Vancouver-Beaverton, OR-WA	6440
3	41.297	Portland-Vancouver-Beaverton, OR-WA	6440
3	41.297	Portland-Vancouver-Beaverton, OR-WA	6440
3	41.297	Portland-Vancouver-Beaverton, OR-WA	6440
3	41.297	Portland-Vancouver-Beaverton, OR-WA	6440
3	41.297	Portland-Vancouver-Beaverton, OR-WA	6440
3	41.2995	Las Cruces, NM	4100
3	41.32925	Syracuse, NY	8160
3	41.32925	Syracuse, NY	8160
3	41.32925	Syracuse, NY	8160
3	42.12775	Fort Smith, AR-OK	2720
3	42.12775	Fort Smith, AR-OK	2720
3	42.12775	Fort Smith, AR-OK	2720
3	42.1335	Visalia-Porterville, CA	8780
3	42.54775	Boise City-Nampa, ID	1080
3	42.54775	Boise City-Nampa, ID	1080
3	42.56075	Sacramento–Arden-Arcade–Roseville, CA	6920
3	42.56075	Sacramento–Arden-Arcade–Roseville, CA	6920

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
3	42.56075	Sacramento–Arden-Arcade–Roseville, CA	6920
3	42.96975	Lafayette, LA	3880
3	42.96975	Lafayette, LA	3880
3	43.09725	Albany-Schenectady-Troy, NY	160
3	43.09725	Albany-Schenectady-Troy, NY	160
3	43.09725	Albany-Schenectady-Troy, NY	160
3	43.09725	Albany-Schenectady-Troy, NY	160
3	43.09725	Albany-Schenectady-Troy, NY	160
3	43.107	Pocatello, ID	6340
3	43.22025	Fayetteville-Springdale-Rogers, AR-MO	2580
3	43.22025	Fayetteville-Springdale-Rogers, AR-MO	2580
3	43.634	Yuba City, CA	9340
3	43.634	Yuba City, CA	9340
3	43.86725	Altoona, PA	280
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	44.562	Minneapolis-St. Paul-Bloomington, MN-WI	5120
3	45.0225	Anchorage, AK	380
3	45.296	Bakersfield, CA	680
3	47.24225	Billings, MT	880
3	48.01025	Sioux City, IA-NE-SD	7720
3	48.01025	Sioux City, IA-NE-SD	7720
3	49.16025	Provo-Orem, UT	6520
3	49.417	Sioux Falls, SD	7760
3	49.417	Sioux Falls, SD	7760
3	51.0815	Bismarck, ND	1010
3	51.0815	Bismarck, ND	1010
4	35.91025	Wilmington, NC	9200
4	35.91025	Wilmington, NC	9200
4	36.54125	Myrtle Beach-Conway-North Myrtle Beach,	5330
4	37.83625	Charleston-North Charleston, SC	1440
4	37.83625	Charleston-North Charleston, SC	1440
4	37.83625	Charleston-North Charleston, SC	1440
4	38.198	New Orleans-Metairie-Kenner, LA	5560
4	38.198	New Orleans-Metairie-Kenner, LA	5560
4	38.198	New Orleans-Metairie-Kenner, LA	5560

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
4	38.198	New Orleans-Metairie-Kenner, LA	5560
4	38.198	New Orleans-Metairie-Kenner, LA	5560
4	38.198	New Orleans-Metairie-Kenner, LA	5560
4	38.198	New Orleans-Metairie-Kenner, LA	5560
4	38.4405	Gulfport-Biloxi, MS	920
4	38.4405	Gulfport-Biloxi, MS	920
4	39.367	Pensacola-Ferry Pass-Brent, FL	6080
4	39.367	Pensacola-Ferry Pass-Brent, FL	6080
4	42.67725	Monroe, MI	2160
4	42.8995	Houston-Baytown-Sugar Land, TX	3360
4	42.8995	Houston-Baytown-Sugar Land, TX	3360
4	42.8995	Houston-Baytown-Sugar Land, TX	3360
4	42.8995	Houston-Baytown-Sugar Land, TX	3360
4	42.8995	Houston-Baytown-Sugar Land, TX	3360
4	42.8995	Houston-Baytown-Sugar Land, TX	3360
4	42.94725	Beaumont-Port Arthur, TX	840
4	42.94725	Beaumont-Port Arthur, TX	840
4	42.94725	Beaumont-Port Arthur, TX	840
4	43.027	Corpus Christi, TX	1880
4	43.027	Corpus Christi, TX	1880
4	44.98875	Houma-Bayou Cane-Thibodaux, LA	3350
4	44.98875	Houma-Bayou Cane-Thibodaux, LA	3350
5	37.48225	Jacksonville, FL	3600

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
5	37.48225	Jacksonville, FL	3600
5	37.48225	Jacksonville, FL	3600
5	37.48225	Jacksonville, FL	3600
5	37.53375	Saginaw-Saginaw Township North, MI	6960
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	37.98675	Virginia Beach-Norfolk-Newport News, VA	5720
5	38.0405	Port St. Lucie-Fort Pierce, FL	2710
5	38.0405	Port St. Lucie-Fort Pierce, FL	2710
5	39.53325	Palm Bay-Melbourne-Titusville, FL	4900
5	39.661	Deltona-Daytona Beach-Ormond Beach, FL	2020

Table 2: MSA Look-up Table

Cluster	Absolute upward mobility	MSA Name	MSA Code
5	39.66275	Dover, DE	2190
5	41.16625	Bay City, MI	6960
5	43.521	Brownsville-Harlingen, TX	1240
5	44.8485	Duluth, MN-WI	2240
5	44.8485	Duluth, MN-WI	2240
6	38.45475	Niles-Benton Harbor, MI	870
6	42.23525	Erie, PA	2360
6	44.12275	Bellingham, WA	860